

**FUNDAMENTALS OF DIFFUSION-BASED  
MOLECULAR COMMUNICATION  
IN NANONETWORKS**

A Dissertation  
Presented to  
The Academic Faculty

By

Massimiliano Pierobon

In Partial Fulfillment  
of the Requirements for the Degree  
Doctor of Philosophy  
in  
Electrical and Computer Engineering



School of Electrical and Computer Engineering  
Georgia Institute of Technology  
December 2013

Copyright © 2013 by Massimiliano Pierobon

**FUNDAMENTALS OF DIFFUSION-BASED  
MOLECULAR COMMUNICATION  
IN NANONETWORKS**

Approved by:

Dr. Ian F. Akyildiz, Advisor  
*Ken Byers Chair Professor in Telecommunications, School of ECE*  
*Georgia Institute of Technology*

Dr. Craig R. Forest  
*Assistant Professor, School of Mechanical Engineering*  
*Georgia Institute of Technology*

Dr. Faramarz Fekri  
*Professor, School of ECE*  
*Georgia Institute of Technology*

Dr. Raghupathy Sivakumar  
*Professor, School of ECE*  
*Georgia Institute of Technology*

Dr. Albert B. Frazier  
*Professor, School of ECE*  
*Georgia Institute of Technology*

Date Approved: August 7, 2013

*To my father, for his faith in my capabilities, his life lessons, and his handed-on experiences, and for his strength and patience while fighting the threatens to his health.*

*To my mother, for her endless care and love for her family, and for her continuous support over my many weaknesses, where she exercised great and endless patience.*

*To them, who gave me the greatest love, support, and courage I could ever ask over this difficult choice of being a free thinker, and coming to the New World to pursue my happiness.*

*For they never hesitated any single moment to give me the strength to leave my motherland and start anew, elsewhere, far from their arms.*

*To them I dedicate this work and all my future endeavors.*

## ACKNOWLEDGMENTS

I would like to express my heartiest thanks to Prof. Ian F. Akyildiz, who has been not only my Ph.D. advisor, but also a mentor and a second father for me, from the very moment I came to Atlanta with my heart full of hope and my mind as wild as a maniac's mirth. He truly shaped my will and my strength, he showed me the right path, he taught me how to focus, concentrate, believe in my potential, and he made me achieve what I would have never thought. The road to his Olympus is steep, and full of ordeals, but there he was waiting for me, at the end of the trail, sure of my final success. Now I am changed, I feel stronger, more confident of who I am, what I want, and where I am going, and this is primarily thanks to Prof. Ian F. Akyildiz, to whom I owe a lot.

My heartiest thanks go also to my travel buddy, Josep Miquel, with whom I shared these years of intense work, and even more intense friendship. Together we walked the trail in the BWN Lab, and we also shared most of our "limited" free time, whether it was for cooking our meals, for praying our God, or for exchanging our ideas. Through the years we became like brothers, and I feel the physical distance of our future commitments will not untangle the strong bound we built.

My sincere and warm thanks go also to the whole BWN Lab family, to past and current members, Ph.D. students, visitors, and Post-Docs, with whom I shared every weekday of my life in these years. Amazing personalities, sincere friendships and open hearts were the true gifts I received from them. People I will never forget in my entire life, whose origins span the worldwide map from North to South, from East to West. The meeting of all these culture and the vivid relationships among us all were a true consolation for me, and many of the bounds we built will surely last forever.

Very special and warm thanks go also to the BWN Lab "extension" in UPC, Barcelona, with whom I spent a part of my summer Ph.D. studies, and whose warmth and friendship made me really feel at home. I will never forget the time spent in that amazing city with

even more amazing people, and surely I will work to make the summertime with N3Cat a reality again in the near future.

My sincere and special thanks go also to the people I collaborated with while working on the MoNaCo project. I had the pleasure to work with amazing faculty, Post-Docs and students with diverse backgrounds, all willing to collaborate for a common research goal. I consider many of them as true friends, and I hope our paths will join again in the future. I feel truly honored to have been part of this team and to have had the opportunity to salute the success of the first two years of this project.

Very very special thanks go to the department staff in the Centergy One building, in particular to Pat Dixon, who is a true friend for me, and with whom I shared so much in these years. To her I owe the many things I learnt of economic research project management, light-hearted attitude and positivity management, self-esteem management, and "vegetable management". Particular thanks go also to Cordai Farrar, whose smiles, gentle attitude and professionalism were true bedrocks for me from the very first day of my stay in the lab.

Finally, I would like to extend my sincere thanks to my Ph.D. committee members, who accepted to serve as guarantors of my graduation, and whose constrictive criticism and invaluable advices helped me take the final leap towards the completion of this life-changing experience.

# TABLE OF CONTENTS

<b>ACKNOWLEDGMENTS</b> . . . . .	iv
<b>LIST OF FIGURES</b> . . . . .	ix
<b>SUMMARY</b> . . . . .	1
<b>CHAPTER 1 INTRODUCTION</b> . . . . .	2
1.1 Background . . . . .	2
1.2 Potential Applications . . . . .	4
1.3 Research Objectives and Solutions . . . . .	6
1.4 Thesis Outline . . . . .	10
<b>CHAPTER 2 PREVIOUS WORK ON MOLECULAR COMMUNICATION</b> .	11
2.1 Molecular Communication Types . . . . .	11
2.2 Diffusion-based Molecular Communication Architectures . . . . .	13
<b>CHAPTER 3 PHYSICAL END-TO-END MODEL OF DIFFUSION-BASED MOLECULAR COMMUNICATION</b> . . . . .	16
3.1 Motivation and Related Work . . . . .	16
3.2 A Basic Design of a Diffusion-based MC System . . . . .	17
3.3 The Emission Process . . . . .	18
3.4 The Diffusion Process . . . . .	23
3.5 The Reception Process . . . . .	27
3.6 End-to-end Normalized Gain and Delay . . . . .	31
3.7 Numerical Results . . . . .	32
3.8 Conclusion . . . . .	37
<b>CHAPTER 4 NOISES IN DIFFUSION-BASED MOLECULAR COMMUNICATION</b> . . . . .	38
4.1 Motivation and Related Work . . . . .	38
4.2 Definition of the Noise Sources in Diffusion-based MC . . . . .	40
4.3 The Particle Sampling Noise . . . . .	41
4.3.1 The Physical Model . . . . .	42
4.3.2 The Stochastic Model . . . . .	46
4.3.3 Numerical Results . . . . .	51
4.4 The Particle Counting Noise . . . . .	54
4.4.1 The Physical Model . . . . .	55
4.4.2 The Stochastic Model . . . . .	58
4.4.3 Numerical Results . . . . .	64
4.5 The Ligand-receptor-binding Noise . . . . .	67
4.5.1 The Physical Model . . . . .	69
4.5.2 The Stochastic Model . . . . .	74

4.5.3	Numerical Results . . . . .	82
4.6	Conclusion . . . . .	91
<b>CHAPTER 5</b>	<b>CAPACITY ANALYSIS OF DIFFUSION-BASED</b>	
	<b>MOLECULAR COMMUNICATION . . . . .</b>	<b>93</b>
5.1	Motivation and Related Work . . . . .	93
5.2	Capacity Analysis through Thermodynamics . . . . .	97
5.2.1	Information-theoretic Definition of Capacity . . . . .	100
5.2.2	Information-theoretic Entropy from Thermodynamic Entropy . . . . .	102
5.2.3	The Input-signal Entropy . . . . .	104
5.2.4	The Equivocation . . . . .	106
5.2.5	The Capacity . . . . .	109
5.2.6	Numerical Results . . . . .	112
5.2.7	Conclusion . . . . .	117
5.3	Capacity Analysis with Channel Memory and Molecular Noise . . . . .	117
5.3.1	Information Capacity of a Diffusion-based MC System . . . . .	121
5.3.2	The Molecule Diffusion as Fick's Diffusion and Particle Location Displacement . . . . .	122
5.3.3	The Fick's Diffusion Mutual Information . . . . .	126
5.3.4	The Particle Location Displacement Mutual Information . . . . .	129
5.3.5	The Capacity . . . . .	135
5.3.6	Numerical Results . . . . .	138
5.3.7	Conclusion . . . . .	142
<b>CHAPTER 6</b>	<b>INTERFERENCE ANALYSIS IN DIFFUSION-BASED</b>	
	<b>MOLECULAR COMMUNICATION . . . . .</b>	<b>144</b>
6.1	Motivation and Related Work . . . . .	144
6.2	Intersymbol Interference and Co-Channel Interference . . . . .	147
6.2.1	Interference Formulas . . . . .	149
6.2.2	Attenuation and Dispersion of Diffusion-Waves . . . . .	151
6.2.3	Interference Analysis . . . . .	153
6.2.4	Numerical Results . . . . .	157
6.2.5	Conclusion . . . . .	160
6.3	Statistical-physical Model of Interference . . . . .	161
6.3.1	Reference Models, Assumptions, and Definitions . . . . .	161
6.3.2	Statistical-physical Interference Modeling . . . . .	166
6.3.3	Log-Characteristic Function and PDF of the Received Power Spectral Density . . . . .	170
6.3.4	Numerical Results . . . . .	175
6.3.5	Conclusion . . . . .	180
<b>CHAPTER 7</b>	<b>A MOLECULAR COMMUNICATION SYSTEM DESIGN</b>	
	<b>VIA BIOLOGICAL CIRCUITS . . . . .</b>	<b>182</b>
7.1	Motivation and Related Work . . . . .	182
7.2	Biological Circuit Design for Diffusion-based MC . . . . .	184

7.2.1	Functional Blocks Description . . . . .	184
7.2.2	Reaction-based Description . . . . .	186
7.3	Deterministic Model . . . . .	189
7.3.1	Functional Block Transfer Functions . . . . .	191
7.3.2	Time Scale Approximation . . . . .	195
7.3.3	Attenuation and Delay Expressions . . . . .	197
7.4	Stochastic Model . . . . .	198
7.4.1	Noise in Chemical Reactions . . . . .	199
7.4.2	Noise in Diffusion . . . . .	210
7.4.3	Output Noise PSD Expression . . . . .	211
7.5	Numerical Results . . . . .	212
7.6	Conclusion . . . . .	215
<b>CHAPTER 8 CONCLUSION . . . . .</b>		<b>217</b>
<b>CHAPTER 9 APPENDICES . . . . .</b>		<b>224</b>
9.1	Particle Binding and Release Rates . . . . .	224
<b>PUBLICATIONS . . . . .</b>		<b>226</b>
<b>REFERENCES . . . . .</b>		<b>228</b>
<b>VITA . . . . .</b>		<b>236</b>

## LIST OF FIGURES

Figure 1	The expected functions of a biological nanomachine realized through the genetic engineering of a bacterium. . . . .	3
Figure 2	Molecular communication types. . . . .	11
Figure 3	Representation of the end-to-end model. . . . .	17
Figure 4	Representation of the emission process. . . . .	18
Figure 5	RC-circuit model of the emission process. . . . .	21
Figure 6	Representation of the diffusion process. . . . .	23
Figure 7	Representation of the reception process. . . . .	27
Figure 8	RC-circuit model of the reception process . . . . .	29
Figure 9	The normalized gain for the emission process $A$ . . . . .	33
Figure 10	The normalized gain for the diffusion process $B$ . . . . .	34
Figure 11	The group delay for the diffusion process $B$ . . . . .	34
Figure 12	The normalized gain for the reception process $C$ . . . . .	35
Figure 13	The group delay for the reception process $C$ . . . . .	35
Figure 14	The normalized gain for the reception process $T$ . . . . .	36
Figure 15	The group delay for the end-to-end model $T$ . . . . .	36
Figure 16	Block scheme of the end-to-end model that includes the noise sources. . .	40
Figure 17	Block scheme of the physical model for the particle sampling noise. . . .	43
Figure 18	Graphical sketch of the transmitter kinetic state $\bar{S}_T(t)$ at time $t$ . $\bar{S}_T(t)$ depends on the particle concentration rate $r_T(t)$ in input through the expressions in (37) and (38). . . . .	45
Figure 19	Block scheme of the stochastic model for the particle sampling noise . . .	47
Figure 20	The particle sampling noise physical model simulation input. . . . .	52
Figure 21	The particle sampling noise physical model simulation output. . . . .	52
Figure 22	The particle sampling stochastic model likelihood. . . . .	53
Figure 23	The Gaussian model likelihood for the particle sampling noise. . . . .	53

Figure 24	Block scheme of the physical model for the particle counting noise. . . . .	55
Figure 25	Graphical sketch of the receiver kinetic state $\bar{S}_R(t)$ at time $t$ . $\bar{S}_R(t)$ depends on the particle concentration $c_R(t)$ in input through the expressions in (63) and (64). . . . .	56
Figure 26	Block scheme of the particle counting noise stochastic model. . . . .	60
Figure 27	The particle counting noise physical model simulation input. . . . .	65
Figure 28	The particle counting noise physical model simulation output. . . . .	65
Figure 29	The particle counting stochastic model likelihood. . . . .	66
Figure 30	The Gaussian model likelihood for the particle counting noise. . . . .	66
Figure 31	Block scheme of the physical model for the ligand-receptor-binding noise. . . . .	70
Figure 32	Graphical sketch of the ligand-receptor kinetic state block. The number of particles $P(t)$ in the receptor space depends on the particle concentration $c_R(t)$ in input through the expressions in (102). . . . .	73
Figure 33	Block scheme of the stochastic chemical kinetics applied to the LIGAND-RECEPTOR BINDING process. . . . .	75
Figure 34	Graphical representation of (116) as a Markov chain. . . . .	78
Figure 35	Graphical representation of (118) as a Markov chain. . . . .	80
Figure 36	The input of the physical model and stochastic model of ligand-receptor-binding noise simulations in terms of particle concentration. . . . .	83
Figure 37	The output of the first set of simulations of the physical model in terms of number of bound chemical receptors (left) and isolated noise contribution (right). . . . .	85
Figure 38	The output of the first set of simulations on the reversible second order reaction model in terms of number of bound chemical receptors (left) and isolated noise contribution (right). . . . .	86
Figure 39	The output of the first set of simulations on the reversible first order reaction model in terms of number of bound chemical receptors (left) and isolated noise contribution (right). . . . .	86
Figure 40	The output of the second set of simulations on the physical model of the ligand-receptor-binding noise in terms of number of bound chemical receptors (left) and isolated noise contribution (right). . . . .	87

Figure 41	The output of the second set of simulations on the reversible second order reaction model in terms of number of bound chemical receptors (left) and isolated noise contribution (right). . . . .	87
Figure 42	The output of the second set of simulations on the reversible first order reaction model in terms of number of bound chemical receptors (left) and isolated noise contribution (right). . . . .	87
Figure 43	The reversible second order reaction log-likelihood for the first set of simulations. . . . .	90
Figure 44	The reversible first order reaction log-likelihood for the first set of simulations. . . . .	90
Figure 45	The log-likelihood of a Gaussian model for the first set of simulations. . .	90
Figure 46	The log-likelihood of a Gaussian model for the second set of simulations.	90
Figure 47	The reversible second order reaction log-likelihood for the second set of simulations. . . . .	91
Figure 48	The reversible first order reaction log-likelihood for the second set of simulations. . . . .	91
Figure 49	Information-theoretic diagram of a diffusion-based MC system. . . . .	96
Figure 50	Sketch of the reference physical model for diffusion-based MC. . . . .	97
Figure 51	Schematic diagram of the diffusion-based MC system with the physical reference model. . . . .	99
Figure 52	Capacity in relation to the transmitter-receiver Distance and for different values of the bandwidth $W$ . . . . .	114
Figure 53	Capacity in relation to the Bandwidth and for different values of the transmitter-receiver distance $d$ . . . . .	115
Figure 54	Capacity in relation to the Bandwidth and different values of the system temperature $T$ . . . . .	116
Figure 55	Sketch of the physical realization of the diffusion-based molecular communication system. . . . .	119
Figure 56	Diagram of the diffusion-based MC system with the Fick's diffusion and the particle location displacement contributions to the molecule diffusion.	123
Figure 57	Capacity in relation to the bandwidth and for different values of the transmitter-receiver distance $d$ . . . . .	138

Figure 58	Capacity in relation to the transmitter-receiver distance and for different values of the bandwidth $W$ . . . . .	140
Figure 59	Capacity in relation to the bandwidth and for different values of the average transmitted power $P_{\mathcal{H}}$ . . . . .	141
Figure 60	Block scheme of the diffusion-based MC system considered in this section of the Ph.D. thesis. . . . .	147
Figure 61	The received pulses as function of the time and the distance in the case of baseband modulation (left) and diffusion wave modulation (right) for the simulation-based for the simulation-based ISI evaluation (upper) and CCI evaluation (lower). . . . .	158
Figure 62	The ISI values for the baseband modulation scheme (upper) and the diffusion wave modulation scheme (lower) from the simulation (left) and from the simple formulas (right). . . . .	159
Figure 63	The CCI values for the the baseband modulation scheme (upper) and the diffusion wave modulation scheme (lower) from the simulation (left) and from the simple formulas (right). . . . .	160
Figure 64	Reference diffusion-based molecular nanonetwork considered for the interference modeling. . . . .	162
Figure 65	PDF $P_{S_Y(\omega)}(s)$ of the received PSD $S_Y(\omega)$ . Different curves refer to different values of the frequency $\omega$ . . . . .	173
Figure 66	PDF $P_{S_Y(\omega)}(s)$ of the received PSD $S_Y(\omega)$ for a range of frequencies $\omega$ from 0 to 2 Hz. . . . .	174
Figure 67	Simulation-based PDF $P_{S_Y(\omega)}(s)$ of the received PSD $S_Y(\omega)$ for a range of frequencies $\omega$ from 0 to 2 Hz. . . . .	176
Figure 68	Probability of interference according to the statistical-physical model. . .	178
Figure 69	Probability of interference according to the simulation environment. . . .	179
Figure 70	Main functional blocks of a biological circuit for diffusion-based molecular communication. . . . .	187
Figure 71	Decomposition of the transfer function of a biological circuit for diffusion-based molecular communication into the transfer functions of each functional block. . . . .	190
Figure 72	Approximation of the decomposition of the transfer function of a biological circuit for diffusion-based molecular communication. . . . .	196

Figure 73 Attenuation of the biological circuit for diffusion-based MC as function of the receiver distance from the transmitter  $r_{Rx}$  and the frequency  $\omega$ . . . . 213

Figure 74 Delay of the biological circuit for diffusion-based MC as function of the receiver distance from the transmitter  $r_{Rx}$  and the frequency  $\omega$ . . . . . 214

Figure 75 PSD of the output noise from the biological circuit for diffusion-based MC as function of the frequency  $\omega$  and the input signal  $PoPS_{in}$ . . . . . 215

## SUMMARY

Molecular communication (MC) is a promising bio-inspired paradigm for the exchange of information among nanotechnology-enabled devices. These devices, called nanomachines, are expected to have the ability to sense, compute and actuate, and interconnect into networks, called nanonetworks, to overcome their individual limitations and benefit from collaborative efforts. MC realizes the exchange of information through the transmission, propagation, and reception of molecules, and it is proposed as a feasible solution for nanonetworks. This idea is motivated by the observation of nature, where MC is successfully adopted by cells for intracellular and intercellular communication. MC-based nanonetworks have the potential to be the enabling technology for a wide range of applications, mostly in the biomedical, but also in the industrial and surveillance fields.

The focus of this Ph.D. thesis is on the most fundamental type of MC, i.e., diffusion-based MC, where the propagation of information-bearing molecules between a transmitter and a receiver is realized through free diffusion in a fluid. The objectives of the research presented in this thesis are to analyze the MC paradigm from the point of view of communication engineering and information theory, and to provide solutions to the modeling and design of MC-based nanonetworks. First, a physical end-to-end model is realized to study each component in a basic diffusion-based MC system design, as well as the overall system, in terms of gain and delay. Second, the noise sources affecting a diffusion-based MC are identified and statistically modeled. Third, upper/lower bounds to the capacity are derived to evaluate the information-theoretic performance of diffusion-based MC. Fourth, a stochastic analysis of the interference when multiple transmitters access the diffusion-based MC channel is provided. Fifth, as a proof of concept, a design of a diffusion-based MC system built upon genetically-engineered biological circuits is analyzed. This research provides fundamental results that establish a basis for the modeling, design, and realization of future MC-based nanonetworks, as novel technologies and tools are being developed.

# CHAPTER 1

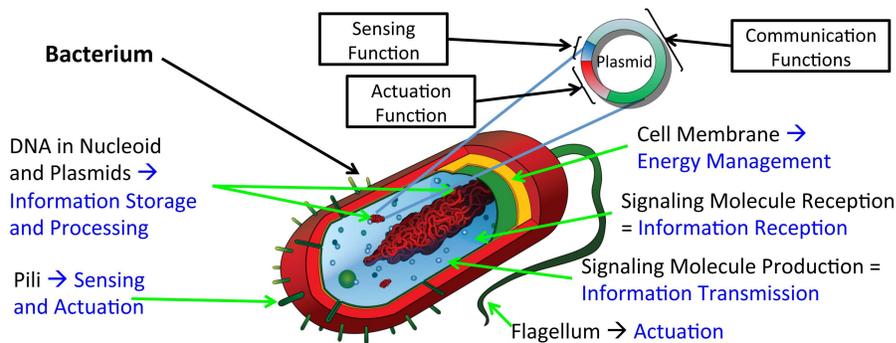
## INTRODUCTION

Molecular communication (MC) [1] is a bio-inspired paradigm where the exchange of information is realized through the transmission, propagation, and reception of molecules. This paradigm was first studied in biology, since it is successfully adopted in nature by cells for intracellular and intercellular communication [2]. MC is considered a promising option for communications in nanonetworks [3], which are defined as the interconnections of intelligent autonomous nanometer-scale devices, or nanomachines. Thanks to the feasibility of MC in biological environments, MC-based nanonetworks have the potential to be the enabling technology for a wide range of applications [3], mostly in the biomedical, but also in the industrial and surveillance fields. The objectives of the research presented in this thesis are to analyze the MC paradigm from the point of view of communication engineering and information theory, and to provide solutions to the modeling and design of MC-based nanonetworks.

### 1.1 Background

Among the more promising research fields of today, nanotechnology is enabling the manipulation of matter at an atomic and molecular scale, from one to a hundred nanometers. One of the goals of nanotechnology is to engineer functional systems based on the unique phenomena and properties of matter at the nanoscale [4]. Currently, a great research effort is spent in the attempt to realize nanoscale machines, also called molecular machines or nanomachines, defined by E. Drexler as “mechanical devices that perform useful functions using components of nanometer-scale and defined molecular structure” [5]. More specifically, nanomachines [3, 6, 7] are expected to have the ability to sense, compute, actuate, manage their energy, and interconnect into networks, termed nanonetworks, to overcome their individual limitations and benefit from collaborative efforts.

Two main types of nanomachines can be identified within the aforementioned definition, namely, synthetic and biological. On the one hand, the synthetic nanomachines are realized either by downscaling from the current micro-scale technologies, such as micro-electronics or micro-electro mechanics, or through the use of chemically synthesized nano-materials [6]. On the other hand, the biological nanomachines are realized either by reusing biological components (e.g., DNA-based memories [8], flagellum-based actuators [9]), or by programming the behavior of biological cells from nature, such as through the genetic engineering of bacteria [10], as illustrated in 1.



**Figure 1. The expected functions of a biological nanomachine realized through the genetic engineering of a bacterium.**

While the engineering of fully synthetic nanomachines is still in its infancy, the research on the genetic engineering of biological cells is currently in rapid progress, thanks to the advancements made by biotechnology [11]. Several key techniques developed under the umbrella of synthetic biology have made possible today the realization of simple biological nanomachines [12]. As illustrated in Figure 1, through the insertion of engineered genetic code in the form of a circular DNA strand (i.e., plasmid) in a bacterium, it will be soon possible to program complete functions, including sensing, actuation, and communication, and have access to the main functionalities of the cell, such as the storage and the processing of information through DNA code, the sensing and actuation through the use of the *pili* (hairlike appendages), the management of the cell energy through the cell membrane, and the transmission and reception of information through the production and the reception of

signaling molecules.

The exchange of information between nanomachines, and their interconnection into nanonetworks, is key to overcome their individual limitations in size, energy and computational capabilities, and benefit from collaborative efforts. In nanonetworks, the applicability of classical communication technologies is limited by several constraints. In particular, the very restricted size of the nanomachines and the peculiarities of the environments in which they are envisioned to operate (e.g., biological scenarios) demand for novel solutions from the perspective of both the choice of the communication medium and the study of suitable communication techniques. While a possible solution to the problem of communication between synthetic nanomachines is suggested by recent studies [6] on nano-structures and on the properties of carbon nano-electronics, the imminent availability of biological nanomachines encourages to study and adopt the communication techniques naturally adopted by biological cells. In this direction, the MC paradigm, inspired by the natural cell communication in biology, where message-carrying molecules are synthesized, emitted, collected, and converted to cellular responses through biochemical processes, is expected to be especially attractive because of its inherent feasibility in a biocompatible environment [1, 3].

## 1.2 Potential Applications

Given the tight integration of MC within the biological environment and its feasibility at the cellular scale (nm -  $\mu\text{m}$ ), MC is studied not only as a candidate for nanonetwork communication, but also as a possible tool for the future nanonetworks to interact with the living organisms and their biological processes. As a consequence, the number of potential applications of MC-enabled nanonetworks is very large. Amongst others, the following three main areas deserve a special attention.

**Biomedical applications**, such as disease control and infectious agent detection [13], smart drug delivery systems [14], and intelligent intrabody systems for monitoring glucose, sodium, and cholesterol [15, 16]. These applications are expected to greatly benefit from

the use of nanomachines deployed over the body (e.g., through tattoo-like patches) or inside the body (e.g., through pills or intramuscular injection). Since MC is naturally adopted by cells, nanonetworks enabled by this paradigm are envisioned to better integrate with the intra-body biological processes and to show higher biocompatibility when compared to other possible solutions.

**Industrial applications**, such as the monitoring and control of microbial formations. As an example, applications based on bacterial biofilms [17], which are used to clean residual waters coming from different manufacturing processes or to treat organic waste [18], could be greatly enhanced by MC-enabled nanonetworks, since microbial organisms naturally produce and respond to molecular stimuli.

**Surveillance applications** will make use of biological and chemical nanosensors that have an unprecedented sensing accuracy [19, 20]. Nanonetworks composed by several MC-enabled nanosensors could serve for surveillance against biological and chemical attacks [20] by detecting toxic or infectious agents diffusing in the environment.

In particular, with specific reference to the use of genetically engineered bacteria [21], and their interconnection in MC-based nanonetworks, the following example applications are envisioned:

- *Intrabody bacteria-based sensor-actuator networks for diagnosis and drug delivery.* Genetically engineered bacteria could be deployed in the human gastrointestinal tract, where they would serve as sensors of particular biomarkers generated by inflammations or ulcers (e.g., Crohn's disease) [22]. Upon sensing an inflammation, these bacteria would communicate with each other, perform a consensus decision-making process, and cooperate to produce drug molecules, finally providing a localized and timely healing of the infection.
- *Microfluidic bacteria-based network-on-chip for chemical analysis.* MC-based nanonetworks of genetically-engineered bacteria could be used to increase the throughput

of a microfluidic device for chemical analysis by enabling on-chip information exchange, where the molecular communication between bacteria could enhance procedures for enzymatic analysis, immunoassays, DNA analysis, and proteomics [23]. The dynamic exchange of molecules in these nanonetworks could also be used to combine multi-stage reactions on a single microfluidic device, thus achieving not only higher throughput and accuracy, but also the automation of chemical analysis that would otherwise require the intervention of an external human operator [24].

- *Bacteria-based environmental monitoring.* Bacteria could be deployed in the environment, where they could detect and communicate the presence of chemicals or changes in the status of other biological organisms [25]. As an example, bacteria could be used to monitor the presence of toxic contaminants in runoff waters from factory plants. The same mechanism could be applied to genetically engineered bacteria from species that are already present in food, (e.g., *Lactobacillus* in yogurt), which could cooperatively decide when the food expires and it is not safe to eat, and trigger a signal easily detectable by humans, such as a change in their color or the generation of fluorescence. This would increase food safety and reduce the waste of food, by effectively monitoring the edibility even after an expiration date.

### **1.3 Research Objectives and Solutions**

The focus of this Ph.D. thesis is on diffusion-based MC, where the propagation of information-bearing molecules between a transmitter and a receiver is realized through free diffusion in a fluid. This choice is motivated by a preliminary analysis, detailed in Chapter 2, which identifies the diffusion-based as the most fundamental type of MC among different options suggested in the literature. As a consequence of the differences between the diffusion-based MC paradigm and classical electromagnetic communication paradigms, the classical communication engineering models and techniques are not directly applicable for the study and the design of diffusion-based MC systems. These differences include, but

are not limited to, the following:

- The process of diffusion-based molecule propagation is based on radically different phenomena with respect to the electromagnetic wave propagation in classical communication systems. While electromagnetic waves operate the propagation of energy at the speed of light, the molecule diffusion process is caused by the random walk of the molecule Brownian motion in a fluid [26, 27]. As a consequence, while an electromagnetic wave propagates in a defined direction, and with negligible delay for most of the terrestrial communication systems, molecules subject to Brownian motion propagate with a random direction and with a high delay for almost all the transmission ranges of interest.
- The biologically-inspired physical processes that can be adopted to transmit and receive information in a diffusion-based MC system are based on different mechanisms with respect to the modulation and reception of electromagnetic radiations in classical communication systems. While in classical systems antennas transmit and receive electromagnetic radiations through moving charges in metallic conductors, in biological cell bio-signaling [2] information is transmitted through the chemical synthesis of signaling molecules, and received through chemical reactions between incoming signaling molecules and chemical receptors.

As a consequence, there is a need of to build a complete understanding of the diffusion-based MC paradigm from the ground up. The research objectives addressed in this Ph.D. thesis, and the proposed solutions, have been identified to specifically target this need, and they are summarized in the following.

The first research objective is to develop of a physical end-to-end modeling of diffusion-based MC, which provides a mathematical characterization of the main physical processes involved in the transmission, propagation, and reception of molecules for the exchange of

information between a transmitter and a receiver. To achieve this objective, a system design, inspired by biochemistry principles related to the living cells and to the mechanism of cell bio-signaling [2], is described as the most basic implementation of diffusion-based MC. This design identifies the three physical processes of molecule emission at the transmitter, molecule diffusion in the fluid medium, and molecule chemical reception. For each process, as well as for the overall end-to-end system, mathematical models are provided in terms of transfer function, and consequently, in terms of gain and delay experienced by an information signal exchanged between the transmitter and the receiver.

The second research objective is to identify and model the noise sources that affect the diffusion-based MC. The modeling of the noise affecting a communication system is fundamental to help the design of the system components, and increase the probability of correct exchange of information between the transmitter and the receiver. By stemming from the aforementioned physical end-to-end modeling, three noise sources affecting diffusion-based MC are identified, namely, the sampling noise, the counting noise, and the ligand-receptor-binding noise, and they are related to the transmitter, the signal propagation in the channel, and the receiver, respectively. Each noise source is described through a mathematical analysis of the physical processes which generate the noise, and stochastically modeled in terms of noise-generating random processes.

The third research objective is to provide an estimate of the achievable performance of a diffusion-based MC system in terms of information capacity. The diffusion-based MC has two main characteristics, namely, a channel with memory and a signal-dependent noise, which notoriously make impossible to find a closed-form analytical expression for the true capacity of a communication system. In the attempt to provide an analytical closed-form expression that relates the performance of a diffusion-based MC system to physical parameters, such as the diffusion coefficient, the temperature, the transmitter-receiver distance, the bandwidth of the transmitted signal, and the average transmitted power, the results included in this Ph.D. thesis are based on some simplifying assumptions. A first expression

is derived from statistical mechanics and equilibrium thermodynamics considerations, and it is based on assumption of having a molecular system in equilibrium. A second expression is based on a pure information-theoretic approach based on the decomposition of the molecule diffusion into two main processes, namely, the Fick's diffusion and the molecule location displacement. These assumptions limit the validity of these closed-form analytical expressions, which have to be considered as upper/lower bounds to the true information capacity.

The fourth research objective is to analyze the interference produced at the reception side of a diffusion-based MC system. Given an information signal sent by a transmitter to a receiver in a diffusion-based MC system, the interference may be either due to distortions of the signal itself or by other concurrent signals coming from different transmitters. The analysis of these types of interference is fundamental to design interference mitigation techniques and increase the performance of a communication systems. In this Ph.D. thesis, first, the InterSymbol Interference (ISI) and the Co-Channel Interference (CCI) are jointly analyzed for a diffusion-based MC system under the assumptions of having a Gaussian-pulse-based information encoding and a limited number of transmitters in predetermined locations. Second, the statistical-physical modeling of the interference is provided through an analytical expression of the Power Spectral Density (PSD) probability distribution of the received signal, independent from the transmitter number, specific transmitter locations, or coding schemes.

The fifth research objective is to provide a possible design of a molecular communication system by stemming from a currently viable technology. In this direction, synthetic biology techniques, and in particular the engineering of biological circuits, allow today to program logical functions within biological cells, thus paving the way for the realization of the aforementioned biological nanomachines. The design of an MC system built upon biological circuits is therefore described in this Ph.D. thesis by identifying a minimal subset of elements to realize diffusion-based MC between biological cells. A mathematical model is

then detailed in terms of transfer functions, from which analytical expressions are derived for the attenuation and the delay experienced by an information signal through the biological circuit. Finally, the most significant noise sources within the biological circuit are identified, stochastically modeled, and quantified in terms of noise Power Spectral Density (PSD) generated at the output of the system.

## **1.4 Thesis Outline**

The rest of this Ph.D. is organized as follows. A preliminary analysis of different MC options from the literature is contained in Chapter 2, which also includes a survey of the results from previous works pertinent to the study of diffusion-based MC. The results obtained through the design and end-to-end modeling of a basic diffusion-based-MC system are presented in Chapter 3, where the contributions of each component of the system are analyzed in terms of gain and delay. In Chapter 4, the most relevant noise sources affecting diffusion-based MC are studied through the mathematical expression of their underlying physical processes, and modeled through the use of statistical parameters. Analytical expressions of upper and lower bounds to the information capacity of diffusion-based MC systems are derived in Chapter 5, first by using tools from thermodynamics, and then through a pure information-theoretic approach. In Chapter 6, an analysis of the interference in diffusion-based MC is provided first for specific cases, and then through a statistical model with general validity. The design of a diffusion-based MC system built upon genetically-engineered biological circuits is presented and modeled in Chapter 7, where implementation-specific results in terms of attenuation, delay, and noise impairments are discussed. A conclusion with the possible future avenues for this research field is provided in Chapter 8. Finally, the publications resulted from the research presented in this Ph.D. thesis are listed.

## CHAPTER 2

### PREVIOUS WORK ON MOLECULAR COMMUNICATION

This chapter of the Ph.D. thesis contains a review of the literature pertinent to the research on MC and its application to nanonetworks. This review is organized as follows. In Section 2.1, different MC options from the literature are detailed on the basis of the type of adopted molecule propagation. In Section 2.2, the results from the literature focused on diffusion-based MC are presented and analyzed to motivate the research proposed in this document.

#### 2.1 Molecular Communication Types

Different types of MC have been studied so far, and they can be characterized on the basis of how molecules propagate through the medium [1]. Three main types of MC are identified from the literature, as shown in Figure 2, as the walkway-based, the advection-based, and the diffusion-based. These categories are also described on the basis of how spontaneous is the propagation of the molecules. The less spontaneous is the type of MC, the more predictable is the path followed by the molecules during their propagation, and *vice versa*.

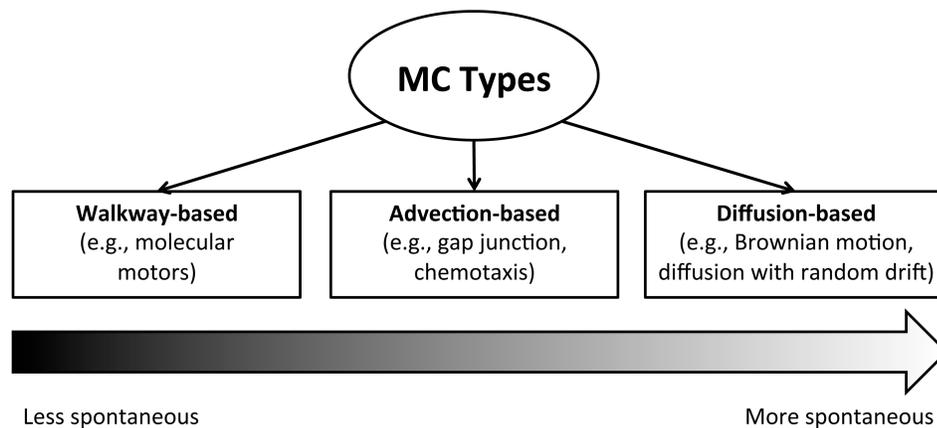


Figure 2. Molecular communication types.

In **walkway-based** MC, the molecules propagate through active transport by following pre-defined pathways connecting the transmitter to the receiver. The most widespread example from the literature of this type of MC is given by the use of molecular motors [28, 29]. Molecular motors [30] are protein filaments that convert chemical energy into kinetic energy, and they are at the basis of the generation of force in biology, e.g., in the muscles. In [28, 29] molecular motors filaments are studied and employed to physically interconnect nanomachines and to generate the force to propagate the information molecules from the transmitter to the receiver.

In **advection-based** MC, the molecules propagate through diffusion in a fluidic medium whose flow, and consequently the transport of the molecules, or advection, is defined and predictable. The use of gap junctions [31], which are nanofluidic pipes connecting the transmitter and the receiver, is one example of this type of MC. Advection-based MC can also be realized by using carrier entities whose motion is constrained on the average along specific paths, despite showing a random component. A good example of this flow-based MC is given by the chemotaxis-based technique in [32, 33], where flagellated bacteria released by the transmitter, and containing information encoded in DNA molecules, are guided to swim towards the receiver, which continuously releases an attractant substance.

In **diffusion-based** MC, the molecules propagate through their spontaneous diffusion in a fluidic medium [34]. Diffusion-based MC was first analyzed in [35, 36, 37], where the transmitted information molecules cover the distance between the transmitter and the receiver by following an unbiased random walk, i.e., the Brownian motion. In diffusion-based MC, the molecules can also be affected by random advection in the fluidic medium. An example of this diffusion-based MC is given in [38], where information molecules are released in the air and are subject to a random and unpredictable component of the wind flow while they propagate.

The focus of the proposed research is on diffusion-based MC. This choice is motivated

by the following arguments. First, molecule diffusion is at the basis of all the aforementioned MC options. As an example, in [29], diffusion is considered as an unavoidable propagation effect superimposed on the molecular motor transport. Moreover, in [32, 33], carrier bacteria are subject to a random walk with the same underlying rules of a biased molecule diffusion. Second, diffusion-based MC can be considered the most general and common MC option in nature, where examples are found in the calcium signaling among cells [39], the pheromonal communication among animals [40, 38], and the synaptic transmission between neurons [2]. As a consequence, the study of diffusion-based MC has the potential to benefit from the observation of nature for the design of bio-inspired and bio-compatible solutions.

## **2.2 Diffusion-based Molecular Communication Architectures**

Different diffusion-based MC architectures for nanonetworks have been proposed and studied in the literature, and they can be classified on the basis of the technique used to encode the information in the diffusing molecules. The three most referenced architectures encode the information in either the time of molecule release, the type of each molecule, or variations in the molecule concentration.

The first type of architecture is theoretically analyzed in [41], where the authors focused on the mathematical modeling of the diffusion channel as a probabilistic contribution in the time of arrival of the molecules at the receiver, while the transmitter and the receiver have simplified ideal models. The results of simulations from [41] in terms of achievable information rate show that, as a consequence of the high uncertainty in the propagation time, this architecture is characterized by a very low capacity.

The second type of architecture is analyzed in [29], where a piece of information is encoded in each single molecule. As a consequence, the information carried by a certain molecule is received only if that molecule reaches the receiver location. Similarly to [41],

only the diffusion channel part is modeled, while the transmitter and the receiver are considered ideal. While no analytical model is provided in [29], the results of simulations reveal very low performance for this architecture if compared to other non-diffusive techniques.

The third type of architecture is treated in [36, 37, 42, 43, 44, 45, 46, 47]. In particular, a simplified receiver model of this architecture is studied in [36], where the diffusion-based channel is coupled with ideal transmitter and receiver. Simulation results in [36] show low values for the system capacity. In [37, 42], a model of a diffusion-based MC receiver is developed by using multiple chemical receptors to read the molecule concentration at the receiver location. The results of this model in terms of capacity are relatively high if compared to [36], especially when an error compensation technique is applied.

The high similarity of the third type of architecture to some biological systems [2], which are characterized by much higher performance than in the aforementioned results, encourages the investigation in this direction.

In particular, very limited research has been conducted to analyze the complete end-to-end (transmitter, channel and receiver) diffusion-based MC architecture and the complete study of the noise sources, the capacity and the interference with closed-form analytical models. In [36], the analytical model of the system is reduced to the diffusion-based channel. In [37, 42, 43] the diffusion process is not captured in terms of molecule-propagation theory and, therefore, the end-to-end model reliability and accuracy are only accounted for the receiver side. In [44], the capacity of a MC system is analyzed on the basis of the effects of the diffusion-based-channel memory, but without accounting for molecular-noise sources and only for the specific case of binary coding. Two different coding techniques are analyzed in [46] in terms of achievable rates, while the diffusion-channel models are oversimplified to a binary or a quadruple channel. Similarly, discrete-memoryless approximations are applied to the molecule-diffusion channel in [48], where the MC capacity is computed for a binary-coding scheme. In [46] the effects of intersymbol and co-channel interference are analyzed in reference to two specific modulation techniques proposed by

the same authors. In [45] the intersymbol interference is characterized in a unicast MC system with binary amplitude-modulation. In [47], interference is studied for another specific modulation technique, based on the transmission order of different types of molecules.

# CHAPTER 3

## PHYSICAL END-TO-END MODEL OF DIFFUSION-BASED MOLECULAR COMMUNICATION

### 3.1 Motivation and Related Work

The objective of the physical end-to-end modeling of diffusion-based Molecular Communication (MC) is to provide a mathematical characterization of the main physical processes involved in the transmission, propagation, and reception of molecules for the exchange of information between a transmitter and a receiver. This characterization is provided in this chapter of the Ph.D. thesis in terms of transfer function and, consequently, by deriving the gain and delay parameters for each physical process, as well as for the overall end-to-end diffusion-based MC system. This characterization does not take into account noise sources that could affect the performance of a diffusion-based MC system, which will be the subject of the next chapter.

Some previous research has been conducted to address the modeling and analysis of diffusion-based MC and the according end-to-end behavior in nanonetworks. In [37, 42], a particle receiver model is developed by taking the ligand-receptor binding mechanism into account [49]. However, in both works, the diffusion process is not captured in terms of molecule propagation theory and, therefore, the end-to-end model reliability and accuracy are only accounted for the receiver side. Moreover, an ideal digital transmitter model is used and the performance evaluation is conducted based on an ideal synchronization between the transmitter and the receiver.

In this chapter, a basic diffusion-based MC system design is provided, which aims at an interpretation of the diffusion-based MC, both in terms of molecule emission/reception and molecule propagation. In the provided diffusion-based MC system design, the desired information modulates the molecule emission at the transmitter side. This modulated signal is then propagated by the diffusion process to the receiver side. The receiver detects the

concentration and generates the received signal. This diffusion-based MC system is divided into three processes: the emission process, the diffusion process and the reception process. Each process is modeled in terms of transfer function, from which the normalized gain and delay are derived. The models of the emission process and the diffusion process are built on the basis of the molecular-diffusion physics [26], while the reception-process model is interpreted by stemming from the theory of the ligand-receptor binding [49]. Finally, the end-to-end normalized gain and delay are derived for the cascade of the three processes, which provide the characterization of the complete diffusion-based MC system. Numerical results are provided to evaluate the normalized gain and delay of the diffusion-based MC system for several different values of the system frequency and the transmission range.

### 3.2 A Basic Design of a Diffusion-based MC System

The end-to-end model is a characterization of a basic diffusion-based MC design in terms of transmission, propagation, and reception of molecules, as sketched in Figure 3.

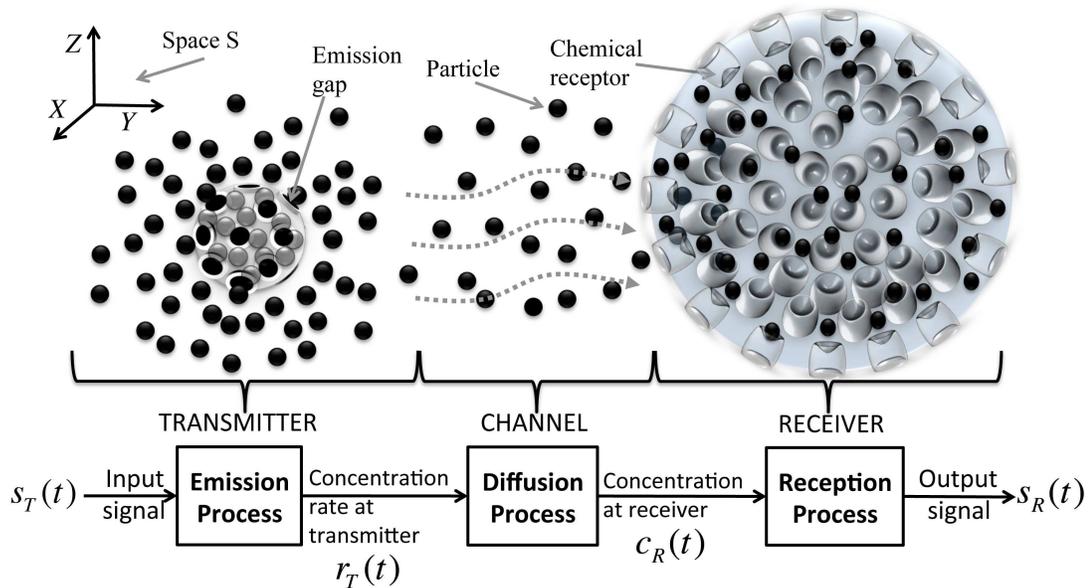


Figure 3. Representation of the end-to-end model.

The molecules are here called particles, since in this preliminary work their chemical properties are not taken into account. The particles are assumed as having no electrostatic

charge, while their mass and shape characteristics are taken into account in the diffusion coefficient  $D$  [26, 27], introduced later in this chapter. The end-to-end model is contained in the three-dimensional space  $S$  indexed through the Cartesian axes  $X$ ,  $Y$ , and  $Z$ . The space  $S$  contains a fluidic medium, and it is initially filled with a homogeneous concentration of particles. The emission process modulates the particle-concentration rate at the transmitter according to an input signal  $s_T(t)$ , function of the time  $t$ . The modulation is achieved through the release/capture of particles from/into emission gaps. The modulated-particle-concentration rate  $r_T(t)$  is the output of the transmitter and the input of the channel. The channel relies on the diffusion process of the particles in the space  $S$  to propagate the signal and deliver the particle concentration  $c_R(t)$  at the receiver. The receiver senses the particle concentration at its location as input, and it recovers the output signal  $s_R(t)$ . The reception process generates the output signal by means of chemical receptors.

### 3.3 The Emission Process

The task of the emission process is to modulate the particle-concentration rate  $r_T(t)$  at the transmitter according to the input signal  $s_T(t)$  of the end-to-end model, as shown in Figure 4, which is based on the following assumptions:

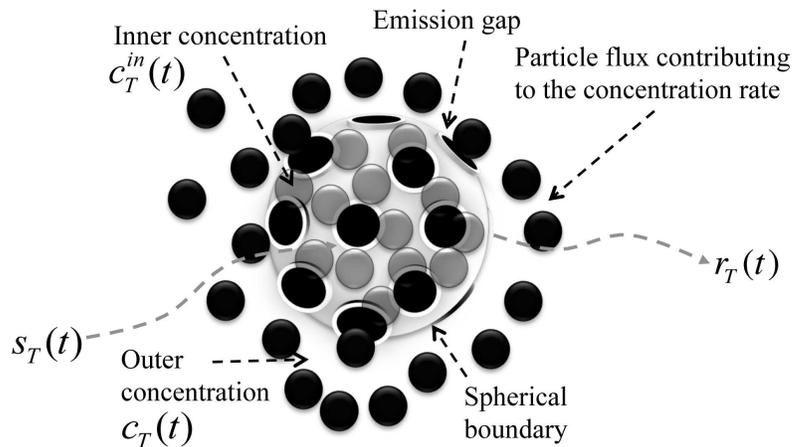


Figure 4. Representation of the emission process.

- The transmitter has a *spherical boundary* that divides the space in proximity of the

transmitter into two areas: the inner area and the outer area.

- The inner concentration  $c_T^{in}(t)$  is the concentration of particles lying in the inner area, whereas the outer concentration  $c_T(t)$  is the concentration of particles lying in the outer area.
- The inner area and the outer area are spatially connected by means of *emission gaps*. An emission gap is an opening in the spherical boundary which allows particles to move through due to their diffusion. The size of an emission gap allows only one particle to pass through at each time instant. Whenever a particle is traversing the emission gap, its movement has only components along the radius of the spherical boundary. As a consequence, the movement of a particle through the emission gap can only be outward (from the inner area to the outer area) or inward (from the outer area to the inner area). The emission gaps are many and homogeneously distributed on the surface of the spherical boundary. The present analysis does not depend on their precise number. We believe it will be important to discuss the impact of the number of emission gaps on the end-to-end model in our future work.
- Whenever there is a difference between the inner concentration  $c_T^{in}(t)$  and the outer concentration  $c_T(t)$ , a movement of particles is stimulated between the inner area and the outer area through the emission gaps.
- The movement of particles between the inner area and the outer area causes a variation in the outer concentration, whose first time derivative is the particle concentration rate at the transmitter location  $r_T(t)$ .
- Particles can be created/destroyed in the inner area in order to reach a desired inner concentration  $c_T^{in}(t)$ . The creation/destruction of particles in the inner area is supposed to be ideally perfect and instantaneous. As a consequence, we do not account for the randomness that can derive from the creation/destruction of particles. We

believe that this is a reasonable approximation that allows us to analyze the contributions coming only from the emission process. Further analysis can be conducted by specifying the processes involved in the creation/destruction of particles. As an example, the creation/destruction of particles could be realized through a cascade of chemical reactions or by the emptying/filling of particle reservoirs located in the inner area.

- The transmitter is supposed to be able to adjust the inner concentration  $c_T^{\text{in}}(t)$  in order to obtain a particle concentration rate  $r_T(t)$  proportional to the input signal  $s_T(t)$  (modulation of  $r_T(t)$  according to  $s_T(t)$ ).

This transmitter characterization is inspired by biochemistry principles related to the living cells and to the mechanism of cell bio-signaling [2]. According to these principles, the spherical boundary is a simplification of the cell plasma membrane, which separates the interior of a cell from the outside environment. The emission gaps are inspired by the channels that permit the selective passage of molecules through the plasma membrane of a cell. As an example, the gated ion channels in the plasma membrane are openings that allow the passage of specific ion molecules between the interior of a cell and the outside environment and, amongst others, they serve for cell-to-cell communication purposes. As stated in [2], those ion molecules, while traversing the gated ion channels, are driven by a force that is a sum of two terms. The first term of the force is a function of the difference between the inside and the outside concentration of the same molecules and it depends on the diffusion. The second term of the force is a function of an electrical potential and it is related to the electrostatic charge carried by the ion molecules. Since, according to our assumption, the particles in our system do not carry any electrostatic charge, when they traverse an emission gap they are driven only by the first term of the force. For this, the difference in the concentration of particles between the inner area and the outer area stimulates the driving force that permits their movement through the emission gaps either outward or inward, as explained above.

The emission process is modeled with an electrical-parallel-resistor-capacitor (RC) circuit [50]. This circuit is shown in Figure 5, where  $I_{in}(t)$  is the input current as a function of the time  $t$ ,  $R_e$  stands for the resistance value,  $C_e$  is the capacitance value and  $I_{out}(t)$  is the output current, which is equal to the current  $I_R(t)$  flowing through the resistance  $R_e$ .

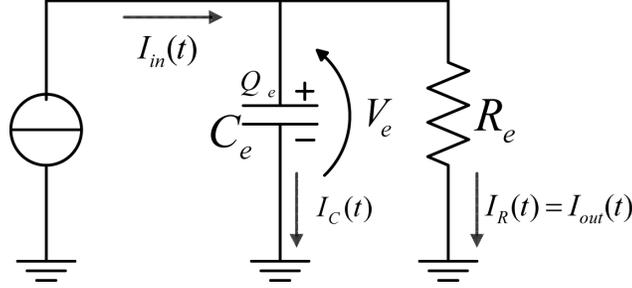


Figure 5. RC-circuit model of the emission process.

Within the RC-circuit model, the input signal  $s_T(t)$ , shown in Figure 4, corresponds to the input current  $I_{in}(t)$ . The particle-concentration gradient  $\nabla_{c_T}(t)$  at the transmitter, defined as the difference between the particle concentration inside and the particle concentration outside the transmitter, is equal to the voltage  $V_e(t)$ . The current  $I_R(t)$  that flows through the resistance  $R_e$  is equal to the particle-concentration rate  $r_T(t)$ , which is the output of the transmitter. The relation between the particle concentration flux  $\bar{J}(\bar{x}, t)$  and the particle-concentration gradient  $\nabla c(\bar{x}, t)$  at time  $t$  and location  $\bar{x}$  is given by the Fick's first law [26, 27] as follows:

$$\bar{J}(\bar{x}, t) = -D\nabla c(\bar{x}, t) , \quad (1)$$

where  $D$  is the diffusion coefficient and it can be considered as a constant value for a specific fluidic medium.

Therefore, since  $\bar{J}_T(t)$  and  $\nabla_{c_T}(t)$  are, respectively, the particle-concentration flux  $\bar{J}(\bar{x}, t)$  and the opposite of the particle-concentration gradient  $-\nabla c(\bar{x}, t)$  at the transmitter, and since

$$I_R(t) = \bar{J}_T(t) \quad V_e(t) = \nabla_{c_T}(t) , \quad (2)$$

then

$$I_R(t) = DV_e(t) , \quad (3)$$

and the constant resistance value becomes

$$R_e = \frac{1}{D} . \quad (4)$$

We relate the capacitor charging/discharging current  $I_C(t)$  at time  $t$  to the difference between  $s_T(t)$ , equal to the input current  $I_{in}(t)$ , and  $r_T(t)$ , equal to the output current  $I_R(t)$ . The voltage applied to the capacitor  $V_e(t)$  is equal to  $\nabla c_T(t)$ . The particle-concentration gradient  $\nabla c_T(t)$  is the difference between the outside particle concentration  $c_T(t)$  and the inside particle concentration  $c_T^{in}(t)$ . Therefore, the time derivative  $d\nabla c_T(t)/dt$  changes according to the net flux of particles that contributes to  $\nabla c_T(t)$ . The net flux of particles is given by the difference between  $s_T(t)$  and  $r_T(t)$ . This results in the following relation:

$$\frac{d\nabla c_T(t)}{dt} = s_T(t) - r_T(t) , \quad (5)$$

and, since  $I_C(t) = s_T(t) - r_T(t)$  and  $V_e(t) = \nabla c_T(t)$ , then  $I_C(t) = dV_e(t)/dt$  and, therefore, the capacitor value becomes

$$C_e = 1 . \quad (6)$$

From the electrical-circuit theory [50], the Fourier transform of the transfer function (FTTF) [51] of the RC circuit  $\mathbf{H}_{RC}(f)$ , equal to the FTTF  $\tilde{\mathbf{A}}(f)$  of the transmitter, has the following expression:

$$\mathbf{H}_{RC}(f) = \tilde{\mathbf{A}}(f) = \frac{\tilde{\mathbf{I}}_{out}(f)}{\tilde{\mathbf{I}}_{in}(f)} = \frac{1}{1 + j2\pi f R_e C_e} , \quad (7)$$

where  $\tilde{\mathbf{I}}_{in}(f)$  and  $\tilde{\mathbf{I}}_{out}(f)$  are the Fourier transforms [51] of the input voltage  $I_{in}(t)$  and output voltage  $I_{out}(t)$ , respectively.

The *normalized gain*  $\Gamma_A(f)$  of the emission process is the magnitude  $|\tilde{\mathbf{A}}(f)|$  of the FTTF  $\tilde{\mathbf{A}}(f)$  normalized by its maximum value  $\max_f(|\tilde{\mathbf{A}}(f)|)$ , which becomes 1 from (7), and it is expressed as

$$\Gamma_A(f) = \frac{|\tilde{\mathbf{A}}(f)|}{\max_f(|\tilde{\mathbf{A}}(f)|)} = \frac{1}{\sqrt{(1 + (2\pi f R_e C_e)^2)}} . \quad (8)$$

The *delay*  $\tau_A(f)$  of the emission process is

$$\tau_A(f) = -\frac{d\phi_A(f)}{df}, \quad (9)$$

where  $\phi_A(f)$  is the phase of the FTTF from (7), and it is expressed as

$$\phi_A(f) = \arctan\left(\frac{\text{Im}(\tilde{\mathbf{A}}(f))}{\text{Re}(\tilde{\mathbf{A}}(f))}\right) = \arctan(-2\pi f R_e C_e), \quad (10)$$

which is computed from the real part  $\text{Re}(\tilde{\mathbf{A}}(f))$  and the imaginary part  $\text{Im}(\tilde{\mathbf{A}}(f))$  of the FTTF  $\tilde{\mathbf{A}}(f)$ .

### 3.4 The Diffusion Process

The diffusion process, as shown in Figure 6, deals with the propagation of the particle-concentration rate  $r_T(t)$  from the transmitter, located at the Cartesian coordinates  $[0, 0, 0]$ , across the space  $S$  by means of free particle-diffusion in the fluidic medium. The particle concentration  $c_R(t)$  at the receiver location  $[x_R, y_R, z_R]$  is considered the output of the diffusion process.

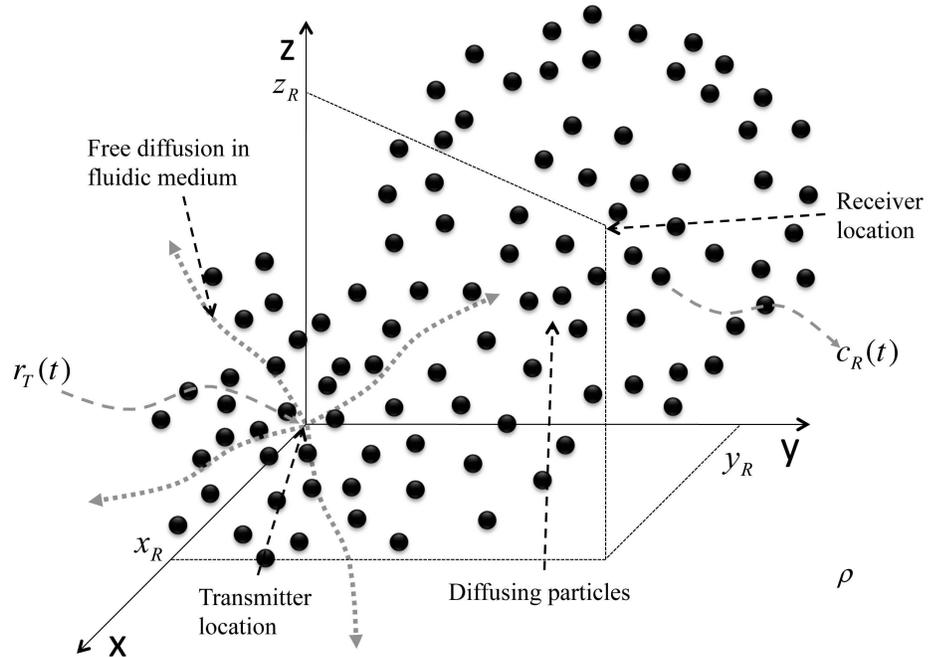


Figure 6. Representation of the diffusion process.

We use the particle concentration distribution flux to study the signal propagation occurring during the diffusion process. According to the Fick's first law [26, 27], the particle concentration flux  $\bar{J}(\bar{x}, t)$  at time instant  $t$  and location  $\bar{x}$ , is equal to the spatial gradient (operator  $\nabla$ ) of the particle concentration  $c(\bar{x}, t)$  occurring at time instant  $t$  and location  $\bar{x}$  multiplied by the diffusion coefficient  $D$ , expressed as follows:

$$\bar{J}(\bar{x}, t) = -D\nabla c(\bar{x}, t) , \quad (11)$$

where  $\nabla c(\bar{x}, t)$  is a vector containing the spatial first derivatives of  $c(\bar{x}, t)$  along the three spatial dimensions.

At time  $t$  we assume to have a particle concentration rate  $r(\bar{x}, t)$  at the location  $\bar{x}$  in the space  $S$  and time  $t$ . The principle of mass/matter conservation allows us to formulate the Continuity Equation [52], which states that the time derivative of the particle concentration  $\partial c(\bar{x}, t)/\partial t$  is equal to

$$\frac{\partial c(\bar{x}, t)}{\partial t} = -\nabla \bar{J}(\bar{x}, t) + r(\bar{x}, t) . \quad (12)$$

Substituting the Fick's first law from (88) into (12), we end up with the inhomogeneous Fick's second law of diffusion. According to this, the time derivative of the particle concentration  $\partial c(\bar{x}, t)/\partial t$  at location  $\bar{x}$  and the time  $t$  is equal to the Laplacian (operator  $\nabla^2$ ) of  $c(\bar{x}, t)$  occurring at time instant  $t$  and location  $\bar{x}$  multiplied by  $D$  (diffusion coefficient) plus the incoming particle concentration rate  $r(\bar{x}, t)$ , expressed as follows:

$$\frac{\partial c(\bar{x}, t)}{\partial t} = D\nabla^2 c(\bar{x}, t) + r(\bar{x}, t) , \quad (13)$$

where  $\nabla^2 c(\bar{x}, t)$  is the sum of the spatial second derivatives of  $c(\bar{x}, t)$  along the three spatial dimensions.

As pointed out in [53], the second Fick's law is in contradiction with the theory of special relativity. The solution of the second Fick's law allows particle concentration information to propagate instantaneously from one point to another point in the space  $S$ , with a so-called super-luminal information propagation speed. In order to overcome this problem,

it was proposed in [54] to add a new term to the Fick's second law accounting for a finite speed of propagation in the concentration information. With this additional term, we obtain the following Telegraph Equation [54]:

$$\tau_d \frac{\partial^2 c(\bar{x}, t)}{\partial t^2} + \frac{\partial c(\bar{x}, t)}{\partial t} = D \nabla^2 c(\bar{x}, t) + r(\bar{x}, t), \quad (14)$$

where  $\tau_d$  is called relaxation time and it has its origin from statistical mechanics of the electrons distribution for heat diffusion [54]. Heat diffusion stems from the same laws underlying the particle diffusion process. Therefore, despite the fact that the Telegraph Equation in (14) was originally formulated for the case of heat transfer, it can also be applied to the diffusion of particle concentration.

We propose to model the processing of the system input  $r(\bar{x}, t)|_{\bar{x}=\mathbf{0}}$  through the linear system denoted by the impulse response  $g_d(\bar{x}, t)$ . This process is a convolution operation performed both in time  $t$  and in space  $\bar{x}$ , expressed as follows:

$$c(\bar{x}, t)|_{\bar{x} \in S} = \int_S \int_{t'=0}^{+\infty} r(\bar{x}', t') g_d(\bar{x}' - \bar{x}, t - t') dt' d\bar{x}', \quad (15)$$

where the system input is the particle concentration rate at transmitter  $r(\bar{x}, t)|_{\bar{x}=\mathbf{0}}$  and the system output is the particle concentration value  $c(\bar{x}, t)$  at any space location  $\bar{x} \in S$  and at any time instant  $t$ .

Since the input particle concentration rate is a non-zero value only at the transmitter, this can be also seen as the multiplication of a Dirac delta in the space  $S$  by the particle concentration rate  $r_T(t)$  at the transmitter. Therefore, the convolution operation is performed only in time, expressed as follows:

$$c(\bar{x}, t)|_{\bar{x} \in S} = \int_{t'=0}^{+\infty} r_T(t') g_d(\bar{x}, t - t') dt', \quad (16)$$

The impulse response  $g_d(\bar{x}, t)$  of the system is the Green's function [55] (the propagator) of the diffusion process occurring between the transmitter and any other location  $\bar{x}$  in the space  $S$ . The Green's function is in this case the diffusion process response to a particle concentration rate at the transmitter given by a Dirac delta in time ( $r_T(t) = \delta(t)$ ). In order

to compute the function  $g_d(\bar{x}, t)$  we study the equations governing the diffusion process and the physical conditions constraining their validity.

The Green's function  $g_d(\bar{x}, t)$  of the Telegraph Equation in (14), which is analytically equivalent to the wave equation in a lossy medium [56, 57], is the analytical solution for the concentration evolution in space and time when  $r(\bar{x}, t)$  is a Dirac delta function both in time  $t$  and in space  $S$ :  $r(\bar{x}, t) = \delta(\bar{x})\delta(t)$ . It is analytically expressed as

$$g_d(\bar{x}, t) = U(t - \|\bar{x}\|/c_d) e^{-\frac{t}{2\tau_d}} \frac{\cosh\left(\sqrt{t^2 - (\|\bar{x}\|/c_d)^2}\right)}{\sqrt{t^2 - (\|\bar{x}\|/c_d)^2}}, \quad (17)$$

where  $\|\bar{x}\|$  is the distance from the transmitter and  $c_d$  is the wavefront speed, defined as  $c_d = \pm\sqrt{D/\tau_d}$ , where  $D$  is the diffusion coefficient,  $\tau_d$  is the relaxation time in (14) and  $U(\cdot)$  is the step function.

The FTTF  $\tilde{\mathbf{B}}(f)$  of the diffusion process is the Fourier transform of the Green's function  $g_d(\bar{x}, t)$  from (17), expressed as

$$\tilde{\mathbf{B}}(f) = \int_{-\infty}^{\infty} g_d(\bar{x}_R, t) e^{-j2\pi ft} dt, \quad (18)$$

where  $\bar{x}_R$  is the receiver location. The values of  $\tilde{\mathbf{B}}(f)$  are computed numerically.

The *normalized gain*  $\Gamma_{\mathbf{B}}(f)$  of the diffusion process is the magnitude  $|\tilde{\mathbf{B}}(f)|$  of the FTTF  $\tilde{\mathbf{B}}(f)$  normalized by its maximum value  $\max_f(|\tilde{\mathbf{B}}(f)|)$ , which is numerically computed from (18):

$$\Gamma_{\mathbf{B}}(f) = \frac{|\tilde{\mathbf{B}}(f)|}{\max_f(|\tilde{\mathbf{B}}(f)|)}. \quad (19)$$

The *delay*  $\tau_{\mathbf{B}}(f)$  of the diffusion process is

$$\tau_{\mathbf{B}}(f) = -\frac{d\phi_{\mathbf{B}}(f)}{df}, \quad (20)$$

where  $\phi_{\mathbf{B}}(f)$  is the phase of the FTTF  $\tilde{\mathbf{B}}(f)$ , as

$$\phi_{\mathbf{B}}(f) = \arctan\left(\frac{\text{Im}(\tilde{\mathbf{B}}(f))}{\text{Re}(\tilde{\mathbf{B}}(f))}\right), \quad (21)$$

which stems from the real part  $\text{Re}(\tilde{\mathbf{B}}(f))$  and the imaginary part  $\text{Im}(\tilde{\mathbf{B}}(f))$ , numerically computed from (18).

### 3.5 The Reception Process

Through the reception process, the receiver senses the particle concentration  $c_R(t)$ , and accordingly modulates the output signal  $s_R(t)$  of the physical end-to-end model. The reception process, sketched in Figure 7, is based on the chemical theory of the ligand-receptor binding [49], and on the following assumptions:

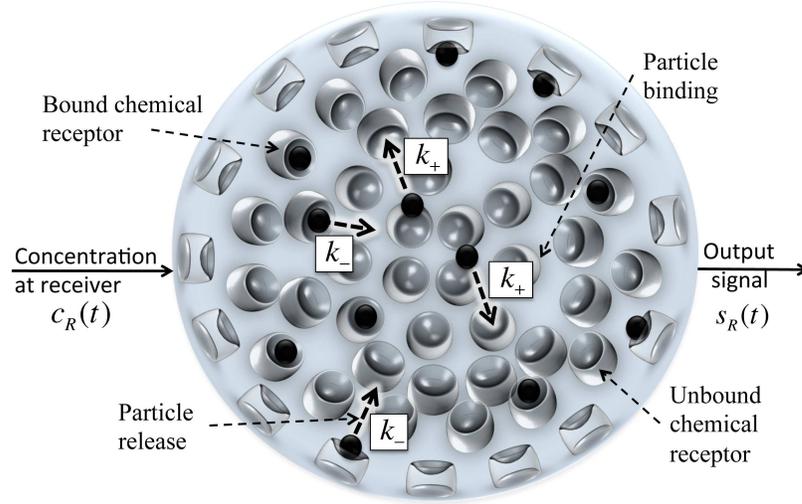


Figure 7. Representation of the reception process.

- The measure of the particle concentration takes place inside the *receptor space*. The receptor space has a spherical shape of radius  $\rho$ .
- The input *particle concentration*  $c_R(t)$  is considered homogeneous inside the receptor space and equal to the particle concentration value at the receiver location.
- The reception is realized by means of  $N_R$  *chemical receptors*.
- The chemical receptors are assumed to homogeneously occupy the volume of the receptor space.
- Each chemical receptor, at the same time instant  $t$ , is exposed to the same particle concentration  $c_R(t)$ .

- The chemical receptors, when exposed to the particle concentration  $c_R(t)$ , can remain in their state, namely, *bound* or *unbound*, or they can change their state by undergoing two possible chemical reactions: the *particle binding* reaction if the receptor was previously unbound, or the *particle release* reaction if the chemical receptor was previously bound to a particle.
- The particle binding occurs with a rate  $k_+$ , while the particle release occurs with the rate  $k_-$ .

When a particle concentration  $c_R(t)$  is present at time  $t$  inside the reception space, the chemical receptors change their states accordingly. The trend is to reach a ratio between the number of bound chemical receptors over the not-bound chemical receptors proportional to  $c_R(t)$  itself. The concentration of chemical receptors inside the receptor space is the number of chemical receptors  $N_R$  divided by the reception space size  $\rho$ . Therefore, we define  $\hat{s}_R(t)$  as the desired of the number of bound chemical receptors over the number of not-bound chemical receptors inside the receptor space with the following expression:  $\hat{s}_R(t) = \rho/N_R c_R(t)$ .

We define  $n_c(t)$  as the number of bound chemical receptors (complexes) inside the emission space at time  $t$ . The time differential in the number of complexes  $dn_c(t)dt$  inside the reception space is equal to the number of receptors  $N_R$  multiplied by the time differential of the system output signal  $ds_R(t)dt$ :

$$dn_c(t)dt = N_R ds_R(t)dt \quad (22)$$

The reception process described above is inspired by cellular systems from biology. According to these systems, the chemical receptors are models of the transmembrane receptors [2] embedded in the plasma membrane of living cells and involved in the signal transduction process. In this section of the report, we do not model the location of the chemical receptors as if they were on the plasma membrane, but we place them homogeneously in the receptor space. This allows us to simplify the treatment related to the

chemical changes which ultimately lead to the signal transduction. The signal transduction in bio-signaling involves the conversion of a chemical change (e.g., the change in concentration of the chemical components in the surrounding environment) into information. From this point of view, the molecules of the biochemical components are modeled by the particles and the biological surrounding environment is the receptor space.

The reception process is modeled through a series-resistor-capacitor (RC) circuit [50]. This circuit is shown in Figure 8, where  $V_{in}(t)$  is the input voltage, function of the time  $t$ ,  $R_r^{ch}$  is the resistance value of the resistor that is active during the capacitance-charging phase,  $R_r^{dis}$  is the resistance value of the resistor that is active during the capacitance-discharging phase,  $C_r$  is the capacitance value, and  $I_{out}(t)$  is the output current, equal to the current  $I_r(t)$  that is charging or discharging the capacitance. In the following, it is assumed that  $R_r^{ch} \approx R_r^{dis}$ . Under this assumption, the diodes can be removed, and a single resistor  $R_r$  is considered.

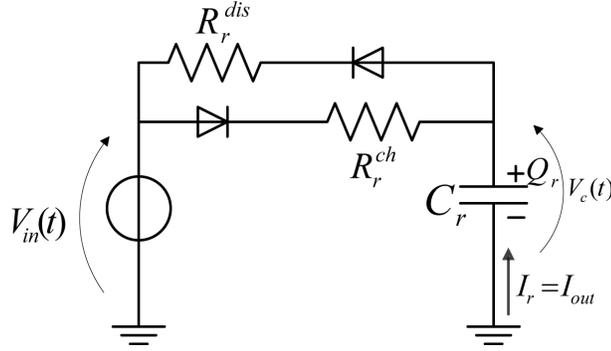


Figure 8. RC-circuit model of the reception process

From the electrical circuit theory [50], the Fourier transform of the transfer function (FTTF) [51] of the RC circuit  $\mathbf{H}_{RC}(f)$  between the input voltage  $V_{in}(t)$  and the output current  $I_{out}(t)$  is

$$\mathbf{H}_{RC}(f) = \frac{\tilde{\mathbf{I}}_{out}(f)}{\tilde{\mathbf{V}}_{in}(f)} = \frac{j2\pi f C_r}{1 + j2\pi f R_r C_r}, \quad (23)$$

where  $\tilde{\mathbf{V}}_{in}(f)$  and  $\tilde{\mathbf{I}}_{out}(f)$  are the Fourier transforms [51] of the input voltage  $V_{in}(t)$  and output current  $I_{out}(t)$ , respectively.

In this scheme, we identify the desired ratio  $\hat{s}_R(t)$  of the number of bound chemical receptors over the number of not-bound chemical receptors inside the receptor space with the input voltage  $V_{in}(t)$  of the RC circuit. The system output signal  $s_R(t)$  is equal to the output current  $I_{out}$ . The number  $n_c$  of bound chemical receptors inside the receptor space is considered as the charge  $Q_r$  stored in the capacitor at time  $t$ . The time differential  $dQ_r(t)dt$  of the charge stored in the capacitor is equal to the capacitor value  $C_r$  multiplied by the time differential  $dV_c(t)dt$  of the output voltage, expressed as follows:

$$dQ_r(t)dt = C_r dV_c(t)dt . \quad (24)$$

The similarity of (22) and (24) implies that  $N_R$  is the total number of chemical receptors inside the receptor space can be considered equal to the capacitor value  $C_r$ , expressed as

$$C_r = N_R . \quad (25)$$

A number of bound chemical receptors inside the reception space generates a proportional ratio between bound chemical receptors and not-bound chemical receptors as in an ideal capacitor, the stored charge  $Q_r(t)$  generates a proportional voltage  $V_c(t)$ . The proportionality constants are  $N_R$  and  $C_r$ , respectively.

The rates of increase and decrease in the number of bound chemical receptors  $V_c(t)$  are related to the rate constant  $k_+$  of binding reaction and the rate constant  $k_-$  of release reaction, respectively. We assume that the probability of a chemical receptor to build/break a complex and to capture/release a particle is affected by the ratio  $V_c(t)$  between the bound chemical receptors and not-bound chemical receptors itself. When  $V_c(t)$  increases, the probability of capturing a particle decreases. When  $V_c(t)$  decreases, the release rate decreases. We assume to have a linear relation, which leads to the following expression:

$$\frac{dn_c}{dt} = (V_{in}(t) - V_c(t))k . \quad (26)$$

When  $V_{in}(t) > V_c(t)$ , then  $k = k_+$ , whereas, when  $V_{in}(t) < V_c(t)$ , then  $k = -k_-$ .

We relate the number of complexes  $n_c$  to the capacitor charge  $Q_r(t)$  stored in the capacitor.  $dn_c(t)/dt$  is therefore related to the capacitor charge current  $I_r(t)$  at time  $t$ , equal to  $dQ_r(t)/dt$ , expressed as follows:

$$I_r(t) = \frac{(V_{in}(t) - V_c(t))}{R_r} \Rightarrow R_r = \frac{1}{k}, \quad (27)$$

where we assume that  $k = k_- \simeq k_+$

From the electrical-circuit theory [50], the FTTF  $\tilde{\mathbf{C}}(f)$  of the transmitter can be expressed as function of the FTTF of the RC circuit as follows:

$$\tilde{\mathbf{C}}(f) = \frac{\rho}{N_R} \mathbf{H}_{RC}(f) = \frac{j2\pi f C_r \rho}{N_R(1 + j2\pi f R_r C_r)}, \quad (28)$$

where  $\tilde{\mathbf{I}}_{in}(f)$  and  $\tilde{\mathbf{I}}_{out}(f)$  are the Fourier transforms [51] of the input voltage  $I_{in}(t)$  and output voltage  $I_{out}(t)$ , respectively.

The *normalized gain*  $\Gamma_C(f)$  of the reception process is the magnitude  $|\tilde{\mathbf{C}}(f)|$  of the FTTF  $\tilde{\mathbf{C}}(f)$  normalized by its maximum value  $\max_f(|\tilde{\mathbf{C}}(f)|)$ , which becomes  $\rho/(N_R R_r)$  from (28). This normalized gain has the following expression:

$$\Gamma_C(f) = \frac{|\tilde{\mathbf{C}}(f)|}{\max_f(|\tilde{\mathbf{C}}(f)|)} = \frac{2\pi f R_r C_r}{\sqrt{(1 + (2\pi f R_r C_r)^2)}}. \quad (29)$$

The *delay*  $\tau_C(f)$  of the reception process is computed similarly to (9), where the phase of the FTTF of (28) is applied in place of the phase  $\phi_A(f)$ , expressed as

$$\phi_C(f) = \arctan\left(\frac{\text{Im}(\tilde{\mathbf{C}}(f))}{\text{Re}(\tilde{\mathbf{C}}(f))}\right) = \arctan\left(\frac{1}{2\pi f R_r C_r}\right), \quad (30)$$

which is computed from the real part  $\text{Re}(\tilde{\mathbf{C}}(f))$  and the imaginary part  $\text{Im}(\tilde{\mathbf{C}}(f))$  of the FTTF  $\tilde{\mathbf{C}}(f)$ .

### 3.6 End-to-end Normalized Gain and Delay

The end-to-end *normalized gain*  $\Gamma_T(f)$  is computed by multiplying the normalized-gain contributions coming from the three processes, expressed as

$$\Gamma_T(f) = \Gamma_A(f) \cdot \Gamma_B(f) \cdot \Gamma_C(f), \quad (31)$$

where  $\Gamma_A(f)$  is obtained from (8), (4) and (6);  $\Gamma_B(f)$  is numerically computed from (18) and (17);  $\Gamma_C(f)$  is obtained from (29), (27) and (25). The end-to-end *delay*  $\tau_T(f)$  is obtained by summation of the delay contributions coming from the three processes, expressed as

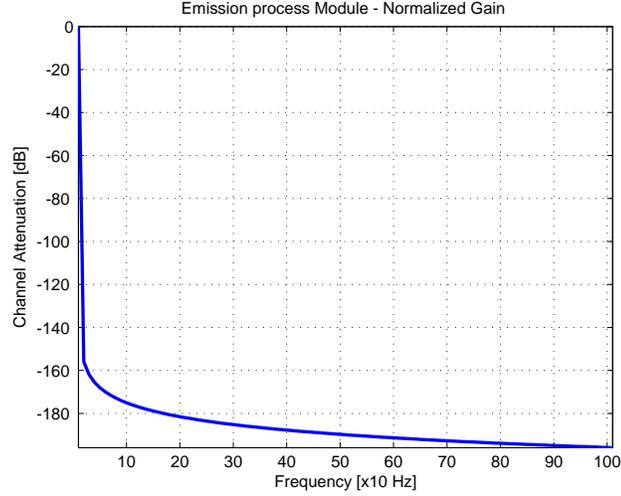
$$\tau_T(f) = \tau_A(f) + \tau_B(f) + \tau_C(f) , \quad (32)$$

where  $\tau_A(f)$  is obtained from (9), (10), (4) and (6);  $\tau_B(f)$  is numerically computed from (18) and (17);  $\tau_C(f)$  is obtained from (9), (30), (27) and (25).

### 3.7 Numerical Results

In this section, we show the numerical results in terms of *normalized gain and delay for each process*, as well as for the *overall end-to-end model*. The frequency spectrum considered in these results ranges from 0 Hz to 1 kHz. Although we believe that there is no biological justification for taking into account this frequency range, we are expecting to study networks of new devices which will be able to exploit the MC end-to-end model by using various modulation techniques, even the ones not used by biological entities. For this, we believe that the results in this frequency range could help the future development of nano-scale communication systems. We do not consider a wider frequency range since the results up to 1 kHz already show clearly the trend of the end-to-end model attenuation and delay as functions of the frequency.

The normalized gain  $\Gamma_A(f)$  for the emission process **A**, shown in Figure 9, is computed from (8), (4) and (6), for a frequency spectrum from 0 Hz to 1 kHz and a diffusion coefficient  $D \sim 10^{-6} m^2 sec^{-1}$  of calcium molecules diffusing in a biological environment (cellular cytoplasm, [58]). The normalized gain  $\Gamma_A(f)$  shows a non-linear behavior with respect to the frequency, as expected from an RC circuit. The curves for the normalized gain in Figure 9 show the maximum value 1 (0 dB) at the frequency 0Hz and they monotonically decrease as the frequency increases and approaches 1 Hz. This phenomenon can be explained considering that if the frequency of the end-to-end model input signal  $s_T(t)$



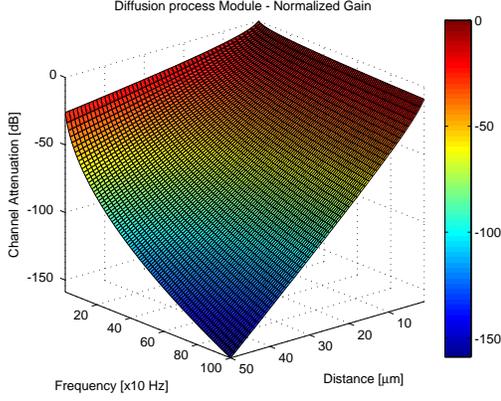
**Figure 9. The normalized gain for the emission process A.**

increases, the resulting modulated particle concentration rate  $r_T(t)$  decreases in its magnitude. This is due to the fact that the particle mobility in the diffusion process between the inside and the outside of the transmitter is constrained by the diffusion coefficient. The higher is the diffusion coefficient, the faster is the diffusion process given a value for the particle concentration gradient between the inside and the outside of the transmitter.

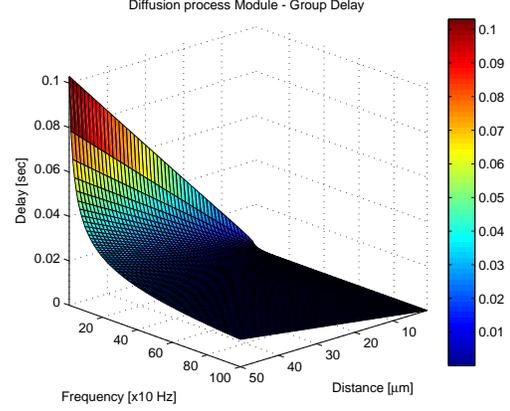
The delay  $\tau_A(f)$  for the emission process **A**, obtained from (9), (10), (4) and (6), shows a constant zero value in the frequency range from 0 Hz to 1 kHz and, for this reason, we omitted its plot here. Consequently, the transmitter does not distort any input signal having a bandwidth contained in the analyzed frequency range.

The normalized gain  $\Gamma_B(f)$  for the diffusion process **B**, shown in Figure 10, is numerically computed from (19), (18) and (17) for a transmitter-receiver distance from 0  $\mu\text{m}$  to 50  $\mu\text{m}$  and a frequency spectrum from 0 Hz to 1 kHz. The diffusion coefficient is the one of calcium molecules diffusing in a biological environment (cellular cytoplasm, [58])  $D \sim 10^{-6} \text{ m}^2\text{sec}^{-1}$ . The relaxation time  $\tau_d$  from (14) is set approximatively to the relaxation time computed for water molecules:  $\tau_d \sim 10^{-9}$  sec. The normalized gain  $\Gamma_B(f)$  shows a non-linear behavior both with respect to the distance  $d$  and the frequency  $f$ . The maximum

value of the normalized gain (0 dB) is at the transmitter location (distance = 0) and for the frequency  $f = 0$ . As the frequency increases, the normalized gain decreases monotonically. The behavior with respect to the distance from the transmitter is monotonically decreasing.



**Figure 10.** The normalized gain for the diffusion process  $B$ .

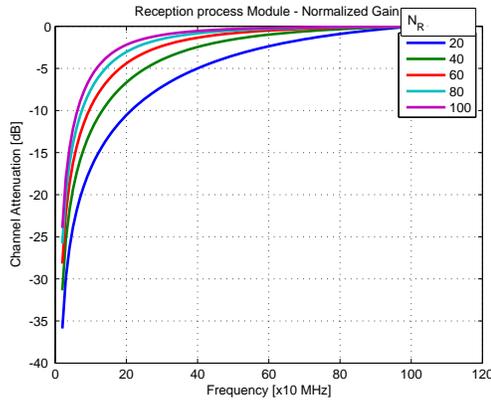


**Figure 11.** The group delay for the diffusion process  $B$ .

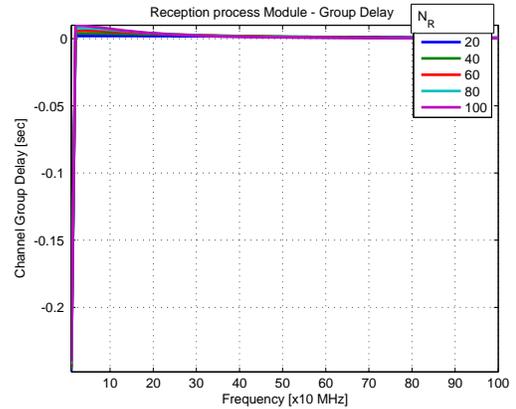
The delay  $\tau_B(f)$  for the diffusion process  $B$ , shown in Figure 11, is numerically computed from (20), (21), (18) and (17). The delay  $\tau_B(f)$  is shown both with respect to the distance and the frequency. For low frequency values, the delay is non-linear with respect to the distance from the transmitter. Therefore, the particle diffusion process  $B$  has a dispersive behavior in the frequency range from 0 Hz to 1 kHz and, consequently, the signal propagating through the molecular diffusion process can be distorted.

The normalized gain  $\Gamma_C(f)$  for the reception process  $C$ , shown in Figure 12, is obtained from (29), (27) and (25), for a variable number of receptors  $N_R$  from 20 to 100, a frequency spectrum from 0 Hz to 1 MHz, rate constants  $k_+ = k_- = 10^8 \text{ M}^{-1}\text{sec}^{-1}$  (see [59]), and  $\rho = 10 \mu\text{m}$ . The reception process normalized gain shows a non-linear behavior with respect to the frequency, as expected from an RC circuit. Each different curve is related to a different value in the number of receptors  $N_R$ . All the curves show the maximum value 1 at the frequency 0 Hz. The normalized gain monotonically increases as the frequency

increases. This phenomenon can be explained considering that if the frequency of the particle concentration  $c_R(t)$  increases, the resulting output signal  $s_R(t)$  increases its magnitude. This is due to the fact that the particle receptors are constrained by the rate of release and binding in the sensing of the particle concentration  $c_R(t)$ . The curves related to lower values of  $N_R$  show lower values of normalized gain throughout the frequency spectrum range. A higher number of receptors  $N_R$  requires a higher number of molecules released or captured in order to reach a desired ratio of bound receptor to non bound receptors. The number of receptors  $N_R$  is also related to the precision of the particle concentration measurement. Then the higher is the number  $N_R$  of receptors inside the receptor space  $S_r$ , the smaller is the minimum concentration variation  $dc_r(t)$  sensed by the reception process and translated into a variation in the system output signal  $s_R(t)$ .



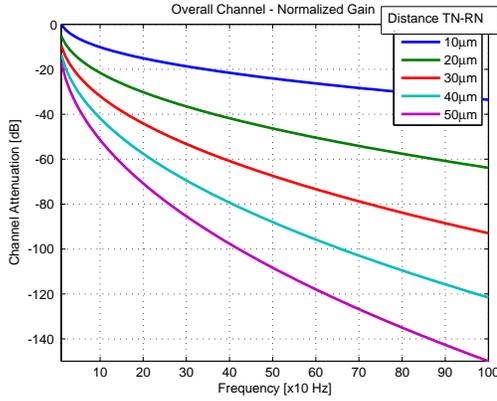
**Figure 12.** The normalized gain for the reception process  $C$ .



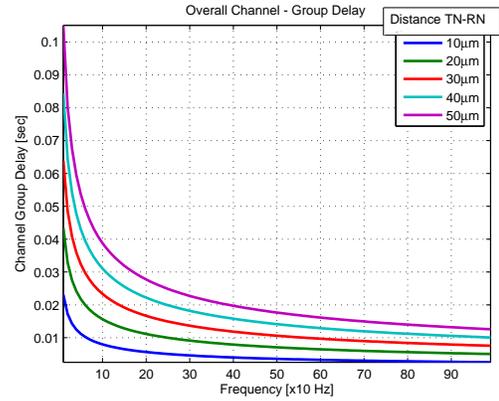
**Figure 13.** The group delay for the reception process  $C$ .

The delay  $\tau_C(f)$  for the reception process  $C$ , shown in Figure 13, is obtained from (30), (27) and (25). The delay  $\tau_C(f)$  curves are shown in Figure 13 with the same input parameters as before. For every curve, the delay has a non-linear behavior with respect to frequency. This means that the shape of the system output signal  $s_R(t)$  is distorted with respect to the particle concentration  $c_R(t)$ . This behavior is enhanced for higher values in the number  $N_R$  of receptors.

The normalized gain  $\Gamma_T(f)$  for the end-to-end model  $T$ , shown in Figure 14, is computed from (31) for a transmitter-receiver distance from  $0 \mu\text{m}$  to  $50 \mu\text{m}$ . The number of receptors is assumed to be  $N_R = 10$ , since it corresponds to the lowest normalized gain for the reception process, while maintaining a reasonable number of receptors at the receiver. The end-to-end normalized gain shows a non-linear behavior with respect to the frequency. Each different curve is related to a different value of the transmitter-receiver distance. All the curves show the maximum value 1 at the frequency 0Hz. The normalized gain monotonically decreases as the frequency increases. If the frequency of the end-to-end model input signal  $s_T(t)$  increases, the resulting output signal  $s_R(t)$  decreases its magnitude. The curves related to higher values of the transmitter-receiver distance show lower values of normalized gain throughout the frequency spectrum range.



**Figure 14.** The normalized gain for the reception process  $T$ .



**Figure 15.** The group delay for the end-to-end model  $T$ .

The delay  $\tau_T(f)$  for the end-to-end model  $T$ , shown in Figure 15, is computed from (32). The delay  $\tau_T(f)$  curves are shown in Figure 15, one for each value of the transmitter-receiver distance. For every curve, each frequency is delayed by a different time. Consequently, the shape of the system output signal  $s_R(t)$  is distorted with respect to the end-to-end model input signal  $s_T(t)$ . This behavior is enhanced for higher values of the transmitter-receiver distance.

### **3.8 Conclusion**

In this chapter of the Ph.D. thesis, a physical end-to-end model for the diffusion-based MC is proposed. For this, a basic diffusion-based MC system design is provided, which aims at an interpretation of the diffusion-based MC, both in terms of molecule emission/reception and molecule propagation. In the provided diffusion-based MC system design, the physical end-to-end model is studied as the composition of three subsequent processes, namely, the particle emission, the particle diffusion, and the particle reception. The normalized gain and delay are computed as functions of the frequency and the transmission range for the three physical processes, as well as for the overall end-to-end model. Numerical results are provided to evaluate the normalized gain and delay of the diffusion-based MC system for several different values of the system frequency and the transmission range.

Typical communication engineering paradigms can be applied to this model in order to study the end-to-end behavior in terms of noise, capacity and throughput. Moreover, in light of the results, several classical modulation schemes could be studied when information is sent over this physical end-to-end model. However, this analysis does not account for stochastic phenomena present in the aforementioned processes that can act as noise on the information signal. In the next chapter, the models for the three processes are further refined by providing an analysis of the noise sources affecting the diffusion-based MC.

# CHAPTER 4

## NOISES IN DIFFUSION-BASED MOLECULAR COMMUNICATION

### 4.1 Motivation and Related Work

The proper study and characterization of the noise is one of the main challenges in the information theoretic analysis of diffusion-based Molecular Communication (MC). The noise in a communication system is defined as an unwanted and unavoidable random component that affects the information-bearing signal [60]. The modeling of the noise sources is fundamental to design the components of a communication system, and increase the probability of correct exchange of information between the transmitter and the receiver. From the point of view of communication engineering, the most useful characterization of the noise sources is achieved, whenever possible, through a stochastic modeling, where the generated noise signal is interpreted as the output of random processes with known statistical parameters. The objective of this chapter of the Ph.D. thesis is to identify the most relevant noise sources affecting diffusion-based MC, and provide stochastic models with a proven physical validity.

Most of the contributions from the literature to the noise analysis for diffusion-based MC are mainly based on the results of simulations, and do not provide stochastic models for the noise sources in terms of random processes. As an example, in [36] the results of simulations show a noise for the diffusion-based MC which follows a non-Gaussian statistics, although the analytical model for this statistics is not investigated. In [29] the noise effects on the diffusion-based MC are resulting only from simulation and there is no analytical model of diffusion-based noise and no stochastic study of its underlying physical phenomena. Moreover, in [29] there is no specific analysis of the noise sources which affect the reception side of a diffusion-based MC system. In [43], the diffusion-based MC reception noise is analyzed in terms of probability of having erroneous digital reception, under

the assumption of a binary squared pulse code modulation signal. As a consequence, the work in [43] addresses the noise analysis for a diffusion-based MC system having specific characteristics in terms of modulation scheme and type of transmitted messages.

Contributions from the biochemistry literature provide descriptions of some physical processes underlying the noise sources in diffusion-based MC systems. Seminal works in biochemistry, such as [61], analyzed how free space diffusion of molecules impairs the proper measurement of the molecule concentration. A more recent contribution to the physical analysis of molecule diffusion and reception in biochemical signaling can be found in [62]. However, these contributions tend to focus on the explanation of natural phenomena, and do not provide suitable models for MC engineering. The work in [63] stems, on the contrary, from the simulation of a biological signal transduction mechanism and its associated noise using tools from communication engineering. However, the analysis of the system is limited to a numerical evaluation of the simulation results using communication engineering parameters (e.g., the Signal to Noise Ratio, SNR). No stochastic models are provided in [63] for the noise sources, but the results are coming from numerical simulations. In [64], the authors develop only a preliminary information theoretic model applied to the study of intracellular communication with the diffusion of calcium ions.

In this chapter of the Ph.D. thesis, three noise sources affecting diffusion-based MC are identified, namely, the sampling noise, the counting noise, and the ligand-receptor-binding noise, and they are related to the transmitter, the signal propagation in the channel, and the receiver, respectively. These noise sources are modeled in a twofold fashion: the physical model provides a mathematical analysis of the physical processes which generate the noise, while the stochastic model aims at capturing those physical processes through statistical parameters. The physical model contains all the physical variables which contribute to the generation of the noise. The stochastic model summarizes the noise generation using random processes and their associated parameters. Sets of noise data realizations are generated through simulation of the physical model. These sets of noise data are then used

to test the stochastic model ability to capture the behavior of the physical processes which generate the noise.

## 4.2 Definition of the Noise Sources in Diffusion-based MC

Three noise sources, which affect the physical end-to-end model defined in Chapter 3, are here identified and studied, namely, the particle sampling noise, the particle counting noise, and the ligand-receptor-kinetics noise. In the following, each noise source is defined with reference to the block scheme in Figure 16.

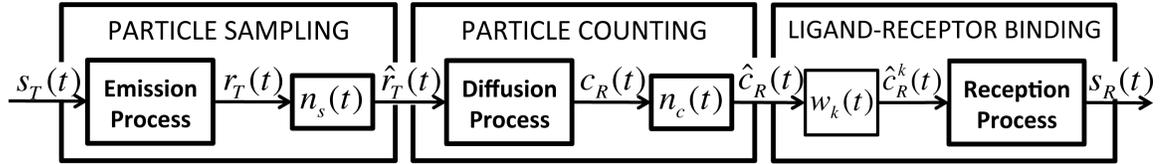


Figure 16. Block scheme of the end-to-end model that includes the noise sources.

The PARTICLE SAMPLING is related to the **Emission Process** at the transmitter. The particle sampling noise is expressed as  $n_s(t)$ . The effect of  $n_s(t)$  is an unwanted perturbation on the output of the emission process  $r_T(t)$ , which results in  $\hat{r}_T(t)$ :

$$r_T(t) \rightarrow \hat{r}_T(t) \quad (33)$$

The particle sampling noise is generated by the PARTICLE SAMPLING, which occurs when the particle concentration rate  $\hat{r}_T(t)$  is being modulated through the emission of the particles. The noise effects arise from the discreteness of the particles that compose the particle concentration rate  $\hat{r}_T(t)$ . The particle concentration rate  $r_T(t)$  in output from the emission process is caused by a particle flux between the transmitter and the external space. Given the discreteness of the particles, the particle concentration rate  $r_T(t)$  is sampled by the particles themselves, resulting in the particle concentration rate  $\hat{r}_T(t)$ . Further details on the analysis for this type of noise are provided in Section 4.3.

The PARTICLE COUNTING is related to the signal propagation due to the **Diffusion Process**. The particle counting noise is expressed as  $n_c(t)$ . The effect of  $n_c(t)$  is an unwanted

perturbation on the output of the diffusion process  $c_R(t)$ , which results in  $\hat{c}_R(t)$ :

$$c_R(t) \rightarrow \hat{c}_R(t) \quad (34)$$

The particle counting noise occurs when the particle concentration value is being measured at the receiver location (PARTICLE COUNTING) and it is due to the randomness in the movement and to the discreteness of the particles. The particle concentration  $c_R(t)$  at the receiver location is computed by counting the number of particles present in the reception space. Fluctuations and imprecisions in counting the particles impair the proper computation of the concentration  $c_R(t)$ . The actual computed concentration  $\hat{c}_R(t)$  differs from  $c_R(t)$ . The analysis for this type of noise is provided in Section 4.4.

The LIGAND-RECEPTOR BINDING is related to the **Reception Process**. The ligand-receptor-binding noise affects the molecular receiver due to random fluctuations in the LIGAND-RECEPTOR BINDING process. Due to the effect of this noise contribution, denoted by  $w_k(t)$ , the particle concentration  $c_R(t)$  is subject to an unwanted perturbation, resulting in  $\hat{c}_R^k(t)$ . As a consequence, this perturbation propagates to the output signal  $\hat{s}_R^k(t)$  after the reception process:

$$c_R(t) \rightarrow \hat{c}_R^k(t) \rightarrow \hat{s}_R^k(t) \quad (35)$$

Further details on the analysis for this type of noise are provided in Section 4.5.

In the following, the analysis of each noise source, which results in both a *physical model* and a *stochastic model*, is detailed. With the former the mathematical expression of the physical process underlying the noise source is obtained, while with the latter the noise source behavior is modeled through the use of statistical parameters.

### 4.3 The Particle Sampling Noise

The model of the emission process provided in Section 3.3 of Chapter 3 does not take into account the discrete nature of the particles. In this chapter, an additional assumption is considered for the particle-emission process:

- The particle flux through the emission gaps in the spherical boundary is composed of discrete particles.

As a result, the relation between the input signal  $s_T(t)$  and the resulting particle-concentration rate, denoted by  $\hat{r}_T(t)$ , is no longer a continuous function as in Section 3.3. During the emission process, as shown in Figure 4, single particles flowing through the emission gaps contribute to the concentration rate  $\hat{r}_T(t)$  with a value  $k_n$  at discrete time instants  $t_n = t_1, t_2, \dots$ . These time instants are not equally spaced, as a consequence of the random nature of the particle motion through the emission gaps. Therefore, the resulting particle-concentration rate  $\hat{r}_T(t)$  assumes values equal to  $k_n$  at randomly-spaced time instants  $t_n$ , and it is zero for any other time instant. This particle-concentration rate is expressed as follows:

$$\hat{r}_T(t) = \sum_{n \in \mathbb{N}} \frac{k_n}{t_n - t_{n-1}} \delta(t - t_n) , \quad (36)$$

where  $\delta(\cdot)$  is a Dirac delta function. According to the Nyquist theorem [65], since the time instants  $t_n$  are randomly spaced, the continuous particle concentration rate  $r_T(t)$  can be reconstructed from the non-uniform sampled particle concentration rate  $\hat{r}_T(t)$  if the bandwidth of  $r_T(t)$  is limited up to frequency  $1/(2\langle t_n - t_{n-1} \rangle)$ , where  $\langle t_n - t_{n-1} \rangle$  is the average interval between two consecutive samples of  $\hat{r}_T(t)$ . As a consequence, given a fixed bandwidth for the system, the degradation caused by the particle sampling noise on the particle concentration rate in output at the transmitter depends on the average rate of the events of single particles flowing between the inner area and the outer area. This event rate corresponds to the particle concentration rate  $r_T(t)$  and the system bandwidth depends on the parameters as defined in Section 3.3. This result is confirmed through the stochastic model of the particle sampling noise, outlined in Section 4.3.2.

#### 4.3.1 The Physical Model

The physical model of the particle sampling noise is represented by the block scheme shown in Figure 17. The signal  $s_T(t)$  is the input of the emission-process block, whose output is the particle-concentration rate  $r_T(t)$ . The physical model of the particle sampling

noise  $n_s(t)$  takes as input the particle-concentration rate  $r_T(t)$  that the emission process would produce in output in the absence of noise. The particle sampling noise  $n_s(t)$  is composed of a decision block and a non-uniform sampler, which have as input the transmitter kinetic state  $\bar{S}_T(t)$ , and a divisor. The output of the particle sampling noise  $n_s(t)$  is the particle-concentration rate  $\hat{r}_T(t)$  affected by noise.

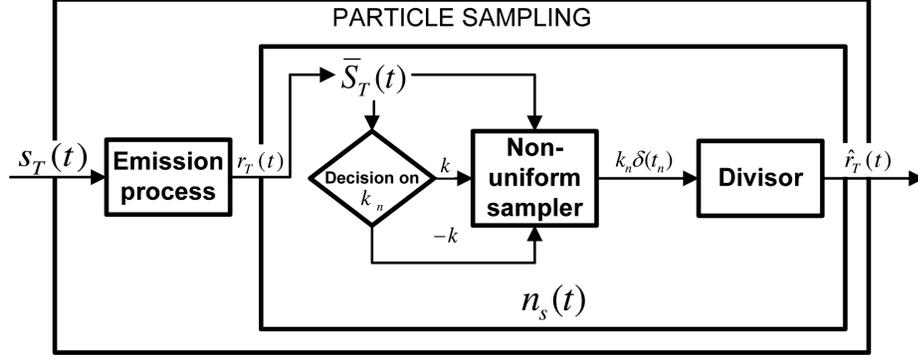


Figure 17. Block scheme of the physical model for the particle sampling noise.

The **transmitter kinetic state**  $\bar{S}_T(t)$ , as shown in Figure 18, is a set composed by the location  $\bar{x}_p(t)$  and the net velocity  $\bar{v}_p(t)$  of each particle  $p$  at time  $t$  present in the surrounding of the transmitter spherical boundary, expressed as

$$\bar{S}_T(t) = \{ \bar{x}_p(t), \bar{v}_p(t) \mid p = 1, \dots, P(t) \}, \quad (37)$$

where  $P(t)$  is the number of particles in the system and varies as a function of the time  $t$ . The net velocity  $\bar{v}_p(t)$  is here defined as the non-isotropic component of a particle speed, in contrast to the Brownian motion in free space which has isotropic components. In order to realistically simulate the transmitter kinetic state  $\bar{S}_T(t)$ , we consider two different contributions to the particle displacement, namely, the Brownian motion and the time integral of the particle net velocity from time instant  $t_0$  to time instant  $t$ . The time instant  $t_0$  corresponds to the beginning of the emission process. The expression of the particle location  $\bar{x}_p(t)$  is written as follows:

$$\bar{x}_p(t) = b_x(t) \hat{i} + b_y(t) \hat{j} + b_z(t) \hat{k} + \int_{t'=t_0}^t \bar{v}_p(t') dt', \quad (38)$$

where the Brownian motion components, namely,  $b_x(t)$ ,  $b_y(t)$  and  $b_z(t)$ , are random variables with normal distribution, zero mean value and variance equal to  $2D\delta t$ , according to the expression of the Wiener process [66], expressed as

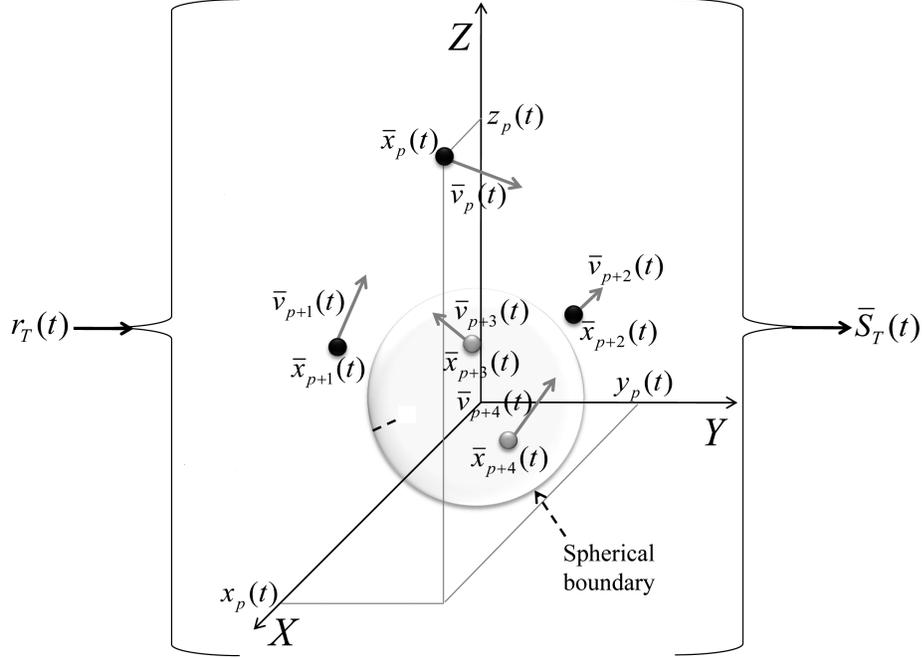
$$b_x(t), b_y(t), b_z(t) \sim \mathcal{N}(0, 2D\delta t), \quad (39)$$

along the versors of the cartesian axes, namely,  $\hat{i}$ ,  $\hat{j}$  and  $\hat{k}$ .  $D$  is the diffusion coefficient and  $\delta t$  is the simulation time step and it depends on how the transmitter kinetic state is sampled during the physical model simulation. The smaller is the time step  $\delta t$ , the closer is the simulation to the real physical phenomenon of particle diffusion. The value of the time step  $\delta t$  defines the time resolution with which we model events concerning particles changing their space area. According to the Nyquist theorem [65], if the value of the time step  $\delta t$  is smaller than  $1/(2B_{r_T})$ , where  $B_{r_T}$  is the bandwidth of the particle concentration rate  $r_T(t)$ , then we can have a perfect simulation of the sampling noise generation as it happens in reality. When the particle is located inside the inner area or the outer area, it is only subject to the Brownian motion. In these cases, the particle speed has only the isotropic components due to the Brownian motion in free space, and its net velocity  $\bar{v}_p(t)$  is equal to zero. When the particle is traversing an emission gap, its movement can only be outward in case of positive rate ( $r_T(t) > 0$ ) or inward in case of negative rate ( $r_T(t) < 0$ ) along the radius of the spherical boundary. In order to quantify the particle net velocity  $\bar{v}_p(t)$ , we consider that the particle concentration rate  $r_T(t)$  is given only by the contribution of the particles traversing the emission gaps. Given a particle concentration rate  $r_T(t)$ , the number of particles traversing the emission gaps in a unit time is given by the transmitter inner concentration  $c_T^{in}(t)$  in case of positive rate ( $r_T(t) > 0$ ) and by the transmitter outer concentration  $c_T(t)$  in case of negative rate ( $r_T(t) < 0$ ), multiplied by their average velocity. When they traverse the emission gap, the particle average velocity corresponds to the net velocity  $\bar{v}_p(t)$ . As a consequence, the particle net velocity  $\bar{v}_p(t)$  is proportional to the particle concentration rate  $r_T(t)$ , divided by the transmitter inner concentration  $c_T^{in}(t)$  in case of positive rate ( $r_T(t) > 0$ ), or divided by the transmitter outer concentration  $c_T(t)$  in case of

negative rate ( $r_T(t) < 0$ ), expressed as follows:

$$\bar{v}_p(t) = \begin{cases} 0 & \text{if } p \text{ in inner or outer} \\ \frac{r_T(t)}{c_T^{\text{in}}(t)\mathbf{1}_{r_T(t)>0} + c_T(t)\mathbf{1}_{r_T(t)<0}} \hat{y} & \text{if } p \text{ in emission gap} \end{cases}, \quad (40)$$

where  $\mathbf{1}_{(\text{condition})}$  is equal to 1 when *condition* is true and 0 otherwise.  $\hat{y}$  is the versor along the radius of the transmitter spherical boundary.



**Figure 18.** Graphical sketch of the transmitter kinetic state  $\bar{S}_T(t)$  at time  $t$ .  $\bar{S}_T(t)$  depends on the particle concentration rate  $r_T(t)$  in input through the expressions in (37) and (38).

The **decision** block assigns the value of  $k_n$  according to the transmitter kinetic state  $\bar{S}_T(t)$ .  $k_n$  is assigned a positive  $k$  value or a negative  $-k$  value according whether there is an event in the kinetic state  $\bar{S}_T(t)$  concerning a particle changing its space area, e.g., from the inner to the outer area, with contribution  $k$  to the rate, or from the outer to the inner area, with contribution  $-k$ , expressed as

$$k_n = \begin{cases} k & \text{if } \bar{S}_T(t) \subset \{\bar{x}_p(t), \bar{v}_p(t) | p \text{ from inner to outer}\} \\ -k & \text{if } \bar{S}_T(t) \subset \{\bar{x}_p(t), \bar{v}_p(t) | p \text{ from outer to inner}\} \end{cases}. \quad (41)$$

The value of  $k$  equals a contribution of one particle to the concentration at the transmitter location or, in other words, it is the constant difference in the particle concentration  $\hat{c}_T(t)$

from consecutive time instants  $t_n, t_{n-1}$ , expressed as

$$k = \hat{c}_T(t_n) - \hat{c}_T(t_{n-1}) . \quad (42)$$

The **non-uniform sampler** block samples at time instants  $t_n$ , which are functions of the transmitter kinetic state  $\bar{S}_T(t)$ . If, at time instant  $t_n$ , there is an event in the kinetic state  $\bar{S}_T(t_n)$  concerning a particle changing its space area, the non-uniform sampler block produces a Dirac impulse at  $t_n$ , with amplitude equal to the current value of  $k_n$ , output from the decision block, expressed as

$$k_n \delta(t - t_n) \text{ if } \bar{S}_T(t_n) \subset \{\bar{x}_p(t_n), \bar{v}_p(t_n) | p \text{ changes space area} \} . \quad (43)$$

The **divisor** block divides the output of the sampler by the time interval between the previous sample at  $t_{n-1}$  and the current sample, which is at  $t_n$ . As a consequence, the output of the divisor block for the time interval  $t_{n-1} < t < t_{n+1}$ , which corresponds to the particle concentration rate  $\hat{r}_T(t)$  affected by noise, is

$$\hat{r}_T(t) = \frac{k_n \delta(t - t_n)}{t_n - t_{n-1}} \text{ for } t_{n-1} < t < t_{n+1} . \quad (44)$$

For a time interval spanning from  $t = 0$  to  $t \rightarrow \infty$  the result is the expression introduced in (36).

Since it is not possible to always have the knowledge of the kinetic state of the system  $\bar{S}_T(t)$  due to the huge amount of information and to the randomness in the particle motion, we cannot analytically compute the value of  $\hat{r}_T(t)$  as function of  $r_T(t)$  from the physical model of the particle sampling noise. Using the physical model provided here, we can only simulate numerically the behavior of the particle sampling noise  $n_s(t)$ .

### 4.3.2 The Stochastic Model

The particle sampling noise can also have another formulation, through statistical parameters, that is suitable when theoretical studies require an analytical expression of the noise. In this formulation, the particle sampling noise  $n_s(t)$  is generated by a random process

$\tilde{n}_s(t)$ , whose contribution corresponds to the difference between the particle-concentration rate  $\hat{r}_T(t)$  affected by noise and the particle-concentration rate  $r_T(t)$  expected in the absence of noise, expressed as follows:

$$\tilde{n}_s(t) = \hat{r}_T(t) - r_T(t) . \quad (45)$$

In Figure 19, the main block scheme of the stochastic model for the particle sampling noise is shown. The random process  $\tilde{n}_s(t)$ , as it is proved in the following, depends on the value of the particle-concentration rate  $r_T(t)$ , output from the emission-process block, which receives the signal  $s_T(t)$  as input. The sum of the random process  $\tilde{n}_s(t)$  and the particle-concentration rate  $r_T(t)$  is the particle-concentration rate  $\hat{r}_T(t)$  affected by the particle sampling noise.

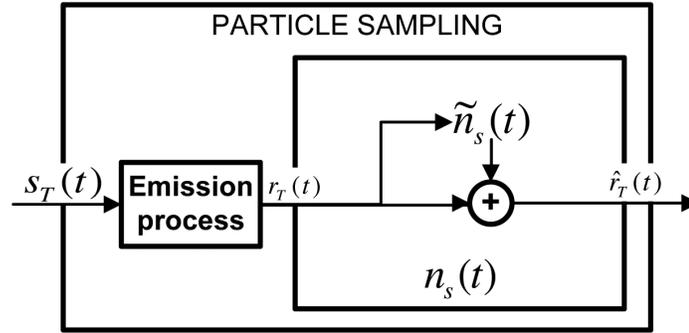


Figure 19. Block scheme of the stochastic model for the particle sampling noise .

In order to properly model the random process  $\tilde{n}_s(t)$  we consider the following assumptions:

- The outer particle concentration at the transmitter  $\hat{c}_T(t)$  increments/decrements its value whenever a single event concerning a particle changing its space area occurs.
- The probability of having two simultaneous events concerning particles changing their space area is zero. In other words, it is unlikely to have two particles crossing the spherical boundary of the transmitter at the same exact time instant. With reference to the physical model of the particle sampling noise from Section 4.3.1, this assumption

translates into the statement: the probability of having two samples from the non-homogeneous sampler at the same time instant is zero. In equation this becomes

$$Pr [t_n - t_{n-1} = 0] = 0 . \quad (46)$$

This assumption is justified by the independency of the Brownian components in the movement of different particles in the space. This assumption directly translates into the property of orderliness for the outer particle concentration  $c_T(t)$  increments/decrements. The property of orderliness states that the probability that the difference between outer particle concentrations  $\Delta$  time apart from each other is higher than the contribution  $k$  from a single particle, tends to zero as  $\Delta$  tends to zero, expressed as

$$\lim_{\Delta \rightarrow 0} Pr [|c_T(t + \Delta) - c_T(t)| > k] = 0 , \quad (47)$$

where  $k$  is defined through (42).

- An event concerning a particle changing its space area (passing through an emission gap) occurring after time  $t$  is independent of any event of the same kind occurring before time  $t$ . This assumption is justified by the property of the Wiener process underlying the particle Brownian motion of having independent increments. As stated in Section 3.3, particles are subject only to the contribution of the Brownian motion when they are located inside the inner area or the outer area. An event concerning a particle passing through an emission gap takes place whenever a particle, due to the Brownian motion, reaches the location of an emission gap: if there is a non-zero particle concentration rate in the outer area, the particle traverses the emission gap with net velocity  $\bar{v}_p(t)$ , given by (40). In other words, given a particle concentration rate in the outer area, which controls the average rate of occurrence of an event of this kind, the statistics of the event is solely dependent on the Brownian motion of the particles. As a consequence, the distribution of the time interval between an event at time  $t - \Delta t$  and another event at time  $t$  is independent from the distribution of the time

interval between an event at time  $t$  and an event at time  $t + \Delta t$ . The two distributions have the same expression from (39), as

$$\begin{aligned} Pr(b_x(t) - b_x(t - \Delta t) = x) &= \frac{1}{\sqrt{2\pi D\Delta t}} e^{-\frac{x^2}{2D\Delta t}} = \\ &= Pr(b_x(t + \Delta t) - b_x(t) = x) \quad , \end{aligned} \quad (48)$$

where  $b_x(t)$  is the motion component along the  $\hat{i}$  versor at time  $t$ ,  $D$  is the diffusion coefficient and  $\Delta t$  is positive. Equation (48) is valid also for the motion components  $b_y(t)$  and  $b_z(t)$  along the versors  $\hat{j}$  and  $\hat{k}$ , respectively. This implies that a particle motion from time  $t$  is independent from any motion of the particle occurred before time  $t$ . Being all the particle independent among each other, events concerning a change in the particle space area show the same independence. As a consequence, the events concerning particles changing their space area have the property of memorylessness.

- The occurrence rate of events concerning particles changing their space area is proportional to the flux of the particles between the inner area and the outer area. The flux of the particles is proportional to the expected particle concentration rate at the transmitter location  $r_T(t)$ .

Under these assumptions [66], the resulting outer particle concentration at the transmitter  $\hat{c}_T(t)$  is a double non-homogeneous Poisson counting process, whose rate of occurrence corresponds to the expected particle concentration rate  $r_T(t)$ . The distribution of the outer particle concentration  $\hat{c}_T(t)$  corresponds to a Poisson counting process with rate of occurrence  $r_T(t)$  whenever the particle concentration rate  $r_T(t)$  is positive. Whenever the particle concentration rate  $r_T(t)$  is negative,  $\hat{c}_T(t)$  is the negative of a poisson counting process with rate of occurrence  $-r_T(t)$ , expressed as

$$\hat{c}_T(t) \sim \begin{cases} \text{Poiss}(r_T(t)) & r_T(t) > 0 \\ -\text{Poiss}(-r_T(t)) & r_T(t) < 0 \end{cases} \quad , \quad (49)$$

When the emission process is subject to the particle sampling noise, the particle concentration rate at the transmitter location  $\hat{r}_T(t_n)$  corresponds to the first finite time difference of

the particle concentration  $\hat{c}_T(t)$ , which is step-wise and, therefore, not derivable, expressed as

$$\hat{r}_T(t_n) = \frac{\hat{c}_T(t_n) - \hat{c}_T(t_{n-1})}{t_n - t_{n-1}} . \quad (50)$$

Since the particle concentration  $\hat{c}_T(t)$  is a double non-homogeneous Poisson counting process, the particle concentration rate at the transmitter location  $\hat{r}_T(t)$  is the first finite time difference of a double non-homogeneous Poisson counting process, whose average value  $\langle \hat{r}_T(t) \rangle$ , where  $\langle \cdot \rangle$  denotes the ensemble average operator, has the same value as the rate of occurrence of the originating double Poisson counting process

$$\langle \hat{r}_T(t) \rangle = r_T(t) , \quad (51)$$

and whose autocorrelation is the expected squared particle concentration rate  $r_T^2(t)$  added to the expected particle concentration rate  $r_T(t)$  itself only for correlation lag  $l$  equal to 0, expressed as follows:

$$\langle \hat{r}_T(t) \cdot \hat{r}_T(t+l) \rangle = r_T^2(t) + r_T(t)\delta(l) , \quad (52)$$

where  $\delta(l)$  is a Dirac delta. Given (45), the random process  $\tilde{n}_s(t)$  has zero average value and its autocorrelation  $R_s(t, l)$  is equal to the expected particle concentration rate  $r_T(t)$  for correlation lag  $l$  equal to 0, expressed as

$$R_s(t, l) = \langle \tilde{n}_s(t) \cdot \tilde{n}_s(t+l) \rangle = r_T(t)\delta(l) . \quad (53)$$

Therefore the random process  $\tilde{n}_s(t)$  is white [66] and its mean squared value is the expected particle concentration rate  $r_T(t)$ , as

$$\langle \tilde{n}_s^2(t) \rangle = \langle \tilde{n}_s(t) \cdot \tilde{n}_s(t+l) \rangle|_{l=0} = r_T(t) . \quad (54)$$

The RMS of the perturbation  $\text{RMS}(\tilde{n}_s(t))$  on the expected particle concentration rate  $r_T(t)$  is equal to the square root of the expected particle concentration rate  $r_T(t)$ , expressed as

$$\text{RMS}(\tilde{n}_s(t)) = \sqrt{r_T(t)} . \quad (55)$$

According to the results in Section 3.3, the relation between the input signal  $T(t)$  and the particle concentration rate  $r_T(t)$  is expressed in the frequency ( $f$ ) domain as

$$\tilde{\mathbf{r}}_T(f) = \tilde{\mathbf{A}}(f)\tilde{\mathbf{T}}(f) , \quad (56)$$

where  $\tilde{\mathbf{T}}(f)$  and  $\tilde{\mathbf{r}}_T(f)$  are the Fourier transforms [51] of the system input signal  $T(t)$  and the particle concentration rate  $r_T(t)$  at the transmitter location, respectively.  $\tilde{\mathbf{A}}(f)$  is the Transfer Function Fourier Transform [51] (TFFT) of the transmitter module. The same relation in the time ( $t$ ) domain becomes

$$r_T(t) = a(t) * T(t) , \quad (57)$$

where  $*$  denotes the convolution operator [51],  $a(t)$  is the impulse response of the transmitter module and  $T(t)$  is the input signal. The formula for the RMS of the perturbation RMS( $\tilde{n}_s(t)$ ) on the signal  $\hat{r}_T(t)$  becomes

$$\text{RMS}(\tilde{n}_s(t)) = \sqrt{a(t) * T(t)} . \quad (58)$$

### 4.3.3 Numerical Results

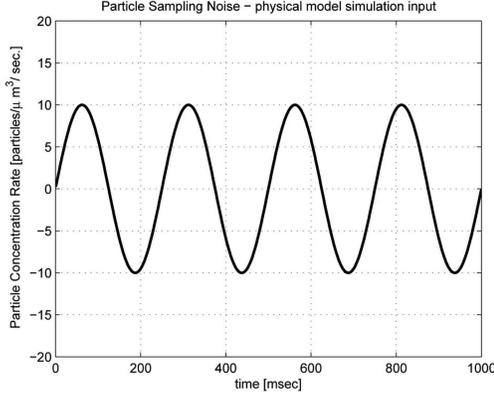
#### *Sampling Noise Simulations*

The simulations of the physical model for the particle sampling noise are computed by applying to the scheme in Figure 17 a sinusoidal signal in the particle concentration rate  $r_T(t)$ , expressed as

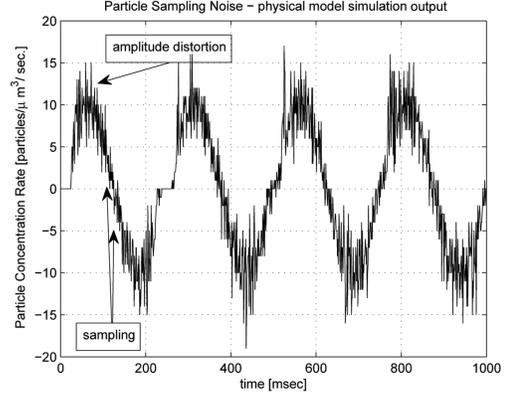
$$r_T(t) = A \sin(2\pi f_a t) , \quad (59)$$

where  $f_a$  is the frequency of the sinusoid in Hz,  $A$  is the value of the maximum particle concentration rate in particles  $\mu\text{m}^{-3}\text{sec}^{-1}$ , and  $t$  is the simulation time index in msec.

The input of the physical model simulation is a sinusoidal particle concentration rate  $r_T(t)$  with frequency  $f_a$  equal to 4 Hz and maximum particle concentration rate  $A$  of 10 particles  $\mu\text{m}^{-3}\text{sec}^{-1}$ , as shown in Figure 20. The radius of the transmitter spherical boundary is  $\rho = 1 \mu\text{m}$ . The simulation runs for 1sec by steps of  $\delta t = 1$  msec. The output noisy particle concentration rate  $\hat{r}_T(t)$  of the physical model simulation is shown in Figure 21.



**Figure 20. The particle sampling noise physical model simulation input.**

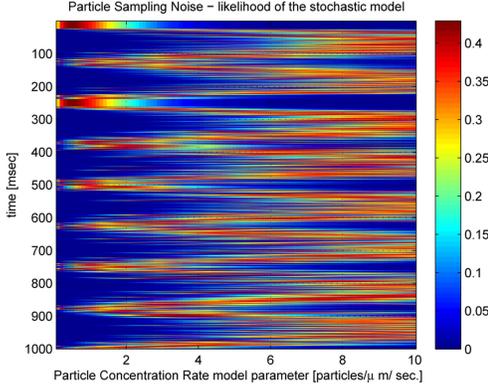


**Figure 21. The particle sampling noise physical model simulation output.**

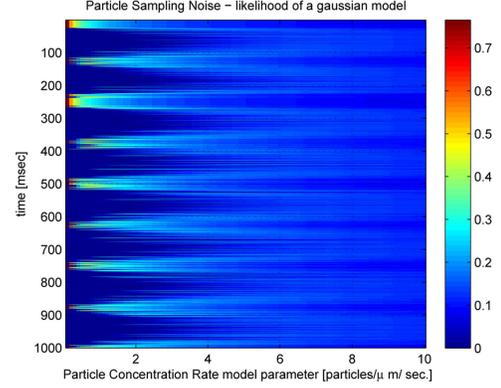
During the simulation, particles are generated inside the transmitter spherical boundary at random locations whenever the particle concentration rate  $r_T(t)$  is positive. Particle deletion is randomly performed inside the transmitter spherical boundary whenever  $r_T(t)$  is negative. Through particle generation and particle deletion we control the number of particles in the system  $P(t)$ , which is a parameter of the transmitter kinetic state  $\bar{S}_T(t)$  shown in (37). The Brownian motion of the particles is modeled according to (38) and having the diffusion coefficient  $D \sim 10^{-6} \text{ cm}^2\text{sec}^{-1}$  of calcium molecules diffusing in a biological environment (cellular cytoplasm, [58]). Samples contributing to the value of  $\hat{r}_T(t)$  are generated by applying (41) and (43) to the transmitter kinetic state  $\bar{S}_T(t)$ . The final results in terms of particle concentration rate  $\hat{r}_T(t)$  is achieved by applying (44).

The particle sampling noise has two different effects on the sinusoidal signal, namely, signal sampling and signal amplitude distortion. Signal sampling is given by the non-homogeneous sampling of the particle concentration rate  $r_T(t)$  in time, as shown in Figure 21. In non-homogeneous sampling, samples are separated by a non-constant time interval. Since in the simulations we apply a constant time step  $\delta t$ , for each time steps the contributions of samples which occur within  $\delta t$  are added. The signal amplitude distortion is given by the constant contribution that each particle gives to the concentration at the

transmitter location, (42), whenever a sample is generated by the non-homogeneous sampling. Constant contributions in non-homogeneous sampling cause sudden changes in the particle concentration rate value, which result in distortions of its amplitude.



**Figure 22. The particle sampling stochastic model likelihood.**



**Figure 23. The Gaussian model likelihood for the particle sampling noise.**

### *Sampling Noise Statistical Likelihood Test*

The statistical likelihood test is applied in order to assess the stochastic model ability to capture the behavior of the physical processes which generate the noise. For this, we compute the likelihood, that is, the probability of the noisy data coming from the physical model simulation  $\hat{r}_T(t)$  given the stochastic model of the particle sampling noise, as defined in Section 4.3.2. In order to evaluate the reliability of the particle sampling stochastic model parameters in (52) and (54), the likelihood probability is evaluated for a range of different values for the parameter  $r_T(t)$  of the Poisson processes in (49):

$$likelihood_{particleSampling} = Pr(\hat{r}_T(t)|Part.Sampl.sto.r_T(t)) \quad (60)$$

where  $r_T(t)$  ranges from 0.1 to 10 particles  $\mu\text{m}^{-3}\text{sec}^{-1}$  for every time instant  $t$ . The results are shown in Figure 22, where it is clearly visible that the highest likelihood value corresponds, for every time instant  $t$ , to the value of  $r_T(t)$  from (59), thus confirming that the best particle concentration rate, parameter of the model, is actually the particle concentration rate in input to the physical model of the particle sampling noise.

This statistical likelihood test results shown in Figure 22 are compared to the results obtained through the use of a Gaussian model in place of the particle sampling noise stochastic model. The Gaussian model, denoted by  $\mathcal{N}(r_T(t), r_T(t))$  has the same expected value and the same variance as the particle sampling noise stochastic model. The likelihood formula is:

$$likelihood_{Gaussian} = Pr(\hat{r}_T(t) | \mathcal{N}(r_T(t), r_T(t))) \quad (61)$$

where  $r_T(t)$  ranges from 0.1 to 10 particles  $\mu\text{m}^{-3}\text{sec}^{-1}$  for every time instant  $t$ . The results in terms of Gaussian model likelihood are shown in Figure 23. When the Gaussian model is applied, the likelihood shows higher values than when using the particle sampling stochastic model, but only at specific time instants. On average, the likelihood values shown in Figure 23 are much lower than the values in Figure 22 and this proves that the particle sampling stochastic model performs better than the Gaussian model. This preliminary result confirms the validity of the particle sampling stochastic model presented in this paper.

#### 4.4 The Particle Counting Noise

The model of the diffusion process provided in in Section 3.4 of Chapter 3 does not take into account the discrete nature of the particles and the randomness of their motion when the concentration  $c(\bar{x}_R, t)$  inside the receptor space is measured. In the present analysis, the following assumptions are introduced:

- The receptor space contains a discrete number of particles.
- Particles may enter/leave the receptor space as a consequence of the diffusion process, even when the concentration  $c(\bar{x}_R, t)$  at the receiver location is maintained at a constant value.

As a result, the measured particle concentration  $\hat{c}_R(t)$  suffers from two effects. The first effect is given by the quantization of the concentration measure, which is due to the discrete number of particles inside the receptor space. The second effect is given by fluctuations in

the concentration measure, which are due to single events of particles entering/leaving the receptor space.

During the reception process, particles present inside the receptor space at time instant  $t$  are counted, and their number  $\hat{N}_p(t)$  is divided by the size of the receptor space, defined in Section 3.5, equal to  $(4/3)\pi\rho^3$ , expressed as

$$\hat{c}_R(t) = \frac{\hat{N}_p(t)}{(4/3)\pi\rho^3}, \quad \hat{N}_p(t) \in \mathbb{N}, \quad (62)$$

where  $\mathbb{N}$  denotes the set of the positive integers.

#### 4.4.1 The Physical Model

The physical model of the particle counting noise is represented through the block scheme shown in Figure 24.

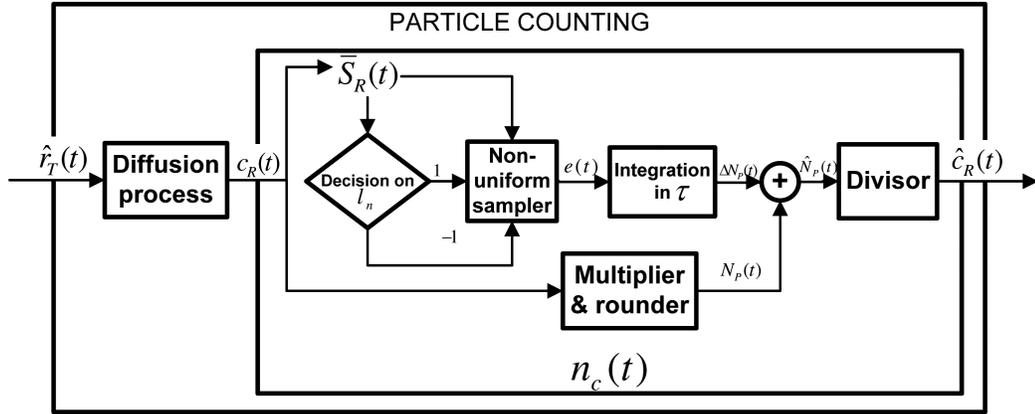
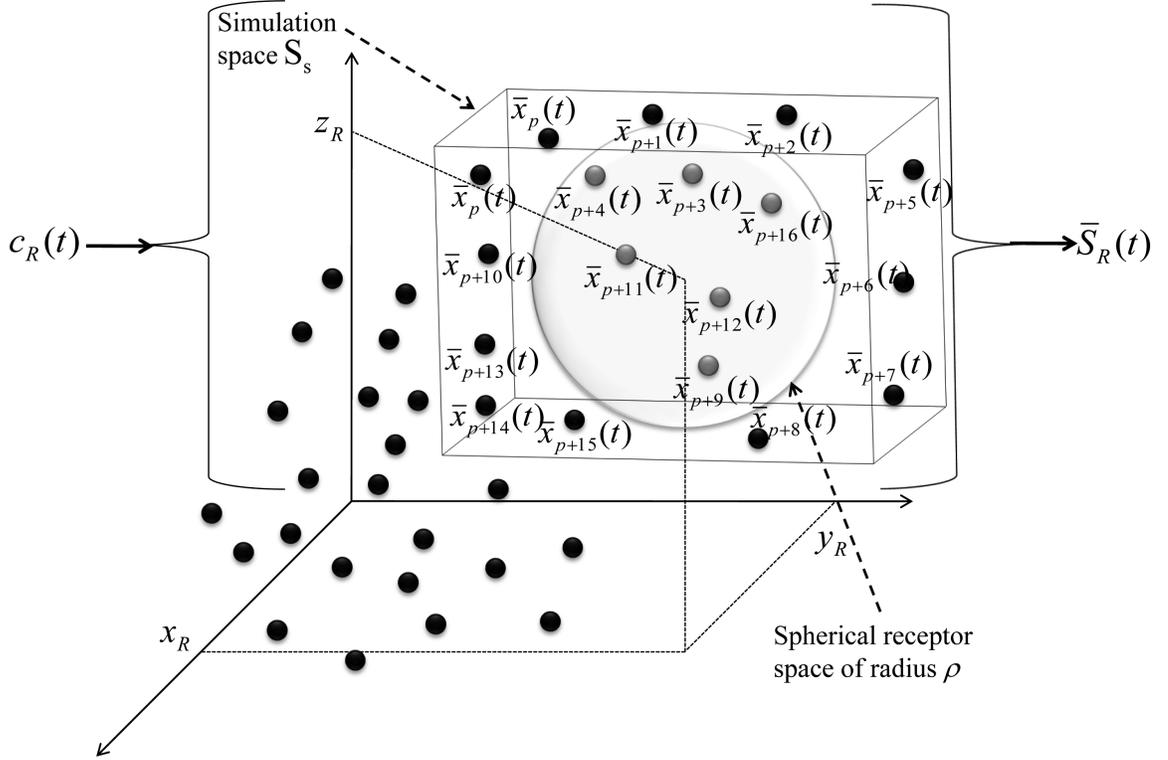


Figure 24. Block scheme of the physical model for the particle counting noise.

The particle-concentration rate  $\hat{r}_T(t)$  is the input of the diffusion-process block, whose output is the true particle concentration  $c_R(t)$ . The physical model of the particle counting noise  $n_c(t)$  takes as input the true particle concentration  $c_R(t)$  that the diffusion process would produce in output in the absence of noise. The particle counting noise  $n_c(t)$  is composed of two branches. The upper branch has a decision block and a non-uniform sampler, which have as input the receiver kinetic state  $\bar{S}_R(t)$ , while the lower branch has a multiplier and rounder block, and it takes as input the true particle concentration  $c_R(t)$ . The two

branches are then added and the result is the input of a divisor. The output of the particle counting noise  $n_c(t)$  is the particle concentration  $\hat{c}_R(t)$  affected by noise.

The **receiver kinetic state**  $\bar{S}_R(t)$ , as shown in Figure 25, is a set composed by the location  $\bar{x}_p(t)$  of each particle  $p$  at time  $t$  present in the surrounding of the receptor space defined in Section 3.5, expressed as follows:



**Figure 25.** Graphical sketch of the receiver kinetic state  $\bar{S}_R(t)$  at time  $t$ .  $\bar{S}_R(t)$  depends on the particle concentration  $c_R(t)$  in input through the expressions in (63) and (64).

$$\bar{S}_R(t) = \{ \bar{x}_p(t) \mid p = 1, \dots, P(t) \}, \quad (63)$$

where  $P(t)$  is the number of particles in the system and varies as a function of the time  $t$ . In order to realistically simulate the receiver kinetic state  $\bar{S}_R(t)$ , we consider the Brownian motion contribution at every time instant  $t$ . The expression of the particle location  $\bar{x}_p(t)$  is written as follows:

$$\bar{x}_p(t) = b_x(t) \hat{i} + b_y(t) \hat{j} + b_z(t) \hat{k}, \quad (64)$$

where the Brownian motion velocity components, namely,  $b_x(t)$ ,  $b_y(t)$  and  $b_z(t)$ , are random variables with normal distribution, zero mean value and variance equal to  $2D\delta t$ , according to the expression of the Wiener process [66], expressed as

$$b_x(t), b_y(t), b_z(t) \sim \mathcal{N}(0, 2D\delta t) , \quad (65)$$

along the versors of the cartesian axes, namely,  $\hat{i}$ ,  $\hat{j}$  and  $\hat{k}$ .  $D$  is the diffusion coefficient and  $\delta t$  is the simulation time step and it depends on how the receiver kinetic state is sampled during the physical model simulation. The smaller is the time step  $\delta t$ , the closer is the simulation to the real physical phenomenon of particle diffusion. The particle number  $P(t)$  is proportional to the particle concentration  $c_R(t)$  multiplied by the size  $size(\mathbf{S}_s)$  of the simulation space  $\mathbf{S}_s$ , shown in Figure 25, which includes the receptor space, expressed as

$$P(t) = c_R(t) size(\mathbf{S}_s) , \quad (66)$$

The **decision** block assigns the value of  $l_n$  according to the receiver kinetic state  $\bar{S}_R(t)$ .  $l_n$  can assume either value 1 or  $-1$  depending whether the kinetic state  $\bar{S}_R(t)$  has an event concerning a particle that is entering or leaving the receptor space, respectively, expressed as

$$l_n = \begin{cases} 1 & \text{if } \bar{S}_R(t) \subset \{\bar{x}_p(t)|p \text{ enters the receptor space}\} \\ -1 & \text{if } \bar{S}_R(t) \subset \{\bar{x}_p(t)|p \text{ leaves the receptor space}\} \end{cases} , \quad (67)$$

The **non-uniform sampler** block samples at time instants  $t_n$ , which are functions of the receiver kinetic state  $\bar{S}_R(t)$ . If, at time instant  $t_n$ , there is an event in the kinetic state  $\bar{S}_R(t_n)$  concerning a particle entering/leaving the receptor space, the non-uniform sampler block produces a Dirac impulse at  $t_n$ , with amplitude equal to the current value of  $l_n$ , in the output  $e(t)$  from the decision block, expressed as

$$e(t) = l_n \delta(t - t_n) \text{if } \bar{S}_R(t_n) \subset \{\bar{x}_p(t_n)|p \text{ ent./leav. rec. space}\} . \quad (68)$$

The **integration** block integrates the output from the nonuniform sampler for a time interval equal to  $\tau$  in the past up to time  $t$ , namely,  $[t - \tau, t]$ , expressed as

$$\Delta N_p(t) = \int_{t-\tau}^t e(t') dt' , \quad (69)$$

where  $\tau$  corresponds to the time interval in which we expect a quasi constant particle concentration and its effect on the particle counting noise is further discussed in Section 4.4.2. The result of the integration block is the perturbation  $\Delta N_p(t)$  at time  $t$  in the number of particles inside the receptor space.

The **multiplier and rounder** block rounds the particle concentration  $c_R(t)$  multiplied by the size of the receptor space  $(4/3)\pi\rho^3$ . The output of this block corresponds to the expected number of particles  $N_p(t)$  contained in the receptor space at time instant  $t$ , expressed as follows:

$$N_p(t) = \text{round} \left[ c_R(t) \left( \frac{4}{3}\pi\rho^3 \right) \right]. \quad (70)$$

The **divisor** block divides the sum of the output coming from the two branches, namely,  $\Delta N_p(t)$  and  $N_p(t)$ , by the size of the receptor space  $(4/3)\pi\rho^3$ . As a consequence, the output of the divisor block corresponds to the particle concentration  $\hat{c}_R(t)$  at the receiver affected by noise, expressed as

$$\hat{c}_R(t) = \frac{N_p(t) + \Delta N_p(t)}{\frac{4}{3}\pi\rho^3} = \frac{\hat{N}_p(t)}{\frac{4}{3}\pi\rho^3}. \quad (71)$$

Since it is not possible to always have knowledge of the kinetic state of the system  $\bar{S}_R(t)$  due to the huge amount of information and to the randomness in the particle motion, we cannot analytically compute the value of  $\hat{c}_R(t)$  as function of  $c_R(t)$  from the physical model of the particle counting noise. Using the physical model provided here, we can only simulate numerically the behavior of the particle counting noise  $n_c(t)$ .

#### 4.4.2 The Stochastic Model

The particle counting noise, similarly to the particle sampling noise, can also have another formulation, through statistical parameters, which is suitable when theoretical studies require an analytical expression of the noise. Statistical parameters for the particle counting noise, such as the RMS value, are provided in [62] without the definition of a complete stochastic model in terms of random processes. The derivation of these statistical parameters in [62] stems from a formulation of the particle counting noise in terms of macroscopic

thermodynamic fluctuations in the system, without accounting for a particle-by-particle analysis. In this paper, we detail the knowledge of the particle counting noise by providing a stochastic model of the noise source. This model is obtained by stemming from the physical model outlined in Section 4.4.1, where the system is modeled in a particle-by-particle fashion. As will be proved in the following, the statistical parameters computed through the stochastic model provided here are in agreement with those from [62].

The particle counting noise  $n_c(t)$  is generated by a random process  $\tilde{n}_c(t)$ , whose contribution corresponds to the difference between the measured particle concentration  $\hat{c}_R(t)$  and the expected particle concentration  $\langle \hat{c}_R(t) \rangle$ , where  $\langle \cdot \rangle$  denotes the ensemble average operator, expressed as

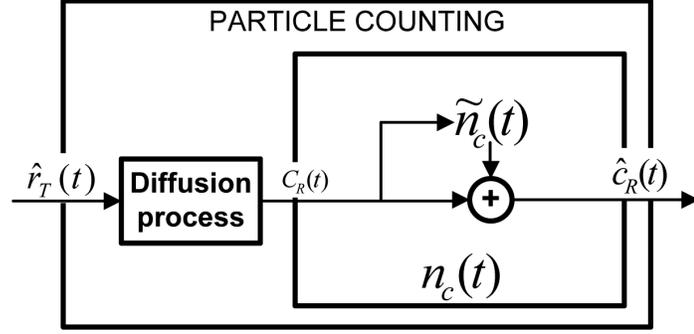
$$\tilde{n}_c(t) = \hat{c}_R(t) - \langle \hat{c}_R(t) \rangle . \quad (72)$$

The expected particle concentration  $\langle \hat{c}_R(t) \rangle$  corresponds to the true particle concentration  $c_R(t)$  that we would measure at the receiver in the absence of the particle counting noise, expressed as

$$\langle \hat{c}_R(t) \rangle = c_R(t) . \quad (73)$$

In other words,  $\tilde{n}_c(t)$  is an unwanted perturbation on the particle concentration measured at the receiver location around its expected value  $c_R(t)$  due to the particle counting noise. In Figure 26 we show the main block scheme of the particle counting noise. The random process  $\tilde{n}_c(t)$ , as it is proved in the following, depends on the value of the particle concentration at the receiver  $c_R(t)$ , output from the diffusion process, which receives the transmitted particle concentration rate  $r_T(t)$  as input. The sum of the random process  $\tilde{n}_c(t)$  and the true particle concentration at the receiver  $c_R(t)$  is the particle concentration affected by the particle counting noise, namely,  $\hat{c}_R(t)$ . In order to properly model the random process  $n_c(t)$  we consider the following assumptions:

- The actual number of particles  $\hat{N}_p(t)$  inside the receptor space at time  $t$  is a random



**Figure 26. Block scheme of the particle counting noise stochastic model.**

process whose average value is the true particle concentration at the receiver multiplied by the size of the receptor space, expressed as

$$\langle \hat{N}_p(t) \rangle = c_R(t) \frac{4}{3} \pi \rho^3 . \quad (74)$$

- It is unlikely to have two particles occupying the same location in space at the same time instant  $t$ . In other words, the probability of having a distance equal to zero between two particles at the time instant  $t$  is zero, expressed as

$$Pr \left[ \|\bar{x}_p(t) - \bar{x}_q(t)\| = 0 \right] = 0 \quad p \neq q, \quad p, q \in [1, \dots, P(t)] , \quad (75)$$

where  $P(t)$  is given by (66),  $\|\cdot\|$  is the Euclidian distance operator and  $p$  and  $q$  are two particles present in the simulation space  $\mathbf{S}_s$  defined in Section 4.4.1. This assumption is justified by the independence of the Brownian components in the movement of different particles in the space. This assumption directly translates into the property of orderliness for the counting process of the number of particles  $\hat{n}_p(t, \bar{x}(t))$  at a location  $\bar{x}(t)$  in the space, expressed as follows:

$$\lim_{\bar{\Delta} \rightarrow 0} Pr \left[ |\hat{n}_p(t, \bar{x}(t) + \bar{\Delta}) - \hat{n}_p(t, \bar{x}(t))| > 1 \right] \rightarrow 0 , \quad (76)$$

where  $\bar{\Delta}$  is a movement in the three directions of the space from  $\bar{x}(t)$  to  $\bar{x}(t) + \bar{\Delta}$ .

- An event concerning a particle which occupies a location in space  $\bar{x}(t)$  is independent of any event of the same kind occurring at another space location  $\bar{x}(t) + \bar{\Delta}$ . This assumption is justified by the property of the Wiener process underlying the particle

Brownian motion of having independent realizations. In other words, the distribution of the distance between the location of a particle in  $\bar{x}(t)$  and another particle in  $\bar{x}(t) + \bar{\Delta}_1$  is independent from the distribution of the distance between the same particle at  $\bar{x}(t)$  and another particle present at location  $\bar{x}(t) + \bar{\Delta}_2$ , where  $\bar{\Delta}_1 \neq \bar{\Delta}_2$ . The two distributions have the same expression from (65), as

$$Pr\left(\|\bar{\Delta}_1\| = x\right) = \frac{1}{\sqrt{2\pi D\Delta t}} e^{-\frac{x^2}{2D\Delta t}} = Pr\left(\|\bar{\Delta}_2\| = x\right) . \quad (77)$$

This implies that the location of a particle is independent from the location of any other particle. As a consequence, the events concerning the location of particles in the space have the property of *memorylessness*.

- The occurrence rate of particle location in the space is proportional to the particle concentration at the receiver location  $c(x_R, y_R, z_R, t)$ , equal to the expected true particle concentration  $c_R(t)$ .

Under these assumptions, the resulting actual number of particles  $\hat{N}_p(t)$  inside the receptor space is a volume non-homogeneous Poisson counting process, whose rate of occurrence corresponds to the expected particle concentration  $c_R(t)$ , expressed as

$$\hat{N}_p(t) \sim \text{Poiss}(c_R(t)) . \quad (78)$$

According to the Poisson process [66] in (216), the expected number of particles  $\langle \hat{N}_p(t) \rangle$  contained in the receptor space can be computed by multiplying the volume Poisson process rate, which is the concentration  $c_R(t)$ , by the size of the receptor space  $(4/3)\pi\rho^3$  and it is in agreement with the assumption made in (213). The variance in the number of particles contained in the receptor space has the same value as  $\langle \hat{N}_p(t) \rangle$  [66], expressed as

$$\langle (\hat{N}_p(t) - \langle \hat{N}_p(t) \rangle)^2 \rangle = c_R(t) \frac{4}{3} \pi \rho^3 . \quad (79)$$

The actual measured particle concentration  $\hat{c}_R(t)$  corresponds to the actual number of particles  $\hat{N}_p(t)$  divided by the size of the receptor space, expressed as follows:

$$\hat{c}_R(t) = \frac{\hat{N}_p(t)}{(4/3)\pi\rho^3} , \quad (80)$$

Therefore, the average  $\langle \hat{c}_R(t) \rangle$  of the actual measured particle concentration is equal to the expected particle concentration  $c_R(t)$ , as

$$\langle \hat{c}_R(t) \rangle = c_R(t) . \quad (81)$$

The variance of the actual measured particle concentration is equal to the expected particle concentration  $c_R(t)$  divided by the size of the receptor space, expressed as

$$\langle (\hat{c}_R(t) - \langle \hat{c}_R(t) \rangle)^2 \rangle = \frac{\langle (\hat{N}_p(t) - \langle \hat{N}_p(t) \rangle)^2 \rangle}{(4/3)\pi\rho^3} = \frac{c_R(t)}{(4/3)\pi\rho^3} . \quad (82)$$

Given (72) and (73), the random process  $\tilde{n}_c(t)$  has zero average value and the RMS of the perturbation  $\tilde{n}_c(t)$  on the actual measured particle concentration  $\hat{c}_R(t)$  is

$$\text{RMS}(\tilde{n}_c(t)) = \sqrt{\langle (\hat{c}_R(t) - \langle \hat{c}_R(t) \rangle)^2 \rangle} = \sqrt{\frac{c_R(t)}{(4/3)\pi\rho^3}} , \quad (83)$$

It is possible to reduce the value of  $\text{RMS}(\tilde{n}_c(t))$  by averaging in time a number  $M$  of measures of the particle concentration  $\hat{c}_R(t)$ , with the following expression:

$$\hat{c}_R(t) = \frac{1}{M} \sum_{m=1}^M \hat{c}_R(t - t_m) . \quad (84)$$

The best results in terms of noise are obtained when the  $M$  measures are statistically independent. For this, we assume independent measures when they are taken at time instants spaced by an interval  $\tau_p$ , as defined in [61]. If we assume to have a quasi-constant expected concentration in a time interval  $\tau$  (which means that the bandwidth of the signal  $c_R(t)$  is less than  $1/\tau$  [51]), the maximum value of  $M$  is equal to the time interval  $\tau$  divided by  $\tau_p$ , expressed as

$$M = \frac{\tau}{\tau_p} , \quad (85)$$

thus, reducing the RMS of the perturbation  $\text{RMS}(\tilde{n}_c(t))$  by a factor  $\sqrt{M}$ , as

$$\text{RMS}(\tilde{n}_c(t)) = \sqrt{\frac{c_R(t)}{(4/3)\pi\rho^3 M}} . \quad (86)$$

The waiting time  $\tau_p$  corresponds to the average time required for a particle to leave the reception space.  $\tau_p$  is equal to the average distance to the spherical boundary, divided by

the velocity of a particle  $v_p$ . The average distance corresponds to the receptor space radius  $\rho$ , expressed as

$$\tau_p = \frac{\rho}{v_p}. \quad (87)$$

The velocity  $v_p$  of a particle comes from the first Fick's law of diffusion [26, 27]. For this, the particle concentration flux  $\bar{J}(\bar{x}, t)$  at time instant  $t$  and location  $\bar{x}$ , is equal to the spatial gradient (operator  $\nabla$ ) of the particle concentration  $c(\bar{x}, t)$  multiplied by the diffusion coefficient  $D$ , expressed as

$$\bar{J}(\bar{x}, t) = -D\nabla c(\bar{x}, t). \quad (88)$$

When we have homogeneous concentration  $\bar{c}$  inside the receptor space and zero concentration outside the receptor space,  $\nabla c(\bar{x}, t)$  is equal to the opposite  $-\bar{c}$  of the concentration divided by the radius  $\rho$  of the receptor space. Further, the particle concentration flux  $\bar{J}(\bar{x}, t)$  is equal, by definition, to the particle concentration  $\bar{c}$  multiplied by the particle velocity  $v_p$ . If we solve (88) for the particle velocity, we obtain

$$v_p = \frac{D}{\rho}. \quad (89)$$

The average time  $\tau_p$  is therefore equal to the radius  $\rho$  squared and divided by the diffusion coefficient  $D$ , as

$$\tau_p = \frac{\rho^2}{D}, \quad (90)$$

which is in agreement with the results from [61, 62]. The final expression for the RMS of the perturbation  $\text{RMS}(\tilde{n}_c(t))$  becomes

$$\text{RMS}(\tilde{n}_c(t)) = \sqrt{\frac{c_R(t)}{(4/3)\pi D \rho \tau}}, \quad (91)$$

where  $c_R(t)$  is the expected measured particle concentration,  $D$  is the diffusion coefficient,  $\rho$  is the radius of the receptor space and  $\tau$  is the time interval in which we expect a quasi-constant particle concentration. The validity of (91) is confirmed by the results from [62], where the authors reach the same expression for the RMS of the particle counting noise by applying a different approach, as explained above.

According to the results in Section 3.4, the relation between the input particle concentration rate  $\hat{r}_T(t)$  and the measured particle concentration  $c_R(t)$  at the receiver location is expressed in the frequency ( $f$ ) domain as

$$\tilde{c}_R(f) = \tilde{\mathbf{B}}(f)\tilde{\hat{r}}_T(f) , \quad (92)$$

where  $\tilde{\hat{r}}_T(f)$  and  $\tilde{c}_R(f)$  are the Fourier transforms [51] of the particle concentration rate  $\hat{r}_T(t)$  and the particle concentration  $c_R(t)$ , respectively.  $\tilde{\mathbf{B}}(f)$  is the Transfer Function Fourier Transform [51] (TFFT) of the propagation module. The same relation in the time ( $t$ ) domain becomes

$$c_R(t) = b(t) * \hat{r}_T(t) , \quad (93)$$

where  $*$  denotes the convolution operator [51],  $b(t)$  is the impulse response of the propagation module and  $\hat{r}_T(t)$  is the input particle concentration rate. The formula for the RMS of the perturbation  $\text{RMS}(\tilde{n}_c(t))$  on the signal  $\hat{c}_R(t)$  becomes

$$\text{RMS}(\tilde{n}_c(t)) = \sqrt{\frac{b(t) * \hat{r}_T(t)}{(4/3)\pi D \rho \tau}} , \quad (94)$$

where  $D$  is the diffusion coefficient,  $\rho$  is the radius of the spherical receptor space, and  $\tau$  is the time in which we expect a quasi-constant particle concentration.

### 4.4.3 Numerical Results

#### *Counting Noise Simulations*

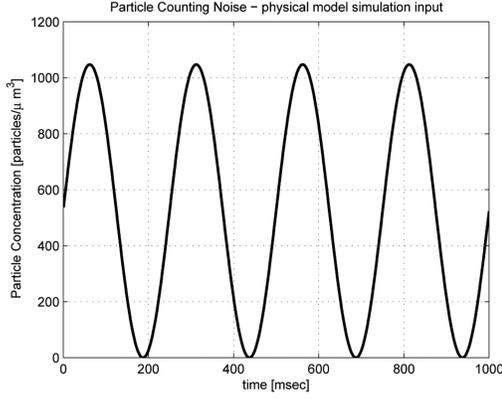
The simulations of the physical model for the particle counting noise are computed by applying to the scheme in Figure 24 a sinusoidal signal in the true particle concentration at the receiver  $c_R(t)$ :

$$c_R(t) = B \sin(2\pi f_b t) + B \quad (95)$$

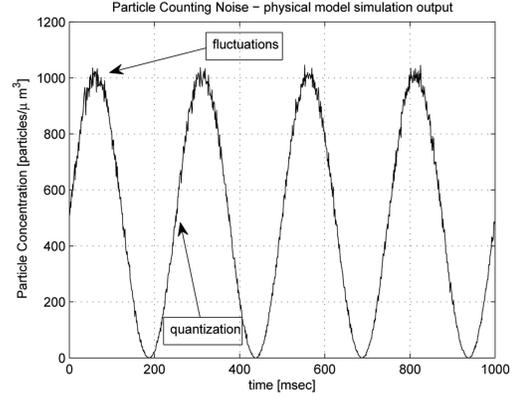
where  $f_b$  is the frequency of the sinusoid in Hz,  $2B$  is the maximum value of the expected particle concentration in particles  $\mu\text{m}^{-3}$ , and  $t$  is the simulation time index in msec.

The input of the physical model simulation is a sinusoidal particle concentration  $c_R(t)$  with frequency  $f_b$  equal to 4Hz and maximum particle concentration  $2B$  of 2000 particles

$\mu\text{m}^{-3}$ , as shown in Figure 27. The radius of the spherical receptor space is  $\rho = 1 \mu\text{m}$ . The simulation runs for 1 sec by steps of  $\delta t = 1 \text{ msec}$ . The output noisy particle concentration  $\hat{c}_R(t)$  of the physical model simulation is shown in Figure 28.



**Figure 27.** The particle counting noise physical model simulation input.

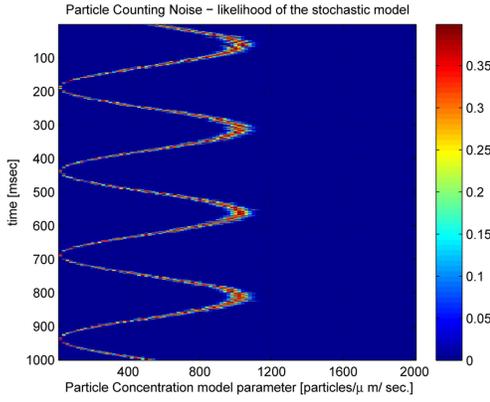


**Figure 28.** The particle counting noise physical model simulation output.

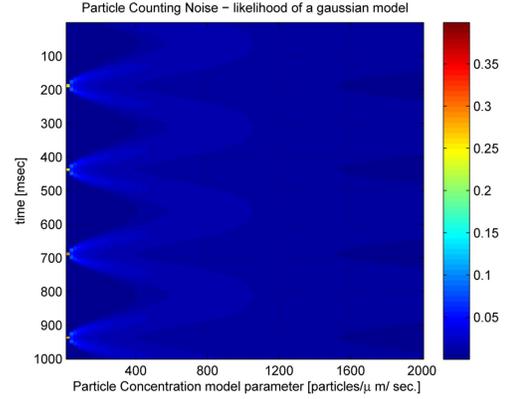
A number  $P(t)$  of particles are deployed according to (66) for each time at random locations inside the simulation space  $S_s$ , shown in Figure 25, which includes the receptor space. The receptor kinetic state is maintained according to (63) and (64), where the Brownian motion of the particles is modeled according to (65). The diffusion coefficient  $D \sim 10^6 \text{ cm}^2\text{sec}^{-1}$  corresponds to the  $D$  of calcium molecules diffusing in a biological environment (cellular cytoplasm, [58]). The upper branch of Figure 24, which generates the contribution  $\Delta N_p$  to the final result, is computed by applying (67) and (68) to the transmitter kinetic state  $\bar{S}_T(t)$ . Equation (69) is applied with a value  $\tau = 1 \text{ msec}$ , equal to a simulation step. The lower branch of Figure 24 gives the second contribution to the final result and includes the computation of  $N_p(t)$  through (70). The final results in terms of particle concentration  $\hat{c}_R(t)$  is achieved by applying (71) to the sum of the outputs from the upper branch and the lower branch.

The particle counting noise is visible through two effects, as shown in Figure 28. The first effect is given by the quantization of the concentration measure by a discrete number of particles inside the receptor space. The second effect is given by fluctuations in the

concentration measure due to single events of particles entering/leaving the receptor space. The latter is more accentuated for high values of the particle concentration. This behavior is a confirmation of the fact that the RMS value of the particle counting noise is proportional to the square root of the true particle concentration  $c_R(t)$ , as shown in (83), (86) and (91).



**Figure 29. The particle counting stochastic model likelihood.**



**Figure 30. The Gaussian model likelihood for the particle counting noise.**

### *Counting Noise Statistical Likelihood Test*

The statistical likelihood test is applied in order to assess the stochastic model model ability to capture the behavior of the physical processes which generate the noise. For this, we compute the likelihood, that is, the probability of the noisy data coming from the physical model simulation  $\hat{c}_R(t)$  given the stochastic model of the particle counting noise, as defined in Section 4.4.2. In order to evaluate the reliability of the particle counting stochastic model parameters in (81) and (82), the likelihood probability is evaluated for a range of different values for the parameter  $c_R(t)$  of the Poisson processes in (216):

$$likelihood_{ParticleCounting} = Pr(\hat{c}_R(t)|Part.Count.sto.c_R(t)) \quad (96)$$

where  $c_R(t)$  ranges from 1 to 2000 particles  $\mu\text{m}^{-3}$  for every time instant  $t$ . The results are shown in Figure 29, where it is clearly visible that the highest likelihood value corresponds, for every time instant  $t$ , to the value of  $c_R(t)$  from (95), thus confirming that the best particle concentration model parameter is actually the particle concentration in input to the physical model of the particle counting noise.

This statistical likelihood test results shown in Figure 29 are compared to the results obtained through the use of a Gaussian model in place of the particle counting noise stochastic model. The Gaussian model, denoted  $\mathcal{N}(c_R(t), c_R(t)/(4/3\pi\rho^3))$  has the same expected value and the same variance as the particle sampling noise stochastic model. The likelihood formula is:

$$likelihood_{Gaussian} = Pr\left(\hat{c}_R(t) | \mathcal{N}\left(c_R(t), \frac{c_R(t)}{(4/3\pi\rho^3)}\right)\right) \quad (97)$$

where  $c_R(t)$  ranges from 1 to 2000 particles  $\mu\text{m}^{-3}$  for every time instant  $t$  and  $\rho = 1 \mu\text{m}$ . The comparison between the Gaussian model likelihood and the particle counting stochastic model drives us to the same conclusions we had for the particle sampling noise. At specific time instants the Gaussian model likelihood shows higher values than when using the particle counting stochastic model but, on average, the likelihood values shown in Figure 30 are much lower than the values in Figure 29. This proves that the particle counting stochastic model performs better than the Gaussian model and it confirms the stochastic model ability to express the behavior of the physical processes underlying the particle counting noise.

## 4.5 The Ligand-receptor-binding Noise

In the model of the reception process provided in Section 3.5 of Chapter 3 the binding reaction and the release reaction are considered as happening only according to deterministic rates:  $k_+$  and  $k_-$ , respectively. This is justified from the viewpoint of the classical chemical kinetics [62], which interprets the time evolution of a chemical system with deterministic Reaction Rate Equations (RREs). The RRE of the chemical system defined by the assumptions from Section 3.5 is expressed as follows:

$$\frac{dn_b(t)}{dt} = k_+c_R(t)(N_R - n_b(t)) - k_-n_b(t), \quad (98)$$

where  $n_b(t)$  is the number of bound chemical receptors,  $k_+$  is the rate of particle binding,  $c_R(t)$  is the particle concentration at the receiver,  $N_R$  is the total number of chemical receptors at the receiver and  $k_-$  is the rate of particle release. In Appendix 9.1, we detail the

mathematical expressions of the particle binding and release rates. Given the the model of the reception process provided in Section 3.5, the first time derivative in the number  $n_b(t)$  of bound chemical receptors can be substituted with the output signal  $s_R(t)$  of the reception process.

The model of the reception process provided in Section 3.5 does not take into account the random fluctuations in the ligand-receptor-binding process. The following additional assumptions are here considered:

- Particles inside the receptor space are discrete and they move according to the Brownian motion.
- The binding reaction can occur only when a particle, subject to the Brownian motion, collides with an unbound receptor.
- The binding reaction occurs only if the kinetic energy of the particle colliding with an unbound receptor is higher than the activation energy  $E_a$ . The kinetic energy  $E_k^p(t)$  of a particle  $p$  at time  $t$  is expressed as follows:

$$E_k^p(t) = \frac{1}{2} |\bar{v}_p(t)|^2 m_p . \quad (99)$$

$\bar{v}_p(t)$  is the velocity of the particle  $p$  at time  $t$  and  $m_p$  is its mass, while  $|\cdot|^2$  denotes the squared absolute value operator.

- Whenever a binding reaction occurs, there is a subtraction of a particle from the reception space. Whenever a release reaction occurs, there is an addition of a particle to the reception space.

As a result, the relation between the particle concentration  $c_R(t)$  and the actual output signal of the reception process, denoted by  $\hat{s}_R^k(t)$ , is subject to random fluctuations. As shown in Figure 7, particles subject to the Brownian motion inside the receptor space contribute to the number of bound chemical receptors, denoted with  $\hat{n}_b(t)$ , only at discrete time instants

$t_n = t_1, t_2, \dots$  that correspond to collision events between the particles themselves and the unbound receptors. Each collision event contributes to  $\hat{n}_b(t)$  according to the coefficient  $k_n$ , which is a function of the kinetic energy of the collided particle. The bound receptors can become unbound according to the particle release rate  $k_-$ , thus decreasing the value of  $\hat{n}_b(t)$ . Therefore, the ligand-receptor kinetics-equation is derived by extending the Reaction Rate Equation (RRE) in (98) in order to account for the random effect of the collisions, expressed as

$$\frac{d\hat{n}_b(t)}{dt} = \left( \sum_n k_n \delta(t - t_n) \right) - k_- \hat{n}_b(t), \quad t_n = t_1, t_2, \dots, \quad (100)$$

where  $\hat{n}_b(t)$  is the number of bound chemical receptors,  $k_n$  is a coefficient related to the particle binding at time instant  $t_n$ ,  $k_-$  is the rate of particle release, and  $\delta(\cdot)$  is a Dirac delta function. The time-first derivative in the number  $\hat{n}_b(t)$  of bound chemical receptors in (100) can be substituted with the output signal of the reception process, denoted as  $\hat{s}_R^k(t)$ . The contribution of the ligand-receptor kinetics creates fluctuations in the output signal  $\hat{s}_R^k(t)$  that are not present in the previous formulation of the reception process in Section 3.3.

#### 4.5.1 The Physical Model

The physical model of the ligand-receptor-binding noise is represented through the block scheme shown in Figure 31. The particle concentration  $c_R(t)$  is the input of the overall reception process +  $w_k(t)$  block, whose output signal is  $\hat{s}_R^k(t)$ . The ligand-receptor-kinetics block is composed of the ligand-receptor kinetic state block, the integration block, and three multiplication blocks.

The **ligand-receptor kinetic state** block, as shown in Figure 31, takes as input the concentration  $\hat{c}_R(t)$  of the particles inside the receptor space and returns the signal  $a(t)$  as output. The ligand-receptor kinetic state block keeps track of the locations  $\bar{x}_p(t)$  of all the particles present inside the receptor space at time  $t$  through the set  $K_P(t)$ , expressed as

$$K_P(t) = \{ \bar{x}_p(t) \mid p = 1, \dots, P(t) \}, \quad (101)$$

where  $P(t)$  is the number of particles in the receptor space at time  $t$  and it is expressed as

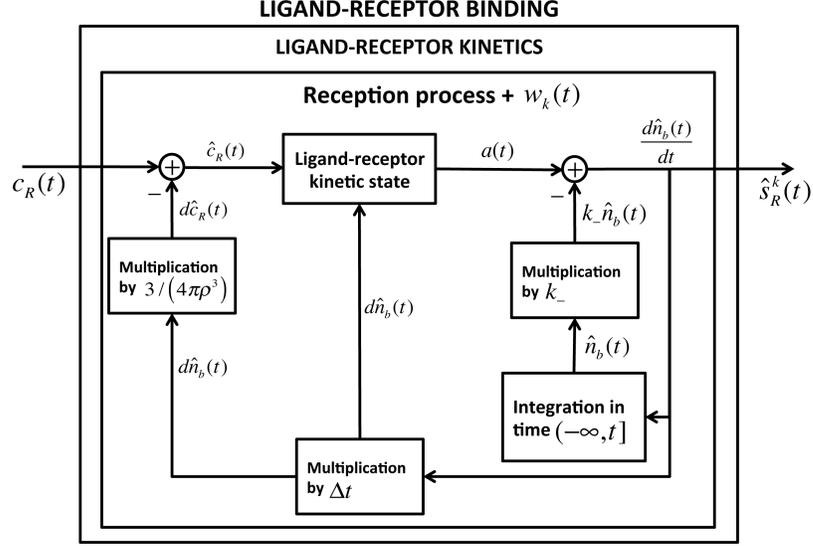


Figure 31. Block scheme of the physical model for the ligand-receptor-binding noise.

follows:

$$P(t) = \text{round}(\hat{c}_R(t) \frac{4}{3}\pi\rho^3), \quad (102)$$

where  $\hat{c}_R(t)$  is the particle concentration at the receiver and  $(4/3)\pi\rho^3$  is the size of the receptor space.  $\rho$  is the radius of the receptor space, and  $\text{round}(\cdot)$  is the operator that rounds the operand to the nearest integer. In order to realistically simulate the evolution of  $K_P(t)$ , we consider the Brownian motion contribution at every time instant  $t$ . The expression of the particle location  $\bar{x}_p(t)$  is written as follows:

$$\bar{x}_p(t) = \bar{x}_p(t - \Delta t) + b_x(\Delta t) \hat{i} + b_y(\Delta t) \hat{j} + b_z(\Delta t) \hat{k}, \quad (103)$$

where the Brownian motion components, namely,  $b_x(\Delta t)$ ,  $b_y(\Delta t)$  and  $b_z(\Delta t)$ , are random variables with normal distribution, zero mean value and variance equal to  $2D\Delta t$ , according to the expression of the Wiener process [66], expressed as

$$b_x(\Delta t), b_y(\Delta t), b_z(\Delta t) \sim \mathcal{N}(0, 2D\Delta t), \quad (104)$$

along the directions of the Cartesian axes, namely,  $\hat{i}$ ,  $\hat{j}$  and  $\hat{k}$ .  $D$  is the diffusion coefficient and  $\Delta t$  is the simulation time step and it depends on how the ligand-receptor kinetics block samples the evolution of  $K_P(t)$  during the physical model simulation. The smaller is the

time step  $\Delta t$ , the closer is the simulation to the real physical phenomenon of the particle Brownian motion. Despite in the simulation we are sampling the Brownian dynamics, the time variable of the number of bound chemical receptor  $n_b(t)$  is kept continuous. This is due to the fact that while collisions between particles and unbound receptors can occur only every  $\Delta t$  time steps, the unbinding reaction occurs continuously according to (100). The ligand-receptor kinetic state block keeps also memory of the locations  $\bar{x}_r(t)$  of the unbound chemical receptors through the set  $K_R(t)$ , expressed as

$$K_R(t) = \{\bar{x}_r(t), r = 1, \dots, n_u(t)\} , \quad (105)$$

where  $n_u(t)$  corresponds to the number of unbound chemical receptors present in the receptor space at time  $t$ . The number  $n_u(t)$  is computed taking into account the time differential  $dn_b(t - \Delta t)$  in the number of bound chemical receptors coming from the lower branch of the block scheme at time  $t - \Delta t$ . The resulting number of unbound chemical receptors  $n_u(t)$  at time  $t$  is recursively computed as follows:

$$n_u(t) = \text{round}(n_u(t - \Delta t) - dn_b(t - \Delta t)) , \quad (106)$$

where  $\text{round}(\cdot)$  is the operator that rounds the operand to the nearest integer. Since for every time instant we assume to have a uniform distribution of both particles and receptors inside the receptors space, the probability of having a collision between a particle and an unbound receptor is uniform. As a consequence, for every time instant, every unbound receptor has the same probability of having a collision with a particle. Whenever there is a collision between a particle and an unbound chemical receptor, which means that the spherical volume of a particle of radius  $r_p$  from the set  $K_P(t)$  has a non-void intersection with the volume of a receptor of radius  $r_R$  from the set  $K_R(t)$ , then the ligand-receptor kinetic block contributes to the output  $a(t)$  with a Dirac delta  $\delta(t - t_n)$  multiplied by the coefficient  $k_n$ . The coefficient  $k_n$  is equal to 1 when the kinetic energy of the colliding particle  $E_k^p(t)$  is higher than the activation energy  $E_a$ , and it is 0 otherwise. In case of  $k_n$  equal to 1, then the collision successfully results in a binding reaction (non-elastic collision). If  $k_n$  is equal to 0, then

the binding reaction does not take place and the particle resumes its Brownian motion from the location of the collision with the same kinetic energy as before the collision occurred (elastic collision). The coefficient  $k_n$  is computed through the following expression:

$$k_n = \begin{cases} 1 & \text{if } E_k^p(t_n) > E_a \\ 0 & \text{otherwise} \end{cases} . \quad (107)$$

The kinetic energy  $E_k^p(t_n)$  of the colliding particle  $p$  at time  $t_n$  is computed through (99), where the velocity  $\bar{v}_p(t_n)$  of the particle  $p$  at time  $t_n$  is computed with the following expression:

$$\bar{v}_p(t_n) = \frac{\bar{x}_p(t_n - \Delta t) - \bar{x}_p(t_n)}{\Delta t} , \quad (108)$$

where  $\bar{x}_p(t_n - \Delta t)$  and  $\bar{x}_p(t_n)$  are the particle location at time  $t_n - \Delta t$  and time  $t_n$ , respectively. The activation energy  $E_a$  is computed as a function of the rate  $k_+$  of particle binding by rearranging (407) and (408) from Appendix 9.1 as follows:

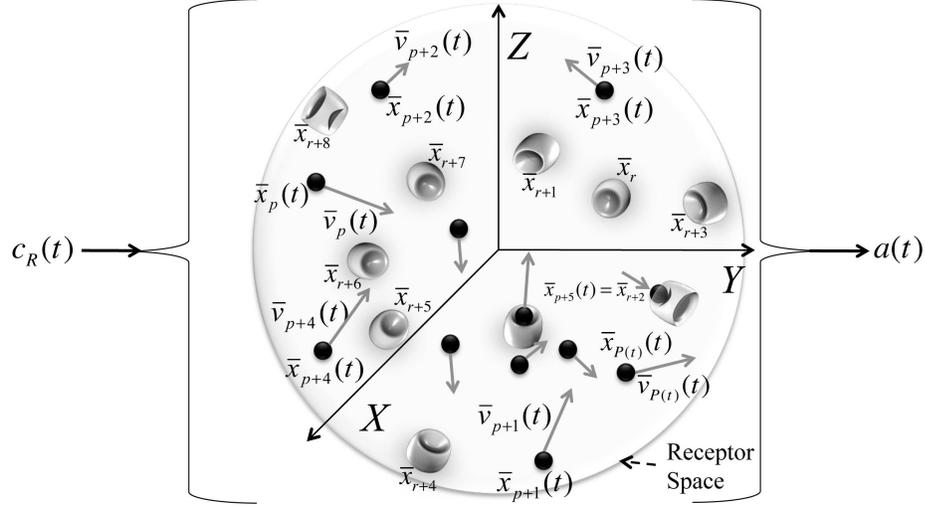
$$E_a = -k_B T \ln \left( \frac{k_+}{Z} \right) , \quad (109)$$

where  $Z$  is computed through (407). The time instant  $t_n$  corresponds to the moment when a collision between a particle and an unbound receptor occurs. As a consequence, the output  $a(t)$  of the ligand-receptor kinetic block is a sum of Dirac deltas  $\delta(t - t_n)$ , each one at a different time instant  $t_n$ , expressed as

$$a(t) = \sum_n k_n \delta(t - t_n) ; \quad t_n \in \{t_i | (K_P(t_i) \cap K_R(t_i)) \neq \emptyset\} . \quad (110)$$

Where  $K_P(t_i) \cap K_R(t_i)$  is the intersection between the set  $K_P(t_i)$  and the set  $K_R(t_i)$ , which contains elements only if there are particle locations with the same value as the locations of unbound receptors.  $\emptyset$  is the void set.

The **multiplication by  $\Delta t$**  block receives as input the first time derivative  $d\hat{n}_b(t)/dt$  at time  $t$  of the number of bound chemical receptors and returns as output its time differential  $d\hat{n}_b(t)$ . The time differential  $d\hat{n}_b(t)$  corresponds, when positive, to the number of particles subtracted from the receptor space due their binding to previously unbound receptors; when



**Figure 32.** Graphical sketch of the ligand-receptor kinetic state block. The number of particles  $P(t)$  in the receptor space depends on the particle concentration  $c_R(t)$  in input through the expressions in (102).

negative, it corresponds to the number of particles added to the receptor space due to their release from previously bound receptors.

The **multiplication by  $3/(4\pi\rho^3)$**  block receives as input the time differential  $d\hat{n}_b(t)$  at time  $t$  in the number of bound receptors, and it outputs the concentration differential  $d\hat{c}_R(t)$  of bound receptors at time  $t$ . The value of  $d\hat{c}_R(t)$  corresponds to the variation in the concentration of particles inside the receptor space given by the binding or the release of particles to/from chemical receptors. The true concentration  $\hat{c}_R(t)$  of particles present at time  $t$  inside the receptor space is given by the following expression:

$$\hat{c}_R(t) = c_R(t) - d\hat{c}_R(t) , \quad (111)$$

where  $d\hat{c}_R(t)$  is the time differential of the particle concentration  $c_R(t)$  in input to the physical model of the ligand-receptor-binding noise.

The **integration** block receives as input the first time derivative  $d\hat{n}_b(t)/dt$  in the number of bound chemical receptors and gives as output the number  $\hat{n}_b(t)$  of bound chemical receptors at time  $t$ . The output of the integration block  $\hat{n}_b(t)$  is therefore the time integral:

$$\hat{n}_b(t) = \int_{-\infty}^t \frac{\hat{n}_b(t')}{dt'} dt' . \quad (112)$$

The **multiplication by  $k_-$**  block receives as input the number  $\hat{n}_b(t)$  of bound receptors at time  $t$  and multiplies it by the rate  $k_-$  of release reaction. The output of the multiplication block is then subtracted from the output  $a(t)$  of the ligand-receptor kinetics block. The result of the subtraction is the first time derivative in the number  $\hat{n}_b(t)$  of bound receptors, expressed as

$$\frac{d\hat{n}_b(t)}{dt} = a(t) - k_- \hat{n}_b(t) , \quad (113)$$

where  $a(t)$  is computed through (110).

Since it is not possible to always have the knowledge of the ligand-receptor kinetic state due to the huge amount of information and to the randomness in the particle motion, we cannot analytically compute the value of  $\hat{S}_R^k(t)$  as a function of  $c_R(t)$  through the physical model of the ligand-receptor-binding noise. Using the physical model provided here, we can only simulate numerically the behavior of the reception noise  $w_k(t)$ .

#### 4.5.2 The Stochastic Model

The ligand-receptor-binding noise can also have another formulation, through the stochastic chemical kinetics. According to this formulation, the reception noise  $w_k(t)$  is generated by a random process, whose contribution corresponds to

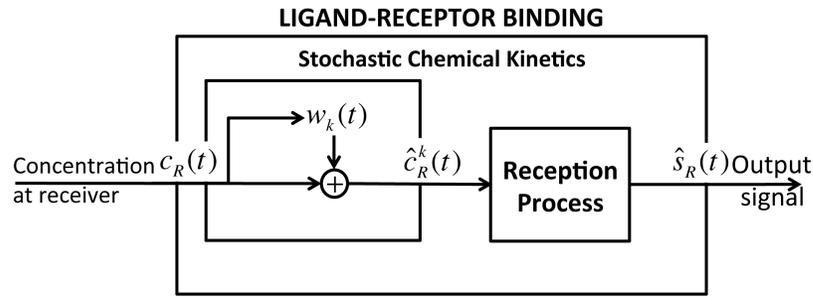
$$w_k(t) = \hat{c}_R^k(t) - c_R(t) , \quad (114)$$

where  $\hat{c}_R^k(t)$  is the actual particle concentration in input to the reception process, and  $\hat{c}_R(t)$  is the expected particle concentration in input to the reception process in the absence of the reception noise. In Figure 33, the main block scheme of the ligand-receptor-binding process is shown, where the stochastic chemical kinetics is applied to model the reception noise. As detailed in the following, the random process  $w_k(t)$  depends on the value of the particle concentration  $\hat{c}_R(t)$  itself, output of the diffusion process. The sum of the random process  $w_k(t)$  and the particle concentration  $\hat{c}_R(t)$  is the particle concentration affected by the reception noise  $\hat{c}_R^k(t)$ .

In the following, we provide the necessary assumptions underlying the stochastic models of the reception noise: the reversible second order reaction and the reversible first order reaction stochastic models. As proved in the following, the latter allows also to find a closed-form solution for the variance  $Var(\hat{c}_R^k(t))$  of the perturbed particle concentration.

#### 4.5.2.1 Stochastic Model Assumptions

In Figure 33 we show the main block scheme of the LIGAND-RECEPTOR BINDING process when the stochastic chemical kinetics is applied to model the reception noise. The random process  $w_k(t)$ , as it is proved in the following, depends on the value of the particle concentration  $c_R(t)$  itself, output from the diffusion process. The sum of the random process  $w_k(t)$  and the particle concentration  $c_R(t)$  is the particle concentration affected by the reception noise  $\hat{c}_R^k(t)$ .



**Figure 33. Block scheme of the stochastic chemical kinetics applied to the LIGAND-RECEPTOR BINDING process.**

In order to properly model the random process  $w_k(t)$  we consider the following assumptions:

- The particles and the chemical receptors inside the receptor space are considered as two different types of molecules (chemical species). For these two chemical species, we assume that the system is “well stirred”, which means that the particles and the chemical receptors have random uniformly distributed locations inside the receptor space.

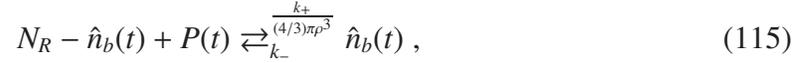
- The assumption of having a “well stirred” system allows us to describe the ligand-receptor binding process only accounting for the populations of the chemical species. We can therefore ignore the description of the system through the ligand-receptor kinetic state defined in Section 4.5.1, which was composed by the locations of the chemical receptors and the locations and velocities of the particles inside the receptor space.
- The populations of the chemical species are described through two quantities: the number of particles  $P(t)$  in the receptor space and the number of bound receptors  $\hat{n}_b(t)$  at each time instant  $t$ . The number of unbound receptors is computed by subtracting the number  $\hat{n}_b(t)$  of bound receptors from the total number  $N_R$  of chemical receptors, namely,  $N_R - \hat{n}_b(t)$ .
- As a consequence of the first assumption, the binding reaction rate is considered as a constant equal to  $k_+$ , since, without accounting for the ligand-receptor kinetic state, it is not possible to know either the kinetic energy of a particle colliding with an unbound chemical receptor, nor when the collision occurs.

Under these assumptions, the value of the population for each chemical species is never known deterministically, but only in probability. The stochastic chemical kinetics studies how the populations of the chemical species evolve in a system due to chemical reactions. This is achieved by the formulation of the chemical system through the Chemical Master Equation (CME) [67]. The CME is a stochastic differential equation that binds together the populations of the chemical species involved in the chemical reactions through a stochastic differential equation. In the following, we consider two different CMEs, namely, the **reversible second order reaction** and the **reversible first order reaction**. Although both of them can represent the ligand-receptor binding, the former CME is the most complete formulation but, due to its complexity, it does not easily provide a closed-form solution. On the contrary, the latter CME is based on further assumptions and it allows for a closed-form

solution to the problem of the stochastic modeling of the ligand-receptor-binding noise.

#### 4.5.2.2 The Reversible Second Order Reaction

The **reversible second order reaction** is able to model the changes both in the number of bound receptors  $\hat{n}_b(t)$  and in the number of particles  $P(t)$  in the receptor space occurring due to the binding and release reactions. Whenever a binding/release reaction occurs, there is a subtraction/addition of a particle from/to the reception space, and this is taken into account in the perturbation of the value of the particle concentration  $c_R(t)$  inside the receptor space. Given this assumption, we can write the reversible second order reaction CME for the ligand-receptor binding, whose schematic relation is as follows:



where  $\hat{n}_b(t)$  is the number of bound chemical receptors,  $P(t)$  is the number of particles in the receptor space and  $N_R - n_b$  is the number of unbound chemical receptors, as defined above.  $(4/3)\pi\rho^3$  is the size of the receptor space and it divides the binding reaction rate  $k_+$  since here we are dealing with the number of particles in the receptor space rather than with the particle concentration  $c_R(t)$ . The formulation of the CME for the ligand-receptor binding states that the first time derivative of the probability of having  $n_b$  bound receptors is equal to the sum of different terms: the probability  $P_{n_b-1}(t)$  of having  $n_b - 1$  bound chemical receptors and having a binding reaction, the probability  $P_{n_b+1}(t)$  of having  $n_b + 1$  bound chemical receptors and having a release reaction, the negative of the probability  $P_{n_b}(t)$  of having  $n_b$  chemical receptors and having either a release reaction or a binding reaction, expressed as follows

$$\begin{aligned} \frac{dP_{n_b}(t)}{dt} = & \frac{k_+}{(4/3)\pi\rho^3} (N_R - n_b + 1)(P(t) - n_b + 1)P_{n_b-1}(t) \\ & + k_-(n_b + 1)P_{n_b+1}(t) - [k_-n_b + k_+(N_R - n_b)(P(t) - n_b)]P_{n_b}(t) \end{aligned} , \quad (116)$$

where:  $k_+ / ((4/3)\pi\rho^3)(N_R - n_b + 1)(P(t) - n_b + 1)$  is the rate of having a binding reaction with  $N_R - n_b + 1$  available unbound receptors and  $P(t) - n_b + 1$  available particles;  $k_-(n_b + 1)$  is

the rate of having a release reaction with  $n_b + 1$  bound receptors;  $k_- n_b$  is the rate of having a release reaction with  $n_b$  bound receptors and  $k_+(N_R - n_b)(P(t) - n_b)$  is the rate of having a binding reaction with  $N_R - n_b$  available unbound receptors and  $P(t) - n_b$  available particles. Equation (116) can be schematically interpreted in terms of Markov chains [66], as shown in Figure 34. According to the theory of Markov chains, each possible value of the number  $n_b$  of bound chemical receptors can represent a finite state in a state chain. In the Markov chain of Figure 34, the probability of having a transition to a higher state number is given by the probability of being in that state and having a binding reaction, while the probability of having a transition to a lower state number is given by the probability of being in that state and having a release reaction. In order to find a closed-form solution to the problem of

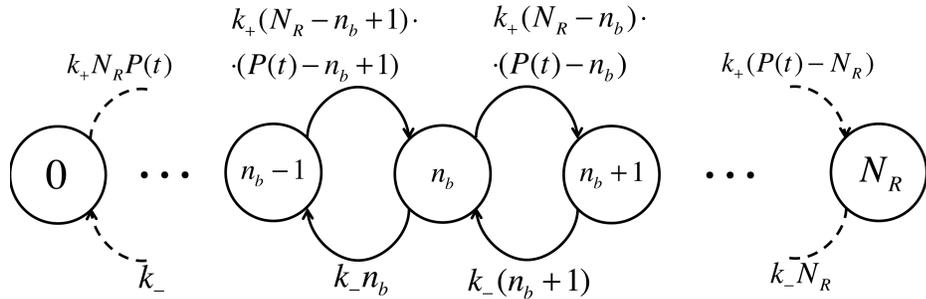


Figure 34. Graphical representation of (116) as a Markov chain.

the stochastic modeling of the ligand-receptor-binding noise it is necessary to add a further assumption to the stochastic model and to interpret the system through the reversible first order reaction CME, which is explained in the following.

#### 4.5.2.3 The Reversible First Order Reaction and Closed-Form Solution

The **reversible first order reaction** is based on a further assumption which is formulated as follows:

- The number of particles  $P(t)$  in the receptor space for any time instant  $t$  is much higher than the number  $N_T$  of chemical receptors.

As a consequence, the particle concentration  $c_R(t)$  in input to the ligand-receptor binding process is not affected by the binding or release reactions occurring between the particles and the chemical receptors. For this, even if, whenever a binding/release reaction occurs, there is a subtraction/addition of a particle from/to the reception space, the perturbation in the value of the particle concentration  $c_R(t)$  inside the receptor space is negligible. Given this assumption, we can write the CME for the ligand-receptor binding process as a reversible first order reaction, whose schematic relation is

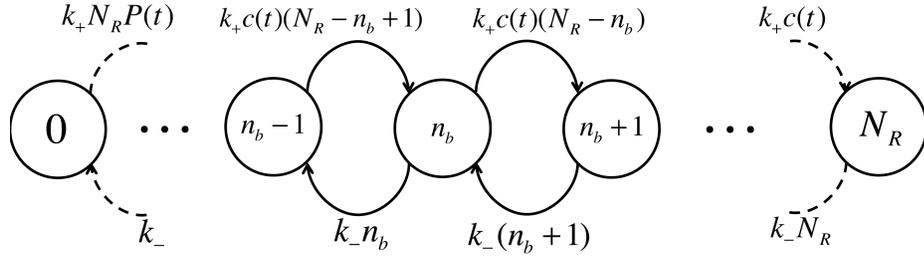


where  $\hat{n}_b(t)$  is the number of bound chemical receptors and  $N_R - \hat{n}_b(t)$  is the number of unbound chemical receptors, as defined above. The formulation of the CME for the ligand-receptor binding states that the first time derivative of the probability of having  $n_b$  bound receptors is equal to the sum of different terms: the probability  $P_{n_b-1}(t)$  of having  $n_b - 1$  bound chemical receptor and having a binding reaction, the probability  $P_{n_b+1}(t)$  of having  $n_b + 1$  bound chemical receptor and having a release reaction, the negative of the probability  $P_{n_b}(t)$  of having  $n_b$  chemical receptors and having either a release reaction or a binding reaction, expressed as follows:

$$\begin{aligned} \frac{dP_{n_b}(t)}{dt} = & c_R(t)k_+(N_R - n_b + 1)P_{n_b-1}(t) + k_-(n_b + 1)P_{n_b+1}(t) \\ & - [k_-n_b + c_R(t)k_+(N_R - n_b)]P_{n_b}(t) \end{aligned}, \quad (118)$$

where:  $c_R(t)k_+(N_R - n_b + 1)$  is the rate of having a binding reaction with  $N_R - n_b + 1$  available unbound receptors;  $k_-(n_b + 1)$  is the rate of having a release reaction with  $n_b + 1$  bound receptors;  $k_-n_b$  is the rate of having a release reaction with  $n_b$  bound receptors and  $c_R(t)k_+(N_R - n_b)$  is the rate of having a binding reaction with  $N_R - n_b$  available unbound receptors. Equation (116) can be schematically interpreted in terms of Markov Chains [66], as shown in Figure 35. We can interpret the Markov chain of Figure 35 in the same way as we did for Figure 34, where, this time, the probability of having a transition to a higher state number does not account for the number of particles  $P(t)$  in the receptor space, but

only for the value of the particle concentration  $c_R(t)$ . The solution to the problem of the



**Figure 35. Graphical representation of (118) as a Markov chain.**

stochastic modeling of the ligand-receptor-binding noise can be found through a similar procedure as in [68]. We express (118) in terms of Probability Generating Function [66]  $F(s, \tau)$ , which is defined as follows:

$$F(s, \tau) = \sum_{n_b=0}^{N_R} P_{n_b}(\tau) s^{n_b} , \quad (119)$$

where  $s$  is an auxiliary variable and  $\tau$  is a time variable which is ranging from  $t$  to  $t+1/(2B_c)$ .  $B_c$  is the bandwidth of the particle concentration  $c_R(t)$  in input to the stochastic model. According to the Nyquist theorem [65], we can sample the particle concentration  $c_R(t)$  with a rate equal to  $2B_c$  without loss of information. During a sampling time interval from  $t$  to  $t + 1/(2B_c)$  spanned by  $\tau$ , we can consider only one sample of the concentration signal  $c_R(t)$ . As a consequence, we can solve (118) treating  $c_R(t)$  as a constant parameter. The CME in (118) can be expressed in terms of Probability Generating Function as follows:

$$\begin{aligned} \frac{\partial F(s, \tau)}{\partial \tau} = & [k_- + (c_R(t)k_+ - k_-)s - k_-s^2] \frac{\partial F(s, \tau)}{\partial s} \\ & + N_R c_R(t) k_+ (s - 1) F(s, \tau) . \end{aligned} \quad (120)$$

We impose to the Probability Generating Function to have at  $\tau = t$  a number of bound receptors  $\hat{n}_b(\tau = t)$  equal to the number of bound receptors  $n_0(t)$  that we would expect in the absence of noise and at chemical equilibrium, given the particle concentration  $c_R(t - 1/(2B_c))$  in the receptor space from the previous time interval  $t-1/(2B_c)$ , which is computed

by setting to zero the derivative  $dn_b(t)/dt$  of the RRE in (98) and solving for  $n_b(t)$  as follows:

$$\hat{n}_b(\tau = t) = n_0(t) = \frac{N_R c_R(t - 1/(2B_c)) k_+}{c_R(t - 1/(2B_c)) k_+ + k_-}, \quad (121)$$

which means that the probability of having  $n_0$  bound receptors at time  $\tau = 0$  is equal to 1. As a consequence, the Probability Generating Function assumes at time  $\tau = 0$  the value  $s^{n_0(t)}$ , expressed as

$$F(s, 0) = s^{n_0(t)}. \quad (122)$$

Accounting for (122), we can solve (120) with respect to the Probability Generating Function  $F(s, \tau)$  as follows:

$$F(s, \tau) = \left\{ \frac{[k_2 s^2 + (k_1 - k_2)s - k_1](\lambda e^{-K(\tau-t)}(s-1) + \lambda + s)^2}{(s-1)(s+\lambda)K(1+\lambda)} \right\}^{\frac{n_0(t)}{2}}, \quad (123)$$

where  $k_- = k_1$ ,  $c_R(t)k_+ = k_2$ ,  $\lambda = k_1/k_2$  and  $K = k_1 + k_2$ . This allows to find the average value  $\langle \hat{n}_b(t) \rangle$  and the variance  $Var(\hat{n}_b(t))$  of the number  $n_b(\tau)$  of bound receptors at time  $\tau$  according to the properties [66] of the Probability Generating Function as follows:

$$\langle \hat{n}_b(\tau) \rangle = \left. \frac{\partial F(s, \tau)}{\partial s} \right|_{s=1}, \quad (124)$$

$$Var(\hat{n}_b(\tau)) = \left. \frac{\partial^2 F(s, \tau)}{\partial s^2} \right|_{s=1} + \left. \frac{\partial F(s, \tau)}{\partial s} \right|_{s=1} - \left( \left. \frac{\partial F(s, \tau)}{\partial s} \right|_{s=1} \right)^2. \quad (125)$$

The final expressions for the average value  $\langle n_b(\tau) \rangle$  and the variance  $Var(n_b(\tau))$  become

$$\langle \hat{n}_b(\tau) \rangle = \frac{n_0(t)}{K} (k_1 e^{-K(\tau-t)} + k_2), \quad (126)$$

$$Var(\hat{n}_b(\tau)) = \frac{n_0(t)(\lambda e^{-K(\tau-t)} + 1)}{1 + \lambda} \left( 1 - \frac{\lambda e^{-K(\tau-t)} + 1}{1 + \lambda} \right) \quad (127)$$

$\tau \in [t, t + 1/(2B_c)]$  .

The perturbed particle concentration  $\hat{c}_R^k(t)$  is computed from the value of the number of bound receptors through the steady state solution of the RRE in (98) as follows:

$$\hat{c}_R^k(t) = \frac{k_- \hat{n}_b(\tau)}{k_+(N_R - \hat{n}_b(\tau))}, \quad (128)$$

which is computed from (121) by substituting  $n_0(t)$  with  $\hat{n}_b(t)$  and  $c_R(t)$  with  $\hat{c}_R^k(t)$  and by solving for  $\hat{c}_R^k(t)$ . By substituting the value of  $n_0(t)$  from (121) into (126) and by applying the approximation  $\exp(-K(\tau - t)) \approx 1$ , the average value of the perturbed particle concentration  $\langle \hat{c}_R^k(t) \rangle$  computed by averaging (128) is equal to the particle concentration  $c_R(t)$  that we would expect in input to the reception process without the contribution of the reception noise, expressed as

$$\langle \hat{c}_R^k(t) \rangle = c_R(t) \quad \tau \in [t, t + 1/(2B_c)] \quad . \quad (129)$$

The variance  $Var(\hat{c}_R^k(t))$  of the perturbed particle concentration can be approximated through the formula for the variance of a function of a random variable of known variance and average [66] as follows:

$$Var(\hat{c}_R^k(t)) \approx \left[ \frac{N_R k_-}{k_+(N_R - \langle \hat{n}_b(\tau) \rangle)^2} \right]^2 Var(\hat{n}_b(\tau)) \quad , \quad (130)$$

valid for  $\tau \in [t, t + 1/(2B_c)]$ , and where  $\langle \hat{n}_b(\tau) \rangle$  and  $Var(\hat{n}_b(\tau))$  are computed through (126) and (127), respectively,  $N_R$  is the total number of chemical receptors,  $k_+$  is the rate of the binding reaction and  $k_-$  is the rate of the release reaction.

### 4.5.3 Numerical Results

In this section, we present a numerical analysis of the ligand-receptor-binding noise models. Sets of noise data realizations are generated through numerical simulation of both the **physical model** and the **stochastic model** of the ligand-receptor-binding noise. The sets of noise data realized using the physical model are then used to assess the performance of the analytical formulations of the reception noise in terms of stochastic model.

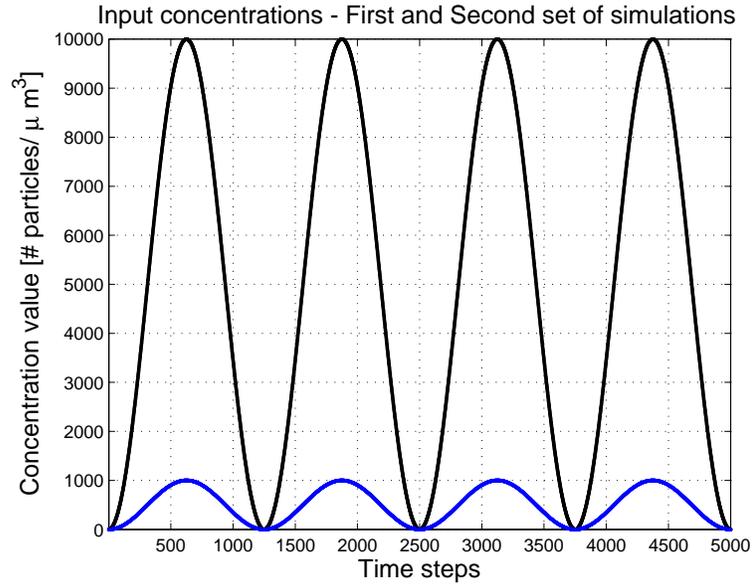
#### *Ligand-receptor Binding Simulations*

The simulations of the physical model are computed by applying a sinusoidal signal in the particle concentration  $c_R(t)$  to the scheme in Figure 31:

$$c_R(t) = A \sin(2\pi f_a t) + A \quad (131)$$

where  $f_a$  is the frequency of the sinusoid in Hz,  $2A$  is the value of the maximum particle concentration, expressed in particles  $\mu\text{m}^{-3}$ , and  $t$  is the simulation time index in msec.

We used a simple modulation waveform, a sinusoid, to produce easy to read graphical results. Moreover, the sinusoidal waveform spans the concentration values from 0 to the maximum value  $2A$  and it allows for the computation of the noise contribution for all the values in this range. Since we do not account for a time correlation model in the noise statistical parameters of the stochastic chemical kinetics, in this paper we are not interested in the analysis of different waveforms in input to the receiver and on their distortion due to the reception noise.



**Figure 36. The input of the physical model and stochastic model of ligand-receptor-binding noise simulations in terms of particle concentration.**

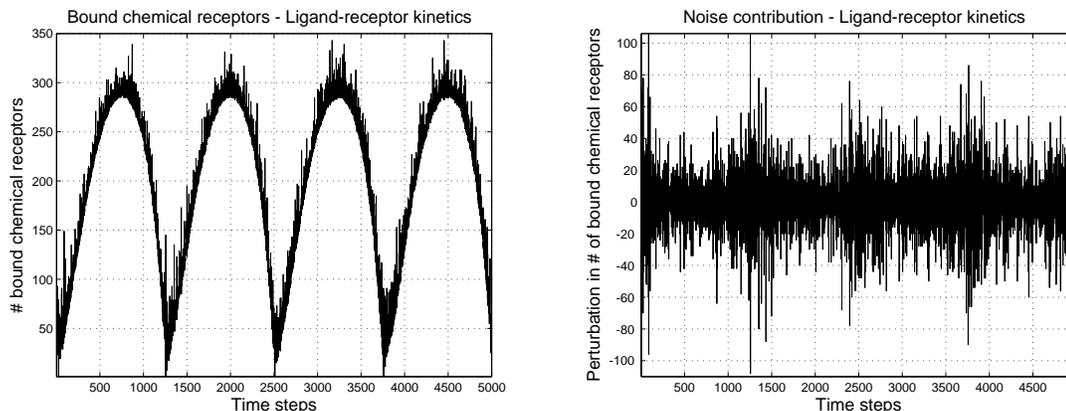
The input of the physical model simulation is the particle concentration  $c_R(t)$  in (131), sum of a sinusoid and a constant value  $A$ , since  $c_R(t)$  cannot have negative values. The sinusoid has frequency  $f_a$  equal to  $4/5$  Hz. We carried out two sets of simulations: in the first set, the amplitude of the input sinusoid is equal to the value  $A$  of 5000 particles  $\mu\text{m}^{-3}$ , while for the second set, the value of  $A$  is 500 particles  $\mu\text{m}^{-3}$ , as shown in Figure 36. The values for the particle concentration  $c_R(t)$  are quantized with respect to the number

of particles ranging from 0 to  $2A$ , even if, due to the high values of the parameter  $A$ , the quantization of the sinusoidal curves is not clearly visible in Figure 36. These two different values in the simulations enable to validate the property of the reversible first order reaction model to approximate the output of the reversible second order reaction model when the number of particles inside the receptor space is much higher than the number of chemical receptors, which is a valid assumption only for the first set of simulations. All the simulations run for  $5\text{sec}$  by steps of  $\Delta t = 1\text{ msec}$ .

The simulations are carried out using the following values for the system parameters: the radius of the reception space is  $\rho = 10\ \mu\text{m}$ , the binding reaction rate is set to  $k_+ = 0.2\ [\mu\text{m}^3/\text{sec}]$  and the release reaction rate is set to  $k_- = 10\ [1/\text{sec}]$ , with reference to [49]; the number of receptors present inside the receptor space is set to  $N_R = 500$ , while the particle diffusion coefficient, used in the physical model, is set to  $D \sim 10^{-6}\ \text{cm}^2\text{sec}^{-1}$  of calcium molecules diffusing in a biological environment (cellular cytoplasm, [58]). The radii of a particle  $r_p$  and a chemical receptor  $r_R$  are set equal to 1 nm.

The ligand-receptor binding is simulated through two different models, namely, the physical model and the stochastic model of the ligand-receptor-binding noise. The former is simulated through the block scheme shown in Figure 31, while the latter is simulated through the CME of the reversible second order reaction from (116) and Figure 34 and through the CME of the reversible first order reaction from (118) and Figure 35.

In the simulation of the **physical model**, particles are generated inside the receptor space at random locations whenever the particle concentration  $c_R(t)$  increases. Particle deletion is randomly performed inside the receptor space whenever  $c_R(t)$  decreases. Through particle generation and particle deletion, we control the number of particles  $P(t)$  in the receptor space, which is a parameter of the ligand-receptor kinetic state block shown in (101). The number of particles  $P(t)$  in the receptor space depends from the particle concentration  $c_R(t)$  through the relation in (102). The Brownian motion of the particles is modeled according to (103). Samples contributing to the value of the number  $\hat{n}_b(t)$  of bound chemical



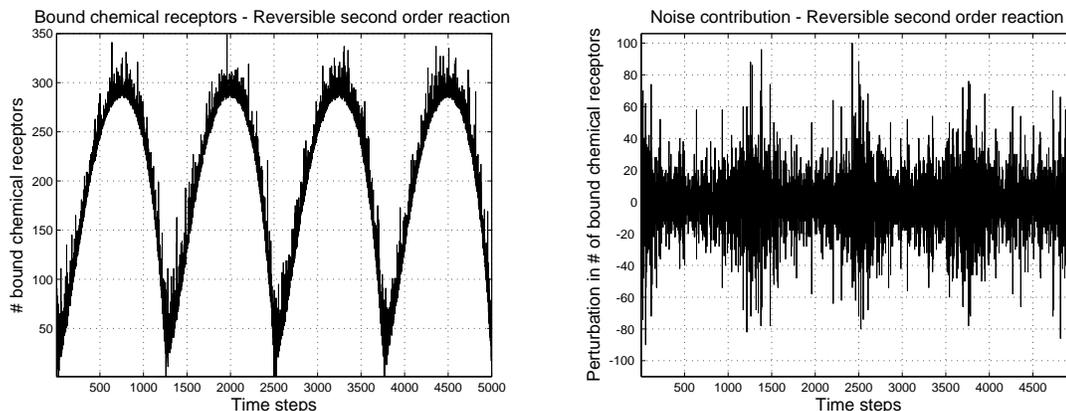
**Figure 37.** The output of the first set of simulations of the physical model in terms of number of bound chemical receptors (left) and isolated noise contribution (right).

receptors are generated by applying (107) and (110) with the knowledge of the results from (99) and (108). The final results in terms of  $\hat{n}_b(t)$  is achieved by applying (112) and (113).

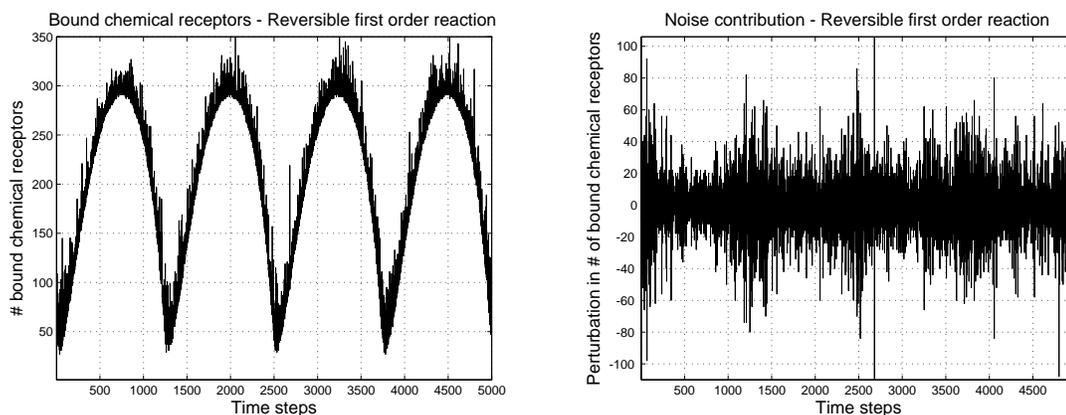
The results of the first set of simulations of the lipphysical model are shown in Figure 37 in terms of the number  $\hat{n}_b(t)$  of bound chemical receptors (left) and in terms of the perturbation of  $\hat{n}_b(t)$  around the average value (right), which corresponds to the isolated noise contribution. Figure 37 shows how the physical model affects the value of  $\hat{n}_b(t)$  more heavily when the input particle concentration  $c_R(t)$  and the resulting  $\hat{n}_b(t)$  have a lower value. This result is a consequence of the fact that when there are fewer particles inside the receptor space, the fluctuations in the number of bound chemical receptors are comparable in magnitude to the average number of bound receptors itself.

For the **stochastic chemical kinetics model**, we reproduce the behavior of the Chemical Master Equations studied in Sec 4.5.2 through simulations of the Markov chains sketched in Figure 34 and Figure 35, respectively. For the **reversible second order reaction** we use (116), while for the **reversible first order reaction** we use (118).

The results of the first set of simulations of the stochastic chemical kinetics model are shown in Figure 38 and Figure 39 in terms of the number  $\hat{n}_b(t)$  of bound chemical receptors (left) and the perturbation of  $\hat{n}_b(t)$  around the average value (right) for the reversible second



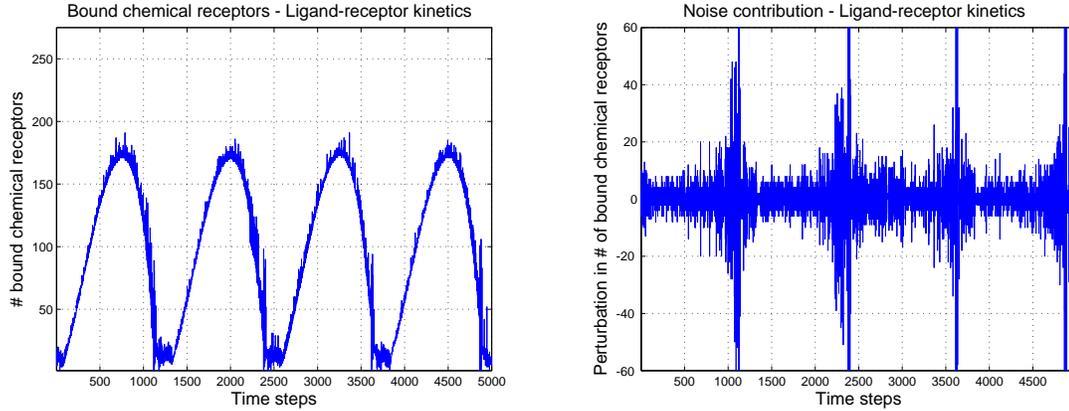
**Figure 38.** The output of the first set of simulations on the reversible second order reaction model in terms of number of bound chemical receptors (left) and isolated noise contribution (right).



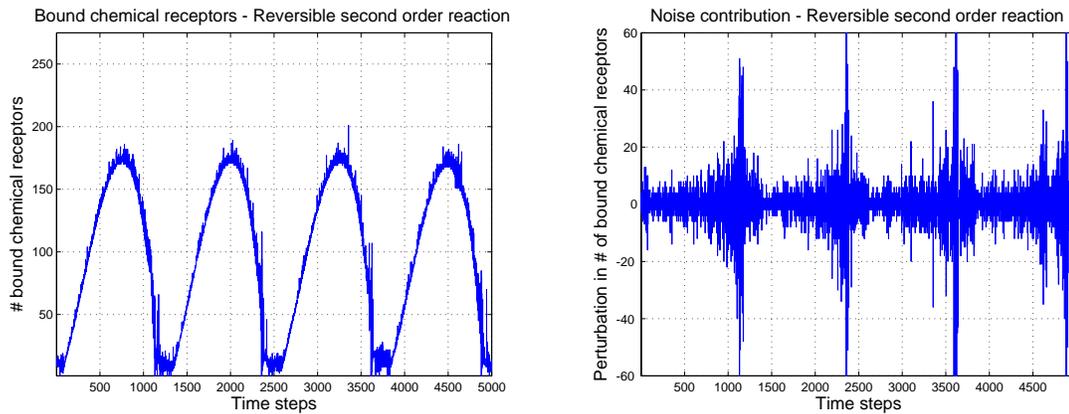
**Figure 39.** The output of the first set of simulations on the reversible first order reaction model in terms of number of bound chemical receptors (left) and isolated noise contribution (right).

order reaction and the reversible first order reaction, respectively. The results for the two types of reactions show similar values to the results of the physical model of the ligand-receptor-binding noise, shown in Figure 37.

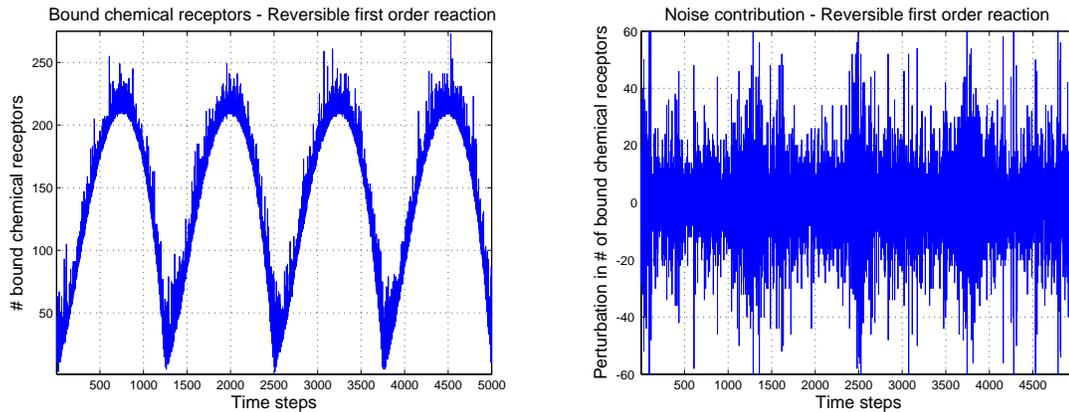
The results of the second set of simulations for the ligand receptor kinetics, the reversible second order reaction and the reversible first order reaction are shown in Figure 40, Figure 41 and Figure 42, respectively. The reversible second order reaction has values closer to the physical model if compared to the reversible first order reaction. This is a consequence of the fact that the reversible second order reaction model accounts for the



**Figure 40.** The output of the second set of simulations on the physical model of the ligand-receptor-binding noise in terms of number of bound chemical receptors (left) and isolated noise contribution (right).



**Figure 41.** The output of the second set of simulations on the reversible second order reaction model in terms of number of bound chemical receptors (left) and isolated noise contribution (right).



**Figure 42.** The output of the second set of simulations on the reversible first order reaction model in terms of number of bound chemical receptors (left) and isolated noise contribution (right).

effects of the binding or release reactions on the input particle concentration  $c_R(t)$ . The reversible first order reaction is an approximation of the real behavior of the physical model of the ligand-receptor-binding noise: the higher is the number of particles inside the receptor space than the number of chemical receptors, the closer is the reversible first order reaction to reality. Since for the second set of simulations we used a number of particles inside the receptor space closer to the number of chemical receptors, the difference of the results from the reversible first order reaction with respect to the results from the physical model are more evident. The reversible first order reaction model overestimates the number of particles present inside the receptor space, while the reversible second order reaction model realistically accounts for a depletion of the particles when these bind to the receptors.

#### *Ligand-receptor Binding Statistical Likelihood Test*

The statistical likelihood test is applied to prove that the analytical formulation of the reception noise in terms of stochastic chemical kinetics provides a good statistical model of the behavior of the physical model. For this, we compute the likelihood, that is, the probability of having a number of bound chemical receptors  $n_b$ , given a stochastic chemical kinetics models defined in Section 4.5.2, and then we compare the results with the value of the number of bound receptors from the simulation of the physical model model.

The likelihood of the stochastic chemical kinetics models is evaluated for a range of different values for the number of bound chemical receptors  $n_b$  as follows:

$$likelihood_{StoChemKin} = Pr(n_b | StoChemKin(\hat{n}_b(t))) \quad (132)$$

where  $n_b$  ranges from 1 to  $N_R$  bound chemical receptors for every time instant  $t$  and  $\hat{n}_b(t)$  is the result of the simulation of the physical model of the ligand-receptor-binding noise. The results are shown for the reversible second order reaction and for the reversible first order reaction in Figure 43 and Figure 44, respectively, for the first set of simulations. The highest likelihood value corresponds, for every time instant  $t$ , to the value of the number of bound

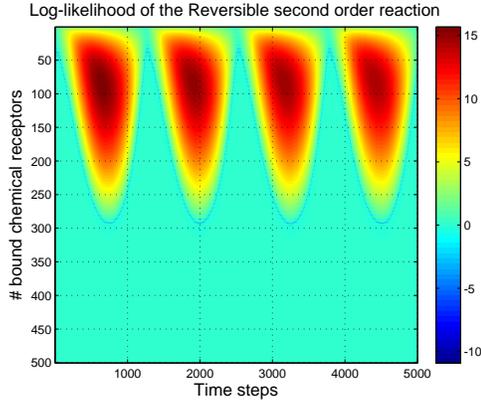
receptors in Figure 37 (left), thus visually confirming that the best particle concentration model parameter for the stochastic chemical kinetics model is actually the number  $\hat{n}_b(t)$  of bound chemical receptors in output from the physical model of the ligand-receptor-binding noise. A noticeable characteristic in Figure 43 and Figure 44 is the asymmetry of the values between a time interval where the input sinusoid increases and a time interval where it decreases. This is also evident in Figure 40, Figure 41 and Figure 42. This phenomenon is created by the difference between the values of the rates  $k_+$  and  $k_-$ , where  $k_+ < k_-$ , which results in a slower increase in the number of bound chemical receptors than a corresponding decrease. This is even more evident for a lower number of particles in the receptor space.

These statistical likelihood test results shown in Figure 43 and Figure 44 are compared to the results obtained through the use of a Gaussian model in place of the stochastic chemical kinetics model. The Gaussian model, denoted by  $\mathcal{N}(\langle n_b(\tau) \rangle, \text{Var}(n_b(\tau)))$ , has the same expected value and the same variance as the stochastic chemical kinetics model for the reversible first order reaction, from (126) and (127), respectively. The likelihood formula is:

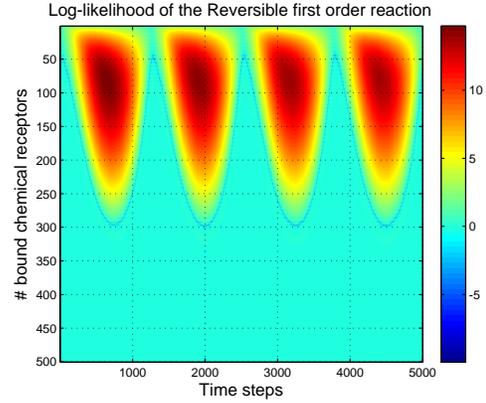
$$\text{likelihood}_{\text{Gaussian}} = \text{Pr}(n_b | \mathcal{N}(\langle n_b(\tau) \rangle, \text{Var}(n_b(\tau)))) \quad (133)$$

where  $n_b$  ranges from 1 to  $N_R$  bound chemical receptors for every time instant  $t$ .  $n_b(\tau)$  is the result of the simulation of the physical model, where  $\tau \in [t, t + 1/(2B_c)]$ . When the Gaussian model is applied, the likelihood shows lower values than when using the stochastic chemical kinetics model. On average, the likelihood values shown in Figure 45 are much lower than the values in Figure 43 and Figure 44 and this proves that the stochastic chemical kinetics model performs better than the Gaussian model. These results confirm the validity of the stochastic chemical kinetics model presented in this paper.

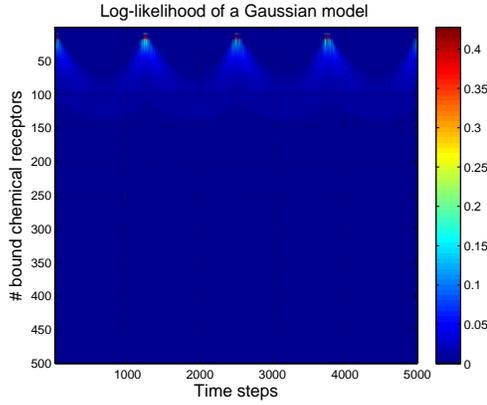
The results of the second set of simulations in terms of statistical likelihood are shown for the reversible second order reaction, the reversible first order reaction and the Gaussian model in Figure 47, Figure 48 and Figure 46, respectively. The same conclusions as for the first set of simulations can be drawn from these results, even if now the likelihood values



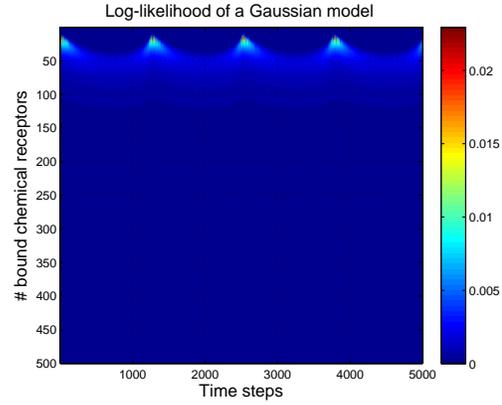
**Figure 43.** The reversible second order reaction log-likelihood for the first set of simulations.



**Figure 44.** The reversible first order reaction log-likelihood for the first set of simulations.



**Figure 45.** The log-likelihood of a Gaussian model for the first set of simulations.

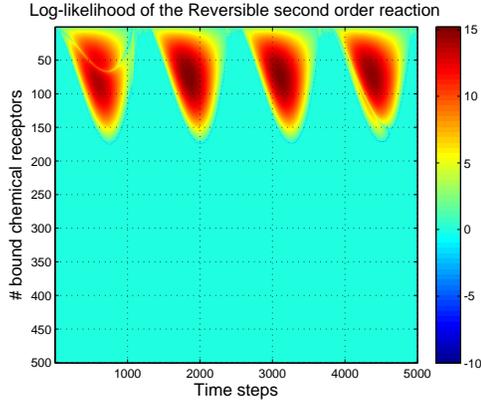


**Figure 46.** The log-likelihood of a Gaussian model for the second set of simulations.

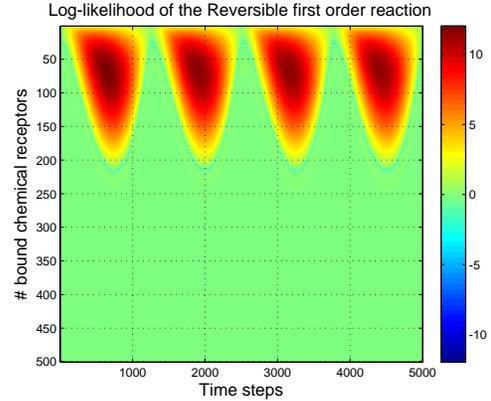
for the reversible second order reaction are noticeably different from the reversible first order reaction. This is given by the fact that for the second set of simulations the number of particles inside the receptor space is not consistently higher than the number of chemical receptors, with the above explained consequences on the reversible first order reaction.

Finally, we computed the Kullback-Leibler (K-L) distance [66]  $K-L(M)$  of each stochastic model  $M$  from the data generated through the physical model of the ligand-receptor-binding noise by applying the following formula:

$$K-L(M) = \int Pr(\hat{n}_b(t)|M) \log [Pr(\hat{n}_b(t)|M)] dt \quad (134)$$



**Figure 47.** The reversible second order reaction log-likelihood for the second set of simulations.



**Figure 48.** The reversible first order reaction log-likelihood for the second set of simulations.

**Table 1.** Kullback-Leibler Distance

	First Simulation Set	Second Simulation Set
<b>Rev. First Order</b>	-64.4435	-64.9863
<b>Rev. Second Order</b>	-65.1680	-66.4959
<b>Gaussian</b>	-89.8191	-115.7583

where  $\hat{n}_b(t)$  is the result of the simulation using the physical model of the ligand-receptor-binding noise. The results in terms of K-L distance are shown in Tab. 1, where values closer to 1 indicate a better matching between the data and the stochastic model. As expected, the values for the stochastic chemical kinetics models are closer to 1 with respect to the values for the Gaussian model. Moreover, the difference in the K-L distance values between the reversible second order and the reversible first order reaction models increases for the second set of simulations, where the assumption for the validity of the reversible first order reaction is no longer valid.

## 4.6 Conclusion

In this chapter of the Ph.D. thesis, the sampling noise, the counting noise, and the ligand-receptor-binding noise are identified as the most relevant diffusion-based-noise sources affecting the diffusion-based-MC transmitter, channel, and receiver, respectively. The analysis of the noise sources results both in physical models and stochastic models. With the

former we aim at the mathematical expression of the physical processes underlying the noise sources, while with the latter we model the noise source behaviors through the use of statistical parameters. For both the two noise sources, the results of the physical models are summarized through block schemes, which expand the end-to-end physical model from Chapter 3. The stochastic models result in noise source characterization in terms of random processes, and in the analytical expression of the Root Mean Square (RMS) perturbation of the noise on the information signal. Simulations are shown to prove that the analytical formulation of the noises in terms of stochastic models is compliant with the generate noise behaviors resulting from simulations based on the physical models.

The analysis of the noise sources provided in this paper and the results in terms of mathematical modeling will serve to expand the knowledge on diffusion-based MC systems and to support further investigation on their performance in terms of capacity and throughput. The results provided in this chapter of the thesis constitute an initial study on the noise sources affecting a basic design of a diffusion-based MC system, and further study is expected in the future to specialize these results in light of more specific diffusion-based MC system implementations.

# CHAPTER 5

## CAPACITY ANALYSIS OF DIFFUSION-BASED MOLECULAR COMMUNICATION

### 5.1 Motivation and Related Work

The theoretical analysis and the modeling of the information capacity in diffusion-based MC are of primary importance to understand the performance of a diffusion-based Molecular Communication (MC) system from an information theoretic perspective. The objective of the research detailed in this chapter is to provide closed-form mathematical expressions that are valid as upper/lower bounds of the true information capacity of an MC system based on free molecule diffusion, independent from any specific coding scheme. Shannon [69] provided the famous mathematical expression of the capacity of a channel affected by additive white Gaussian noise, which has a general validity for classical electromagnetic (EM)-based communication. As detailed later in this chapter, the diffusion-based MC has two main characteristics, namely, a channel memory and a signal-dependent noise, which limit the applicability of the aforementioned classical capacity expression. The channel memory is given by the persistent effects in the channel of previous transmissions, while a signal-dependent noise is given by a correlation between the transmitted information signal and the characteristics of the noise-generating stochastic processes present in the communication system. The impossibility of finding a closed-form analytical expression for the true capacity of such a communication system, even when affected by only one of these characteristics, is a well-known argument in information theory. As a consequence, in the attempt to provide an analytical closed-form expression that relates the performance of a diffusion-based system to physical parameters, such as the diffusion coefficient, the temperature, the transmitter-receiver distance, the bandwidth of the transmitted signal, and the average transmitted power, the results included in this chapters are based on some simplifying assumptions, mentioned in the following. These assumptions limit the validity of these

expressions, which have to be considered as upper/lower bounds to the true information capacity.

Up to date, some contributions from the literature have attempted to study the information capacity in diffusion-based MC systems, but often these are focused on specific modulation and coding schemes, or do not take into account the aforementioned channel memory and signal-dependent noise characteristics of the diffusion-based MC. The work in [35] addresses for the first time the capacity of MC systems by emphasizing the need for its mathematical analysis, but no concrete solutions are proposed. In [37, 43], the MC capacity is computed for a specific binary coding scheme and by taking into account the molecular receiver model, but without modeling the molecule diffusion propagation. An analysis of the molecular achievable rate is conducted in [70] by assuming a single instantaneous emission of molecules from the transmitter, a deterministic diffusion channel and a detailed chemical model of the receiver, but the effects of an emission of molecules over time is not considered. In [44], the capacity of a MC system in case of binary coding is properly analyzed on the basis of the effects of the channel memory, but without accounting for molecular noise sources. The same authors complement in [71] their first work by including the contribution on the capacity of a ligand-receptor binding reception of binary coded molecular signals. Moreover, in [72] and [73], they further extend their work by considering the effects on capacity of an MC system realized through genetically engineered bacteria, where multiple bacteria act collectively as a network node. To achieve these results, the authors take into account the noise sources in the biochemical processes of molecule generation and reception, in the contexts of both binary and M-ary coding schemes. Two different coding techniques are analyzed in [46] in terms of achievable rates, while the diffusion channel models are reduced to a binary or a quadruple channel. Similarly, discrete memoryless approximations are applied to the molecule diffusion channel in [48], where the MC capacity is computed for a binary coding scheme.

In the first part of this chapter, solutions from statistical mechanics and equilibrium thermodynamics are applied to derive a preliminary upper-bound expression to the diffusion-based MC capacity. This preliminary expression is derived through the simplifying assumption of having a molecular system in equilibrium, and the dynamic effects of the diffusion-based channel are not taken into account, in particular the channel memory. This derivation is based on the interpretation of a diffusion-based MC system as at the crossroad of two different disciplines, namely, information theory and statistical mechanics. While information theory [69] focuses on the quantification of the information in a communication channel, statistical mechanics [74] studies the thermodynamic behavior of systems composed of a large number of particles. While in the past years some efforts [75, 76] have been devoted to merge information theory and statistical mechanics, they were not directed towards the interpretation of a communication system. Moreover, these contributions tend to focus on the explanation of natural phenomena using information theory as a tool, and they do not provide suitable models for MC engineering. In particular, the relationship between thermodynamic entropy and information entropy is used to derive a closed-form expression of an upper bound to the true capacity in diffusion-based MC, as a function of the parameters from statistical mechanics, namely, the volume, the temperature, the number of molecules, as well as the bandwidth of the system and the thermodynamic power spent at the transmitter for releasing molecules.

In the second part of this chapter, unlike the previous contribution, a lower-bound expression of the capacity is provided by taking into account both the channel memory and the signal-dependent noise, termed molecular noise. For this, we decompose the molecule diffusion into two main processes: i) the Fick's diffusion, which captures solely the effects of the channel memory; ii) the particle location displacement, which isolates the molecular noise. The properties of these two processes allow to analyze them as a cascade of two separate communication systems. We compute the information capacity by assuming that the transmitter can modulate the emission of molecules in the space according to any

possible time continuous input message, differently from previous contributions where the transmitter is assumed to modulate (e.g., binary coding) impulses according to discrete input messages (e.g., binary digital messages). As a consequence, this lower-bound expression to the information capacity is independent from any specific coding scheme, and it is expressed as a function of the average transmitted power, which corresponds to the thermodynamic power spent at the transmitter for molecule emission.

The information-theoretic diagram of a diffusion-based MC system is shown in Figure 49, and it is composed by the classical [77] cascade of information source, transmitter, channel, receiver and destination. The **Information Source** produces *messages* to be communicated to the destination. The type of message depends on the particular application in which the diffusion-based MC system is deployed. In case of intelligent drug delivery applications [78], the message can be a time sequence of ON/OFF values that trigger/stop the release of the drug molecules. In nanomachine communication [35] the message can be any function of the time carrying data such as nanomachine states [1] or sensory measurements [6]. The **Transmitter**, the **Channel**, and the **Receiver**, which are based on the molecule emission, molecule diffusion and molecule reception, respectively, are within a **Physical System**, whose underlying laws and parameters affect how these components are physically realized. The **Destination** is the recipient of the *messages* coming from the receiver. Upon reception of a message, the destination reacts according to the meaning and to the particular application.

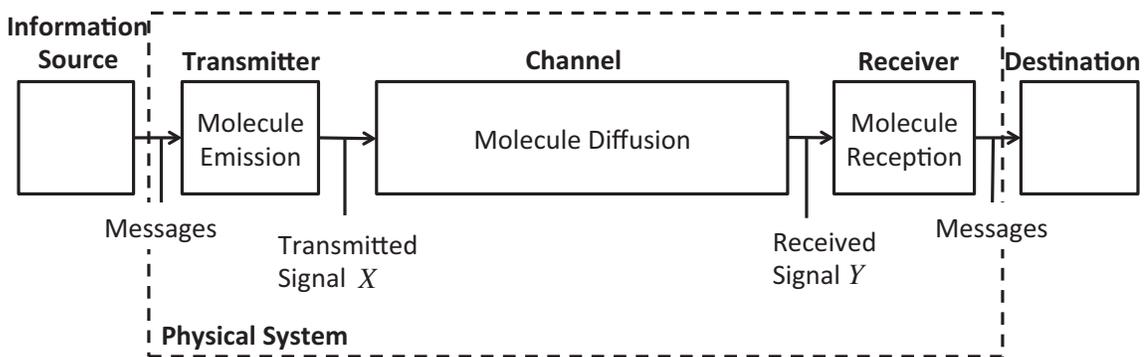


Figure 49. Information-theoretic diagram of a diffusion-based MC system.

## 5.2 Capacity Analysis through Thermodynamics

In a diffusion-based MC, the transmitter, the channel, and the receiver are function of **physical parameters** and depend on how the diffusion-based MC system is physically realized. The physical parameters are variables that characterize a physical realization of the MC diffusion-based system. Physical parameters can be considered, e.g., the temperature or the chemical composition of the environment in which diffusion-based MC is performed. For this, we define a physical reference model that allows to identify general physical parameters in a diffusion-based MC system.

The **physical reference model** embodies the basic characteristics of a diffusion-based MC system through the definition of the physics underlying the diffusion-based molecule exchange. We are interested in the analysis of the theoretical best performance limit [69] for the communication system (diffusion-based MC capacity). For this, we study an ideal physical realization of the MC system through the use of an ideal gas. The physical reference model based on the ideal gas is sketched in Figure 50 and it is defined through the following statements:

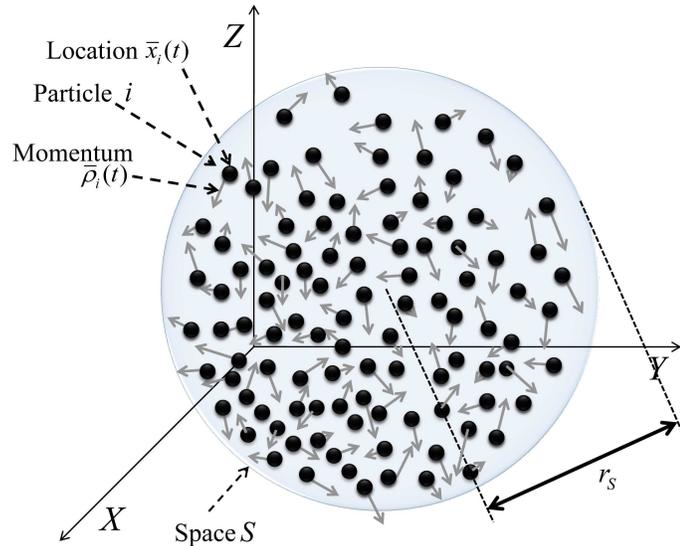


Figure 50. Sketch of the reference physical model for diffusion-based MC.

- The physical reference model is contained in the *space S*, whose shape is spherical with

radius  $r_S$ .

- All the molecules are considered as mono-atomic.
- The molecules have negligible spatial dimension when compared to the size of the space  $S$ . As a consequence, in the physical reference model they are equivalent to zero-dimensional point *particles*. From now on we will refer to particles when talking about molecules in the reference system.
- Each particle  $i$  is characterized by two quantities, namely, the location  $\bar{x}_i$  and the momentum  $\bar{\rho}_i$ . The location of a particle is a vector  $\bar{x}_i = [x_i, y_i, z_i]$  containing the values of the three space coordinates where the particle is located. The momentum  $\bar{\rho}_i$  of a particle is the product of the velocity  $\bar{v}_i$ , which is a vector containing all the velocity components  $\bar{v}_i = [v_i^x, v_i^y, v_i^z]$ , by the particle mass  $m$ . The set containing the locations and momenta of all the particles in the system define the Phase Space ( $\Phi$ ) of the system as follows:

$$\Phi = \{\bar{x}_i, \bar{\rho}_i | i = 1, \dots, N_p\}, \quad (135)$$

where  $N_p$  is the total number of particles in the system.

- Each particle is randomly-moving in the space following the Brownian motion [79] random process. According to the Brownian motion, the distribution  $f_{d_i}$  of the distance traveled by a particle  $i$  in any direction in a time interval  $t$  is equal to

$$f_{d_i}(d_i, t) = \frac{1}{\sqrt{4\pi Dt}} e^{-\frac{d_i^2}{4Dt}}, \quad (136)$$

where  $d_i$  corresponds to  $x_i, y_i, z_i$  or any other direction and  $D$  is the diffusion coefficient. Due to the Brownian motion [79], the distribution  $f_{\bar{\rho}}$  of the momenta of the particles in the system is

$$f_{\bar{\rho}}(\bar{\rho}) = \left( \frac{1}{2\pi m k_b T} \right)^{\frac{3}{2}} e^{-\frac{\|\bar{\rho}\|^2}{2mk_b T}}, \quad (137)$$

where  $\|\cdot\|$  is the Euclidian distance operator,  $k_b$  is the Boltzmann constant,  $m$  is the particle mass and  $T$  is the absolute temperature of the system.

Given the above statements, the particles of the physical reference model behave according to the theory of the ideal gases [80]. Due to the simplicity of the model, the ideal gas is the most simple diffusion-based molecular system and enables us to find a closed-form expression for an upper bound to the capacity. The ideal gas concept allows us to define the **physical parameters** for the physical reference model when it is in a state of thermodynamic equilibrium. The thermodynamic equilibrium is defined when the system has homogeneous distribution of the particle locations in the *space S* and the distribution of the particle momenta follows (137). When the system is in thermodynamic equilibrium we can define the following physical parameters: the temperature  $T$ , the pressure  $P$ , the volume  $V$  and the number of particles  $N_p$ . The physical parameters completely define the state of the system and they are bound by the Ideal Gas Law.

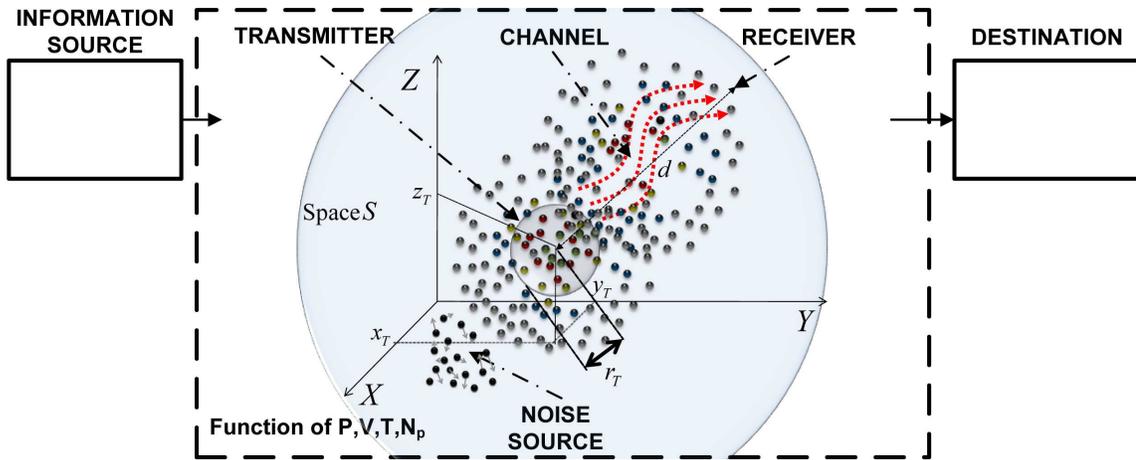


Figure 51. Schematic diagram of the diffusion-based MC system with the physical reference model.

In Figure 51 we show the elements from the diagram in Figure 49 that are function of the physical parameters. In the following, those elements are detailed in light of the definition of the physical reference model.

The **TRANSMITTER** has a spherical shape of radius  $r_T$ , where  $r_T \ll r_S$ . The location of the transmitter corresponds to the center  $[x_T, y_T, z_T]$  of the spherical space  $S$ . The transmitter can modulate two different molecular features for information transmission over the channel. Those features are the local particle number and the particle type inside the

transmitter spherical shape. The local concentration is modulated by releasing or absorbing particles into/from the space  $S$  at the transmitter location. The particle type modulation is realized by releasing particles of different types.

The **CHANNEL** propagates the signal from the transmitter to the receiver by means of particle diffusion, as detailed in Section 3.4, which stems from the expression of the Brownian motion in (136). From now on, we assume to have the same particle diffusion coefficient  $D$  for all the particles released by the transmitter.

The **RECEIVER** is point-wise (zero-dimensional). The location of the receiver is at a distance  $d$  from the transmitter, where  $r_T < d < r_S$ . The receiver has the ability to read the particle concentration and the incoming particle type at its location. The reading of the particle concentration is realized by counting all the particles present around the receiver location. The particle types are recognized on the basis of some chemical features (e.g., atomic composition) that characterize the particles.

The information capacity of a communication system is defined as the maximum rate of transmission at the information source that allows the reception of all the sent information at the destination. *The goal of the research work detailed in this section of the Ph.D. thesis is the study of the information capacity as a function of the physical parameters that control the diffusion-based MC systems in the physical reference model.*

### 5.2.1 Information-theoretic Definition of Capacity

The information **capacity** of a communication system is expressed by the general formula from Shannon [69]. The general formula defines the information capacity as the maximum difference between the entropy  $H(x)$  of the signal  $x$  in input to the channel and the equivocation  $H_Y(x)$  as follows:

$$C = \max_{f_X(x)} \{H(x) - H_Y(x)\} , \quad (138)$$

where the maximum is found with respect to the probability density function  $f_X(x)$  of all the possible values of the input signal  $x$ .

**Definition 5.2.1** The *entropy*  $H(x)$  of the input signal  $x$  is defined [69] as the opposite of the integral of the probability density function  $f_X(x)$  multiplied by its base 2 logarithm in the space of all the possible values of the input signal  $x$ , expressed as

$$H(x) = - \int f_X(x) \log_2 (f_X(x)) dx, \quad (139)$$

where  $H(x)$  is the entropy expressed in bits per transmitted sample [69] [bit/sample]. If we assume that the system has a bandwidth  $W$ , and by approximating the response of the system within this bandwidth as flat, the entropy  $H'(x)$  of the input signal expressed in [69] bits per second [bit/sec] is the entropy in [bit/sample] multiplied by the maximum rate of samples per second, which is equal to  $2W$ . The formula is [69] as follows:

$$H'(x) = 2WH(x), \quad (140)$$

**Definition 5.2.2** The *equivocation*  $H_Y(x)$  is defined as the entropy of the signal  $x$  in input given the output signal  $y$ . The equivocation  $H_Y(x)$  is computed as the opposite of the integral of the joint input-output distribution  $f_{X,Y}(x,y)$  multiplied by the base 2 logarithm of the probability density function  $f_{X|Y}(x|y)$  of the input signal  $x$  given the output signal  $y$ . The integral is computed over all the possible values of the input signal  $x$  and the output signal  $y$ :

$$H_Y(x) = - \int \int f_{X,Y}(x,y) \log_2 (f_{X|Y}(x|y)) dx dy, \quad (141)$$

In the following, we study the entropy of the input signal (Section 5.2.3), the equivocation (Section 5.2.4) and the upper-bound expression of the capacity (Section 5.2.5) of a diffusion-based MC system as functions of the physical parameters such as the volume, the temperature, the number of molecules. For this, in Section 5.2.2 we investigate the similarity between two entropies that can be defined in a MC system, namely, **information entropy and thermodynamic entropy**.

### 5.2.2 Information-theoretic Entropy from Thermodynamic Entropy

The thermodynamic entropy is defined by Gibbs [81] as a measure of the disorder in a thermodynamic system when it is in a macrostate. A thermodynamic system can be described in a twofold way, namely, through the definition of its macrostate or its microstate. A macrostate corresponds to the values of the overall thermodynamic variables of the system, such as pressure, volume, temperature and number of particles. A microstate is defined by all the values of the mechanical variables present in the system, such as particle locations and particle velocity. A system defined by a macrostate can assume many different microstates. As a consequence, a macrostate corresponds to a distribution of microstates. The entropy  $S$  associated to a thermodynamic system in a macrostate  $i$  is defined as the opposite of the Boltzmann constant [81]  $K_b$  multiplied by the sum over the space of all possible microstates  $\Psi$  of the probability  $p_i$  for the system of being in a microstate  $i$ , multiplied by the natural logarithm of the same probability, expressed as

$$S = -K_b \sum_{i \in \Psi} p_i \ln(p_i) , \quad (142)$$

The physical reference model introduced at the beginning of Section 5.2 is a thermodynamic system. The macrostate of the physical reference model can be defined through the values of its thermodynamic variables when it is in thermodynamic equilibrium. The microstate of the physical reference model is defined by the locations and the momenta of all the particles in the system. Therefore, the microstate corresponds to the values contained in the phase space  $\Phi$ , (135) defined at the beginning of Section 5.2. The phase space of the reference model can assume continuous values (locations and momenta are continuous). As a consequence, the microstates have a continuous distribution when the physical reference model is in a macrostate. The thermodynamic entropy  $S_r$  of the physical reference model is given by the formula in (142), where the distribution is expressed as a probability density function  $f_\Phi(\phi)$  and the sum is substituted with an integral over all the possible values

$\Psi$  for the phase space  $\phi$ , expressed as follows:

$$S_r = -K_b \int_{\phi \in \Psi} f_{\Phi}(\phi) \ln(f_{\Phi}(\phi)) d\phi, \quad (143)$$

where the integral is computed over  $6N_p$  dimensions:  $3N_p$  dimensions for the particle locations and  $3N_p$  dimensions for the particle momenta.  $N_p$  is the number of particles present in the system.

As stated at the beginning of Section 5.2, the physical reference model behaves according to the ideal gas theory. According to the Sackur-Tetrode equation [76], the entropy of an ideal gas in thermodynamic equilibrium has a closed-form expression as function of the thermodynamic variables. Therefore, the entropy  $S_r$  of the physical reference model can be computed using the Sackur-Tetrode equation, expressed as follows

$$S_r = N_p K_b \left[ \ln \left( \frac{V}{N_p} \left( \frac{2\pi m K_b T}{h^2} \right)^{\frac{3}{2}} \right) + \frac{5}{2} \right], \quad (144)$$

where  $N_p$  is the number of particles present in the system,  $K_b$  is the Boltzmann constant [81],  $V$  and  $T$  are the volume and the absolute temperature of the system, respectively,  $m$  is the particle mass and  $h$  is the Planck's constant [76].

The thermodynamic entropy formula in (143) can be reduced to the information entropy formula in (139) if the Boltzmann constant  $K_b$  is removed and the logarithm  $\ln$  is set to  $\log_2$ . The comparison between thermodynamic entropy and information entropy has already been tackled as a research topic [76]. The main conclusions of this research state that despite the mathematical formulation is the same, the physical meaning of the two formulas is fundamentally different. This difference is related to the probability density function to which the formula applies. In the case of the thermodynamic entropy, the probability density function  $f_{\Phi}(\phi)$  is a function of the thermodynamic variables of the system (macrostate). In the case of the information entropy, the probability density function  $f_X(x)$  is the probability of having the input signal value  $x$  out of all the possible values, without being a function of any physical variable. In a diffusion-based MC system, the similarity between the two formulas is not only formal. The microstate of the physical reference

model can be also seen as an input signal value. We are only able to control the thermodynamic variables of the physical reference model but not directly its microstate. Therefore, the probability density function of having a microstate out of all the possible microstates is a function of the thermodynamic variables. In conclusion, in the case of a diffusion-based MC system, we can interpret the thermodynamic entropy as an information entropy where the input signal values are the microstates of the physical reference model. The Boltzmann constant  $K_b$  relates only to the conventional units of the temperature [81] and it has no meaning from the Information Theory point of view. The logarithm  $\ln$  is converted into the  $\log_2$  because [69] the units of the information entropy are  $[bit/sample]$  or  $[bit/sec]$ .

As a consequence, the information entropy of the physical reference model has the same expression as the thermodynamic entropy in (144), where the first Boltzmann constant  $K_b$  is removed and the logarithm  $\ln$  is set to  $\log_2$ , expressed as

$$H_{ref} = N_p \left[ \log_2 \left( \frac{V}{N_p} \left( \frac{2\pi m K_b T}{h^2} \right)^{\frac{3}{2}} \right) + \frac{5}{2} \right]. \quad (145)$$

### 5.2.3 The Input-signal Entropy

The input signal in the physical reference model introduced at the beginning of Section 5.2 corresponds to the modulation of the **molecular features** operated by the **transmitter**. As described at the beginning of Section 5.2, the molecular features that can be modulated are the local particle number and the particle type.

The modulation of the particle features affects the value of the entropy of the physical reference model. The distribution of the particle locations in the space  $S$  is not homogeneous due to both the modulation of the local particle number and the particle type. As a consequence, the modulation shifts the physical reference model away from the state of thermodynamic equilibrium. Therefore, the entropy of the physical reference model cannot be computed through the formula in (145).

A local entropy can be defined for the transmitter when it modulates the molecular features. The local information entropy depends on the molecular features and it can be

computed from (145). The local entropy  $H_T$  at the transmitter is a function of the set  $M$  of all possible particle types, of the number  $N_m$  for each particle type  $m$  from the set  $M$ , the absolute temperature  $T$  of the system and the transmitter volume  $V_T$ , expressed as

$$H_T = \sum_{m \in M} N_m \left[ \log_2 \left( \frac{V_T}{N_m} \left( \frac{2\pi m K_b T}{h^2} \right)^{\frac{3}{2}} \right) + \frac{5}{2} \right], \quad (146)$$

where the transmitter volume  $V_T$  is

$$V_T = \frac{4}{3} \pi r_T^3. \quad (147)$$

Equation (146) takes into account that, according to the Dalton's law [80], in thermodynamics the contributions to the entropy coming from different types of particles of ideal gases (mixture of ideal gases) are independent. As a consequence, (146) is the sum of each contribution coming from the application of (145) to each particle type at the transmitter.

The total entropy  $H_{ref}^{mod}$  of the physical reference model when the transmitter is modulating the particle features can be written as the sum of two contributions. The first contribution is the transmitter local entropy  $H_T$  from (146). The second contribution is the entropy  $H_{ref}$  of the physical reference model from (145) without the transmitter modulation, expressed as

$$H_{ref}^{mod} = H_T + H_{ref}. \quad (148)$$

The input signal entropy corresponds to a measure of the amount of information contained in the input signal. The formula in (148) contains the information entropy of the system composed by the transmitter and the physical reference model. In practice, the input signal does not contain all the entropy  $H_{ref}^{mod}$  since the transmitter is not able to control the exact value of the phase space  $\Phi$  of the system. While the transmitter modulates, it inserts information in the system. This information, according to the Second Law of Thermodynamics [80], eventually will fade out when the system will reach a new thermodynamic equilibrium state, characterized by a higher entropy. In the thermodynamic equilibrium all the thermodynamic variables are constant and homogeneous and, therefore,

there is no transmitted information. The entropy  $H_{ref}^{new}$  of the physical reference model in the new thermodynamic equilibrium is computed from the formula in (145), expressed as

$$H_{ref}^{new} = N_{tot} \left[ \log_2 \left( \frac{V}{N_{tot}} \left( \frac{2\pi m K_b T}{h^2} \right)^{\frac{3}{2}} \right) + \frac{5}{2} \right], \quad (149)$$

where  $N_{tot}$  is the total number of particles in the system, expressed as

$$N_{tot} = N_p + \sum_{m \in M} N_m. \quad (150)$$

$N_p$  is the number of particles present in the physical reference model before modulation. In order to quantify the input signal entropy, we subtract the total entropy  $H_{ref}^{mod}$  of the physical reference model at modulation from the entropy  $H_{ref}^{new}$  of the physical reference model in the new thermodynamic equilibrium, expressed as follows:

$$H(x) = H_{ref}^{new} - H_{ref}^{mod}, \quad (151)$$

where  $H(x)$  is the input signal entropy in [*bit/sample*].

If we consider a bandwidth  $W$  for the system, we can transmit  $2W$  samples per second without equivocation. According to (140) the input signal entropy in [*bit/sec*] becomes

$$H'(x, W) = 2W \left( H_{ref}^{new} - H_{ref}^{mod} \right). \quad (152)$$

#### 5.2.4 The Equivocation

The received signal in the physical reference model introduced at the beginning of Section 5.2 corresponds to the reading of the changes in the **molecular features** operated by the **receiver**. As described at the beginning of Section 5.2, the receiver is considered as point-wise, therefore it only reads the changes in the molecular features that occur at the exact receiver location.

The propagation of the signal from the transmitter to the receiver affects the value of the local entropy at the receiver. The variation in the local entropy at the receiver  $H_R$  at instant  $t$  and distance  $d$  corresponds to the variation in the entropy of the spherical surface

at instant  $t$  and distance  $d$ , divided by the spherical surface area  $4\pi d^2$  as

$$H_R = \frac{H_R^{sph}}{4\pi d^2}. \quad (153)$$

The variation in the entropy in a spherical surface at distance  $d$  from the transmitter and at time  $t$  can be computed as the entropy of an ideal gas from (145) with an equivalent number of particles  $N_{eq}^m$  and an equivalent volume  $V_{eq}$ . The variation in the entropy  $H_R^{sph}(d, t)$  in a spherical surface at instant  $t$  at distance  $d$  is

$$H_R^{sph} = \sum_{m \in M} N_{eq}^m \left[ \log_2 \left( \frac{V_{eq}}{N_{eq}^m} \left( \frac{2\pi m K_b T}{h^2} \right)^{\frac{3}{2}} \right) + \frac{5}{2} \right]. \quad (154)$$

The equivalent number of particles  $N_{eq}^m$  of the ideal gas corresponds to the number of transmitted particles of type  $m$  that in a time  $t$  reach a distance  $d$  from the transmitter. This can be computed from the particle diffusion process that is at the basis of the particle propagation. The number of particles  $N_R$  that diffuse from the transmitter to the receiver located at a distance  $d$  is equal to the number  $N_T$  of the particles at the transmitter multiplied by the probability for each particle of shifting by a distance  $d$  at a time instant  $t$  after the instant of transmission, given by (136). This is expressed as

$$N_{eq}^m = N_m \frac{1}{\sqrt{4\pi D t}} e^{-\frac{d^2}{4Dt}}. \quad (155)$$

The value of the equivalent volume  $V_{eq}$  divided by the equivalent number of particles  $N_{eq}^m$  in (154) is the inverse of a the concentration of type  $m$  particles at the spherical surface at instant  $t$  and distance  $d$ . This concentration is equal to the concentration  $N_m/V_T$  of type  $m$  particles inside the transmitter multiplied by the probability in (136)

$$\frac{N_{eq}^m}{V_{eq}} = \frac{N_m}{V_T} \frac{1}{\sqrt{4\pi D t}} e^{-\frac{d^2}{4Dt}} = \frac{N_{eq}^m}{V_T}. \quad (156)$$

The equivocation formula in (141) corresponds to the entropy of the input signal when the receiver knows the value of the output signal. As a consequence, if the receiver is placed at a distance  $0 \leq d \leq r_T$  from the transmitter, then the increase in entropy is zero, since the output signal corresponds to the input signal. If the receiver is at a distance  $d \geq r_T$  from the

transmitter, then the increase in the entropy at instant  $t$  depends on the number of particles that reach the receiver location in a time  $t$ . The increase in the entropy at the receiver corresponds exactly to the equivocation. In other words, the equivocation is the entropy increase in the transmitted signal during propagation, computed by subtracting the entropy per unit volume  $\frac{H_T}{V_T}$  of the transmitted signal from the entropy  $H_R(d, t)$  of the received signal

$$H_Y(x) = H_R(d, t) - \frac{H_T}{V_T}, \quad (157)$$

where  $H_R$  is computed through (153),  $H_T$  through (146) and  $V_T$  through (147).

If we consider a bandwidth  $W$  for the system, we can receive up to  $2W$  samples per second without equivocation. In order to find the expression of the equivocation as a function of the bandwidth  $W$  it is necessary to find the number of particles  $N_R(d, W)$  that diffuse from the transmitter to the receiver located at a distance  $d$  and for a bandwidth  $W$ . This corresponds to (155), where the time variable is substituted with the inverse of the frequency of sample reception  $2W$ , that corresponds to the time interval between two consecutive transmitted samples. This is valid only under the hypothesis of not having influences between the diffusion of two different samples. In practice, this is not physically realistic since particles diffusing after the transmission of a sample will inevitably be present in the system at the time of transmission of the consecutive sample. Since taking into account this latter effect would complicate more the treatment concerning the equivocation and since the equivocation will result in a higher value, in this section of the Ph.D. thesis we rely on the assumption of not having influences between consecutive samples. **This assumption will result in an overestimation of the capacity (upper bound to the true capacity).**

The variation in the entropy  $H_{R,W}^{sph}$  in a spherical surface with bandwidth  $W$  at distance  $d$  is computed through

$$H_{R,W}^{sph} = \sum_{m \in M} N_{eq,W}^m \left[ \log_2 \left( \frac{V_{eq}}{N_{eq,W}^m} \left( \frac{2\pi m K_b T}{h^2} \right)^{\frac{3}{2}} \right) + \frac{5}{2} \right], \quad (158)$$

where  $N_{eq,W}^m$  is equal to the number of particles that diffuse from the transmitter to the

receiver located at a distance  $d$  and for a bandwidth  $W$

$$N_{eq,W}^m = N_m \sqrt{\frac{2W}{4\pi D}} e^{-\frac{2Wd^2}{4D}}. \quad (159)$$

As a consequence, the variation in the local entropy at the receiver  $H_R(W)$  with bandwidth  $W$  and distance  $d$  is

$$H_R(W) = \frac{H_{R,W}^{sph}}{4\pi d^2}. \quad (160)$$

The expression for the equivocation formula in [bit/sec] becomes

$$H_Y(x, W) = 2W \left( H_R(W) - \frac{H_T}{V_T} \right), \quad (161)$$

where  $H_R(W)$  is computed through (160),  $H_T$  through (146) and  $V_T$  through (147).

### 5.2.5 The Capacity

The information capacity of the diffusion-based MC system, given the physical reference model detailed at the beginning of Section 5.2, is expressed by the formula in (183), where the input signal entropy  $H(x)$  is given by (151) and the equivocation is given by (157). The final upper-bound expression of the capacity becomes

$$C = \max_{f_X(x)} \left\{ H_{ref}^{new} - H_{ref}^{mod} - \left( H_R - \frac{H_T}{V_T} \right) \right\}, \quad (162)$$

where  $H_{ref}^{new}$  is computed through (149),  $H_{ref}^{mod}$  through (148),  $H_R$  through (153) and  $H_T$  through (146). The transmitter volume  $V_T$  is computed through (147).

If we consider a bandwidth  $W$  for the system, the upper-bound expression of the information capacity  $C(W)$  of the diffusion-based MC system is expressed as a function of the bandwidth  $W$  by the formula in (183), where the input signal entropy  $H(x, W)$  is given by (166) and the equivocation  $H_Y(x, W)$  is given by (170), as follows:

$$C(W) = \max_{f_X(x)} 2W \left\{ H_{ref}^{new} - H_{ref}^{mod} - \left( H_R(W) - \frac{H_T}{V_T} \right) \right\}, \quad (163)$$

where  $H_{ref}^{new}$  is computed through (149),  $H_{ref}^{mod}$  through (148),  $H_R(W)$  through (160) and  $H_T$  through (146). The transmitter volume  $V_T$  is computed through (147).

A closed-form expression for the upper bound to the capacity of the diffusion-based MC system is given by the input signal probability density function  $f_X(x)$  that maximizes the quantity  $H(x) - H_Y(x)$  from (183). Such a value for  $f_X(x)$  can be found by setting a constraint on the total transmitted power. The value for  $f_X(x)$  corresponds to the distribution of the total transmitted power among the possible values for the input signal that gives the maximum input signal entropy  $H(x)$  and the minimum equivocation  $H_Y(x)$ . This distribution is found by associating the input signal entropy  $H(x)$  to the power  $P_{\mathcal{H}}$ .  $P_{\mathcal{H}}$  is the transmitter enthalpy power. The transmitter enthalpy power corresponds to the transmitted power to perform diffusion-based MC. If we consider a bandwidth  $W$  for the system, the transmitter enthalpy power  $P_{\mathcal{H}}$  becomes

$$P_{\mathcal{H}} = 2\mathcal{H}W, \quad (164)$$

where  $\mathcal{H}$  is the transmitter enthalpy.

**Definition 5.2.3** *The transmitter enthalpy is defined as the energy necessary to insert  $N$  particles in the system and to heat these particles up to a temperature  $T$  when the system has the pressure  $P$  and the volume  $V$ . In our case, the transmitter enthalpy can be computed as*

$$\mathcal{H} = PV + \frac{3}{2}K_bT \sum_{m \in M} N_m, \quad (165)$$

where  $P$  and  $V$  are the pressure and the volume of the physical reference model, respectively.  $M$  is the set of all possible particle types that the transmitter can emit,  $N_m$  is the number of particles of type  $m$ ,  $K_b$  is the Boltzmann constant and  $T$  is the absolute temperature of the system.

In order to find a closed-form expression for the upper bound to the capacity starting from (175), it is necessary to find a distribution of the transmitter enthalpy power among the  $M$  particle types so that the input signal entropy  $H(x)$  from (151) is maximized and the equivocation  $H_Y(x)$  from (157) is minimized.

The input signal entropy  $H(x)$  as function of the transmitter enthalpy power is

$$H(x, W) = 2W \left( H_{ref}^{new} - H_{ref}^{mod} \right), \quad (166)$$

where  $H_{ref}^{new}$  is computed through (149) and  $H_{ref}^{mod}$  through (148). The expression of the entropy  $H_{ref}^{mod}$  will depend on the fraction of the power  $P_{\mathcal{H}}^m$  assigned to each particle type  $m$ . The entropy  $H_{ref}^{mod}$  can be written as function of the transmitter power by expressing  $N_m$  as function of  $P_{\mathcal{H}}^m$  as follows:

$$H_T = \sum_{m \in M} N_m^P \left[ \log_2 \left( \frac{V_T}{N_m^P} \left( \frac{2\pi m K_b T}{h^2} \right)^{\frac{3}{2}} \right) + \frac{5}{2} \right], \quad (167)$$

where  $M$  is the set of possible input signals,  $V_T$  is the receiver volume from (147),  $K_b$  is the Boltzmann constant,  $T$  is the absolute temperature of the system,  $m$  is the particle mass,  $h$  is the Planck constant and  $N_m^P$  is given by

$$N_m^P = \frac{P_{\mathcal{H}}^m - 2WPV}{3WK_bT}, \quad (168)$$

where  $P_{\mathcal{H}}^m$  is the fraction of the power assigned to each particle type  $m$  and  $P$  and  $V$  are the pressure and the volume of the physical reference model, respectively. The maximum of the input signal entropy  $H(x, W)$  corresponds to the distribution of the power that results in the minimum of  $H_T$  from (167). The minimum of  $H_T$ , denoted by  $H_T^{min}$ , corresponds to an even distribution of the power among all the  $M$  types of particles that the transmitter can send. Therefore, for each particle type out of the  $M$  possible, the fraction of power  $P_{\mathcal{H}}^m$  is given by the total transmitter enthalpy power  $P_{\mathcal{H}}$  divided by the number of particle types  $M$ , as follows:

$$P_{\mathcal{H}}^m = \frac{P_{\mathcal{H}}}{M} \rightarrow H_T = H_T^{min}. \quad (169)$$

The equivocation  $H_Y(x)$  as function of the transmitter enthalpy power is

$$H_Y(x, W) = 2W \left( H_R(W) - \frac{H_T}{V_T} \right). \quad (170)$$

$H_R(W)$  is computed through (160) and  $H_T$  through (146). Both  $H_T$  and  $H_R(W)$  depend on the fraction of power  $P_{\mathcal{H}}^m$  assigned to each particle type  $m$ .  $H_T$  is expressed by (167) and

$H_R(W)$  is given by (160) where  $H_{R,W}^{sph}$  is

$$H_{R,W}^{sph} = \sum_{m \in M} N_{eq,P_{\mathcal{H}}}^m \left[ \log_2 \left( \frac{V_T}{N_{eq,P_{\mathcal{H}}}^m} \left( \frac{2\pi m K_b T}{h^2} \right)^{\frac{3}{2}} \right) + \frac{5}{2} \right], \quad (171)$$

where  $N_{eq,P_{\mathcal{H}}}^m$  is

$$N_{eq,P_{\mathcal{H}}}^m = \frac{P_{\mathcal{H}}^m - 2WPV_T}{(3/2)K_b T}. \quad (172)$$

The optimal distribution of power in (169) minimizes  $\frac{H_r}{V_T}$ , and it minimizes also  $H_{R,W}^{sph}$  and, consequently,  $H_R(W)$ .  $H_R(W)$ , expressed in (160), corresponds to the entropy of the transmitted particles per unit volume at a distance  $d$  and time  $t$ , while  $\frac{H_r}{V_T}$  is the entropy per unit volume at the transmitter at time  $t = 0$ . We can logically assume that if  $d$  is sufficiently large, the entropy  $\frac{H_r}{V_T}$  is negligible with respect to the contribution of  $H_R(W)$ , expressed as

$$H_R(W) \gg \frac{H_T}{V_T}; \quad (173)$$

which is valid also for large values of  $W$ , as can be observed by computing the limit for  $W \rightarrow \infty$  of the expression in (160), after substitution with (158) and (159). Therefore, the minimum of the equivocation  $H_Y(x)$  corresponds roughly to the minimum of  $H_R(W)$ , denoted as  $H_R^{min}(W)$ , which is again given by the even power distribution in (169).

A closed form expression for the upper bound to the capacity of the diffusion-based MC system is

$$C(W) = 2W \left[ H_{ref}^{new} - H_{ref}^{mod,min} - \left( H_R^{min}(W) - \frac{H_T^{min}}{V_T} \right) \right], \quad (174)$$

where  $H_{ref}^{new}$  is the computed through (149),  $H_R(W)$  through (160) and (149), and  $H_T$  through (146), (149) and (146) are evaluated with (169). The transmitter volume  $V_T$  is computed through (147) and  $H_{ref}^{mod,min}$  is

$$H_{ref}^{mod,min} = H_T^{min} + H_{ref}, \quad (175)$$

where  $H_{ref}$  is detailed in (145).

## 5.2.6 Numerical Results

In this section, we provide numerical results for the upper bound expression to the capacity in MC nanonetworks. All the results are computed for a common set of parameters, whose

values are assigned as follows. The total transmitter enthalpy power  $P_{\mathcal{H}}$  is set to  $1 \mu\text{W}$ . Note that the latter is only a reference value, since so far we do not know how much transmitter power a nanomachine will be able to provide. The radius  $r_S$  of the spherical space  $S$  in which the system is immersed is set to  $1\text{cm}$ , while the radius of the transmitter  $r_T$  is set to  $1 \mu\text{m}$ . The mass of the particles  $m$  is set to  $1.66053878283 \times 10^{-27} \text{ kg}$ , which is the standard atomic mass unit [82]. The number  $N_p$  of particles present in the system is set equal to a number of molecules in a mole [80]  $6.0221417930 \times 10^{23}$ . We set the number  $M$  of molecules types that the transmitter can release to 5. The Boltzmann constant [81] is  $K_b = 1.380650424 \times 10^{-23} [\text{Joule}/\text{K}]$  and the Planck constant is [76]  $h = 6.6260689633 \times 10^{-34} [\text{Joule} \cdot \text{sec}]$ .

All the results come from the evaluation of (175), (160), (149) and (146), given the condition of having an even distribution of the transmitted power among all the types of particles, (169).

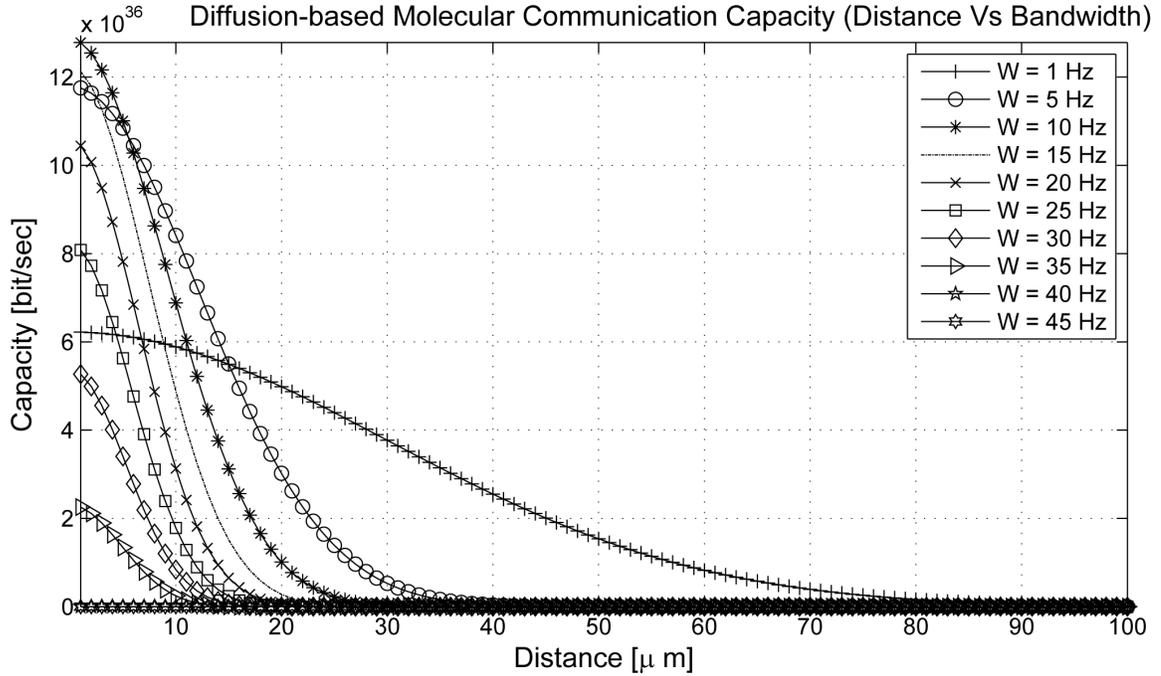
The results in Figure 52, as well as Figure 53 and Figure 54, show extremely high values for the capacity which are on the order of magnitude of  $10^{36} [\text{bit}/\text{sec}]$ .

Although those numbers refer to an upper bound to the true capacity, which could be achieved only with a theoretical optimal modulation scheme at the transmitter, they can be physically explained as follows.

The computation of the capacity in this section of the Ph.D. thesis considers a transmitted sample as any combination of the number of particles of any possible type out of  $M$ , bounded by the transmitter power using (168).

In our numerical analysis, by taking into account (169) and (168) we achieve a maximum number of molecules  $N_m^P$  on the order of magnitude of  $10^{12}$  per each type  $m$ .

Therefore, the number of combinations we can achieve using the optimal modulation scheme is extremely high and, consequently, also the capacity.



**Figure 52.** Capacity in relation to the transmitter-receiver Distance and for different values of the bandwidth  $W$ .

#### 5.2.6.1 Distance Vs Bandwidth

In Figure 52 we show the values of the upper-bound expression of the capacity of a diffusion-based MC nanonetwork in relation to the distance  $d$  between the transmitter and the receiver locations and different values of the bandwidth  $W$ .

We evaluate the capacity in  $[bit/sec]$  for a distance ranging from  $1 \mu m$  to  $100 \mu m$ . The different lines refer to different bandwidth  $W$  values, from 1 Hz to 45 Hz. The temperature  $T$  of the system is set to a standard room temperature of  $25^\circ C$  and the diffusion coefficient  $D$  is set to  $10^{-9} [m^2/sec]$ , as explained in Chapter 3.f

The curves in Figure 52 show the maximum values of the upper-bound to the true capacity for every value of the bandwidth when the distance between the transmitter and the receiver locations is equal to the minimum considered value, namely,  $1 \mu m$ . As the bandwidth  $W$  increases, the Gaussian shape of the curves, related to (136), decreases its variance and the maximum capacity value when the distance is  $1 \mu m$ . This behavior can be explained by noticing that, as the bandwidth of the input signal  $W$  increases, the higher

frequencies are affected by higher attenuation in the channel, and the received signal power, whose average is kept constant at the transmitter, gets lower. As a consequence, increasing the bandwidth  $W$  shifts also the maximum distance with non-zero capacity closer to the transmitter.

### 5.2.6.2 Bandwidth Vs Distance

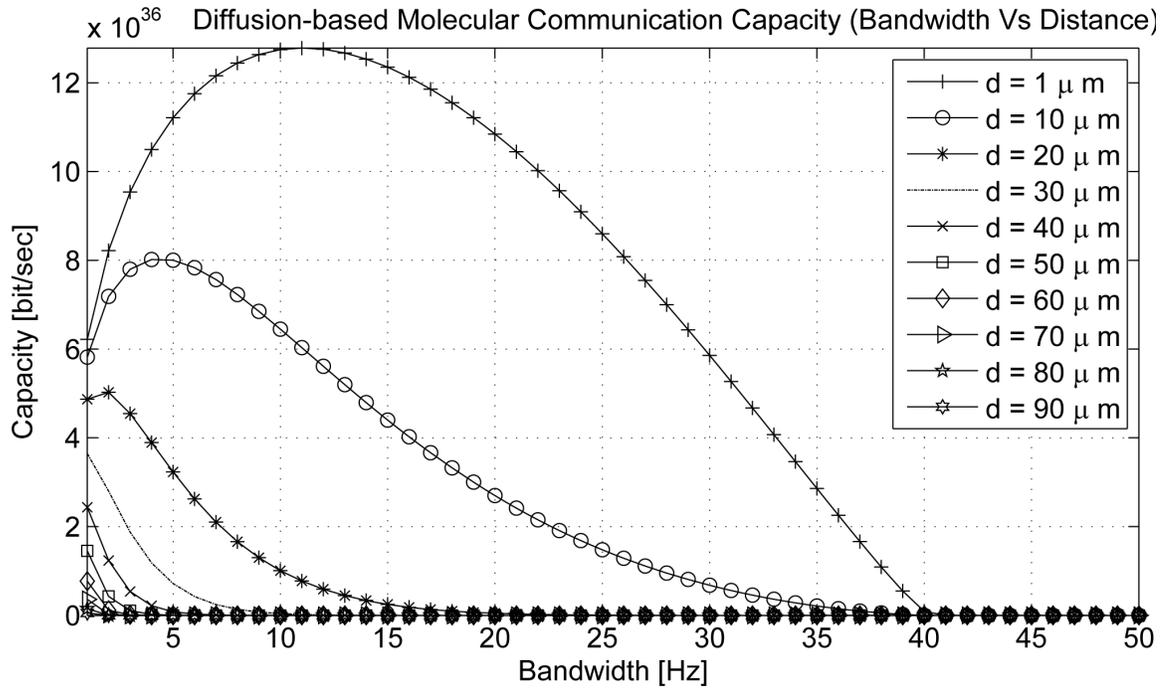


Figure 53. Capacity in relation to the Bandwidth and for different values of the transmitter-receiver distance  $d$ .

The upper-bound expression of the capacity of a diffusion-based MC system is shown in Figure 53 in relation to the bandwidth  $W$  and different values of the transmitter-receiver distance  $d$ .

We evaluate the capacity in [bit/sec] for a bandwidth  $W$  ranging from 1 Hz to 45 Hz and different lines refer to different distance values, from  $1 \mu\text{m}$  to  $100 \mu\text{m}$ . The temperature  $T$  of the system is set to a standard room temperature of  $25^\circ\text{C}$  and the diffusion coefficient  $D$  is set to  $10^{-9}$  [ $\text{m}^2/\text{sec}$ ], as explained in Chapter 3.

Figure 53 shows a bandwidth value which maximizes the capacity. This maximum capacity value appears only for short distances  $d < 30 \mu\text{m}$ , it is different for each distance  $d$

value and it decreases as the distance increases. As a consequence, we understand that for each distance value there is an optimal bandwidth in Hz that minimizes the equivocation in diffusion-based molecular communication.

The curves in Figure 53 show also a maximum bandwidth value over which the capacity is zero. This value is not changed by different distance values and for every curve it is around 40 Hz.

### 5.2.6.3 Bandwidth Vs Temperature

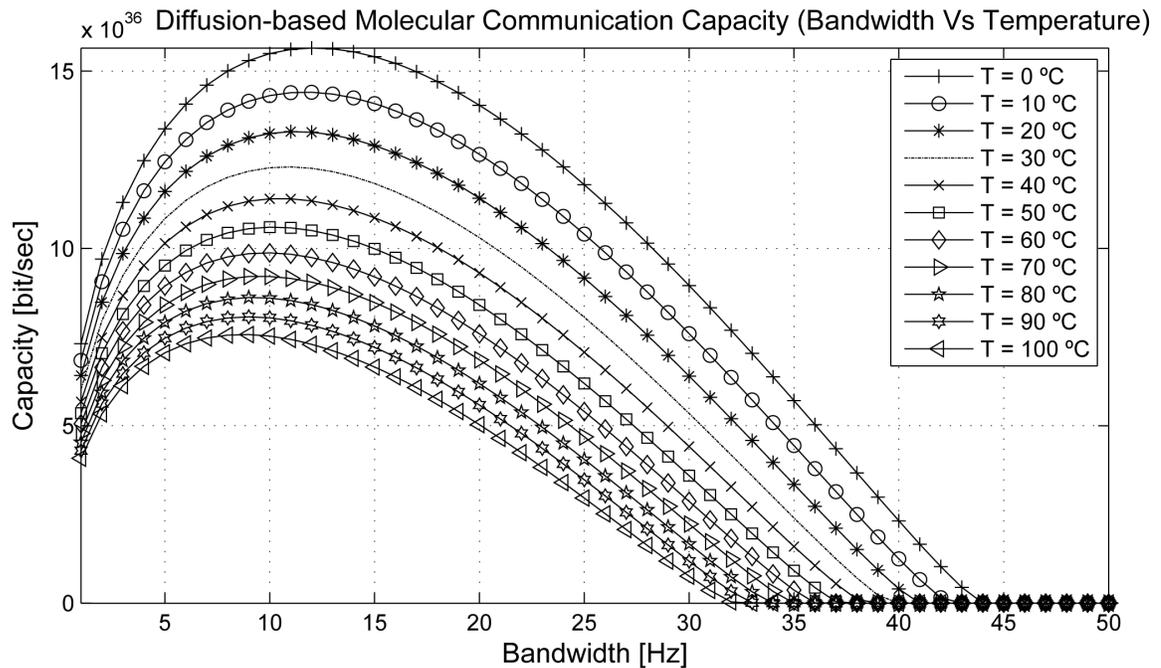


Figure 54. Capacity in relation to the Bandwidth and different values of the system temperature  $T$ .

In Figure 54 we show the upper-bound expression of the capacity dependent on the bandwidth  $W$  ranging from 1 Hz to 45 Hz and the temperature  $T$ . Different lines refer to different system temperature  $T$  values, from 0 °C to 100 °C. The distance between the transmitter and the receiver locations is set to 50  $\mu\text{m}$  and the diffusion coefficient  $D$  is set to  $10^{-9}$  [ $\text{m}^2/\text{sec}$ ], as explained in Chapter 3, for a temperature of 25 °C as a reference, and it is varied according to the actual temperature values by following the Einstein-Stokes equation [79].

Figure 54 shows a bandwidth value which maximizes the capacity. This maximum

capacity value is different for each temperature  $T$  value and it decreases as the temperature increases. As a consequence, we understand that for each temperature value there is an optimal bandwidth in Hz that minimizes the equivocation in diffusion-based molecular communication.

The curves in Figure 54 show also a maximum bandwidth value over which the capacity is zero. This value is different for each temperature and it ranges from 32 Hz at 100 °C to 44 Hz at 0 °C.

### 5.2.7 Conclusion

In this section of the Ph.D. thesis, a mathematical expression of an upper bound to the information-theoretic capacity in diffusion-based MC is provided as a function the pressure, the volume, the temperature, the number of molecules, the bandwidth of the system, and the transmitter power. Further investigation will be carried out in the future on finding more stringent upper bounds to the performance. The most interesting result stands in the order of magnitude of these capacity values, since they are extremely high if compared to capacity values in classical EM-communication systems. These results confirm the validity of the sentence by R. Feynman "There's Plenty of Room at the Bottom" [83], and explain the growing interest around diffusion-based MC shown by the research community in the last couple years.

## 5.3 Capacity Analysis with Channel Memory and Molecular Noise

Differently from the previous section, the **Physical System** considered in this section of the Ph.D. thesis is sketched in Figure 55 and it is based on the following considerations:

- The diffusion-based MC channel is in a three-dimensional space indexed by the three axes X, Y, Z and it has infinite extent in all three dimensions. This space is filled with a fluidic medium having viscosity  $\mu$ . The fluidic medium does not have flow currents or turbulence, therefore the propagation of the molecules between the transmitter and the receiver is solely realized by the Brownian motion.

- All the molecules in the system, which are emitted by the transmitter, are indistinguishable and equivalent to spherical *particles* of radius  $r$  and mass  $m$ , where  $r \ll d$ ,  $d$  being the distance between the transmitter and the receiver in the diffusion-based MC system. As a consequence, from now on we will refer to particles when talking about molecules in the physical system.
- The transmitter is considered point-wise (size equal to zero) and at location  $\mathbf{T} = (T_X, T_Y, T_Z)$  in the three-dimensional space.
- Once emitted from the transmitter, every particle moves independently from the others and according to its Brownian motion in the fluidic medium. The Brownian motion of a molecule is referred to as the random motion of the particles suspended in a fluid and its formulation according to the Langevin equation [84] states that the location  $p_i^n(t)$  of the particle  $n$  at time  $t$  along any  $i$  of the three dimensional axes X, Y, Z obeys the following stochastic differential equation:

$$m \frac{\partial^2 p_i^n(t)}{\partial t^2} = -6\pi\mu r \frac{\partial p_i^n(t)}{\partial t} + f_i(t), \quad i \in \{X, Y, Z\}, \quad (176)$$

where  $m$  is the particle mass,  $\partial^2(\cdot)/\partial t^2$  and  $\partial(\cdot)/\partial t$  are the second and first time derivative operators, respectively,  $\mu$  is the viscosity of the fluid,  $r$  the radius of the particle and  $f_i(t)$  is a random process whose probability density function is Gaussian and has correlation function  $\langle f_i(t)f_j(t') \rangle$  given by

$$\langle f_i(t)f_j(t') \rangle = 12\pi\mu r k_B T \delta_{i,j} \delta(t - t') \quad i, j \in \{X, Y, Z\}, \quad (177)$$

where  $\langle \cdot \rangle$  is the average operator,  $k_B$  is the Boltzmann constant,  $T$  is the absolute temperature of the fluid, considered homogeneous throughout the space, and  $\delta_{i,j}$  is equal to 1 if  $i = j$  and zero otherwise;  $\delta(t - t')$  is the Dirac delta function.

- The receiver detects a signal which is proportional to the concentration of the incoming particles. The receiver location is at a distance  $d$  from the transmitter.

In the following, the components included in the physical system are described in light of the aforementioned considerations.

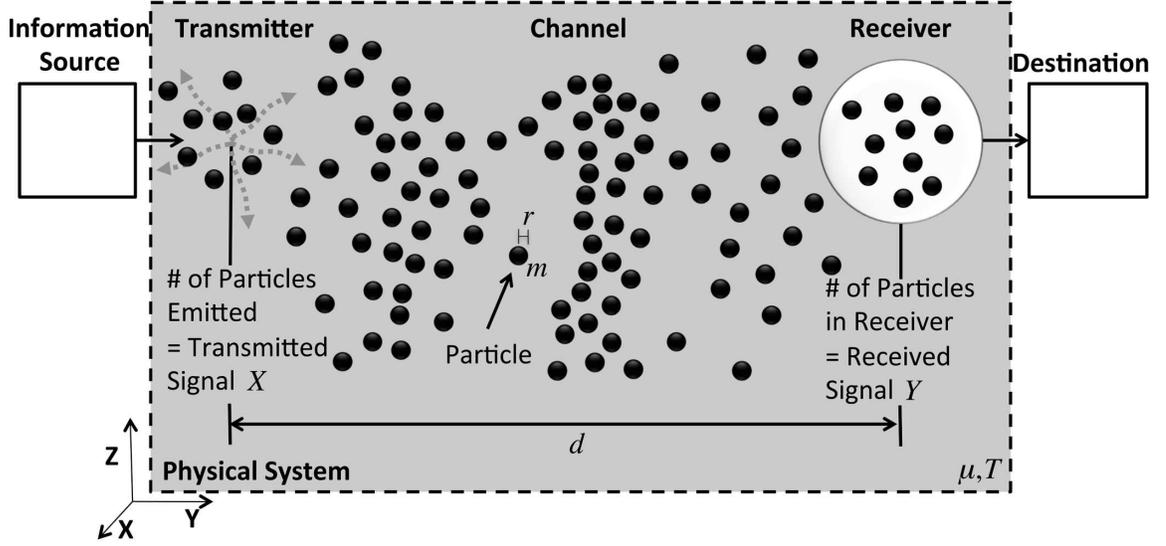


Figure 55. Sketch of the physical realization of the diffusion-based molecular communication system.

The **Transmitter** processes the messages from the information source and produces a signal suitable for the transmission over the channel. The *transmitted signal*, denoted by  $X$ <sup>1</sup>, is here defined as the number of particles  $n_T(t)$  emitted into the space as a function of the time  $t$ , expressed as

$$X := n_T(t), \quad t > 0. \quad (178)$$

At the time  $t$  of emission of a particle, denoted by  $\bar{n}$ , its location  $\mathbf{p}^{\bar{n}}(t) = (p_X^{\bar{n}}(t), p_Y^{\bar{n}}(t), p_Z^{\bar{n}}(t))$  corresponds to the location of the transmitter  $\mathbf{T} = (T_X, T_Y, T_Z)$ , expressed as

$$\mathbf{p}^{\bar{n}}(t) = \mathbf{T}, \quad \bar{n} = \int_0^t n_T(\tau) d\tau, \quad t > 0, \quad (179)$$

where  $\mathbf{T}$  is the vector of the three-dimensional coordinates  $(T_X, T_Y, T_Z)$  of the transmitter.  $\bar{n}$  is here an index assigned to each particle on the basis of the order in which they are emitted. This index serves only for the mathematical formulation of their propagation

<sup>1</sup>For the information capacity analysis of this communication system, the transmitted signal  $X$  is considered as a band-limited random process whose value at every time instant  $t$  is a realization of the random variable  $n_T(t)$ . As a consequence, the entropy of  $X$  is found by decomposing  $X$  into a band-limited ensemble of functions, as detailed in Section 5.3.3.

through the Langevin equation in (176), while, as mentioned above, particles are identical and indistinguishable in the physical system.

The **Channel** propagates the signal from the transmitter to the receiver by means of molecule diffusion, which is the result of the collective translation by Brownian motion of many particles from an area in which they are more dense to an area of lower density. This results in the propagation of the particles emitted by the transmitter throughout the three-dimensional space. This propagation can be expressed as the translation of the three-dimensional coordinates from the location  $\mathbf{T}$  of the transmitter to a location  $\mathbf{p}^n(t)$  at time  $t$  computed by applying (176) to each particle  $n$  from the set  $\mathcal{N}_T(t)$ , as

$$\mathbf{T} \rightarrow \mathbf{p}^n(t), \forall n \in \mathcal{N}_T(t), \quad (180)$$

where  $\mathcal{N}_T(t)$  is the set containing all the indexes of the particles emitted by the transmitter from time 0 to time  $t$ :

$$\mathcal{N}_T(t) = \left\{ \int_0^{t'} n_T(\tau) d\tau \mid 0 < t' < t \right\}. \quad (181)$$

The **Receiver** reconstructs the messages (sent by the transmitter) from the *received signal*  $Y$ , which is proportional to the concentration of incoming particles. In this section of the Ph.D. thesis, we assume an ideal receiver where the received signal  $Y$  is defined as the time-varying number of particles that are present inside a spherical volume  $V_R$  centered at the receiver location and with radius  $R_{V_R} \ll d$ , where  $d$  is the distance between the transmitter and the receiver. This choice makes the results of this section of the Ph.D. thesis independent from any specific techniques for the reception (e.g., the chemical ligand-binding reception detailed in Section 3.5). As a consequence, the received signal  $Y$  is expressed as the number of particles emitted by the transmitter from time instant 0 to time instant  $t$  whose location  $\mathbf{p}^n(t)$  is inside the volume  $V_R$ , as

$$Y := \# \{n \in \mathcal{N}_T(t) : \mathbf{p}^n(t) \in V_R\}, \quad t > 0, \quad (182)$$

where  $\#\{.\}$  stands for the cardinality (number of elements) of the set enclosed in the brackets.

### 5.3.1 Information Capacity of a Diffusion-based MC System

The **capacity**  $C$  of a communication system in [*bit/sec*] is defined as the maximum rate of transmission between the information source and the destination, where this maximum is with respect to all possible signals produced by the transmitter [69]. This is expressed by the general formula from Shannon [77], which defines the capacity as the maximum mutual information  $I(X; Y)$  between the transmitted signal  $X$  and the received signal  $Y$  with respect to the probability density function  $f_X(x)$  in all the possible values of the transmitted signal, expressed as follows:

$$C = \max_{f_X(x)} \{I(X; Y)\} . \quad (183)$$

The **mutual information**  $I(X; Y)$  in [*bit/sec*] is defined as:

$$I(X; Y) = H(X) - H(X|Y) = H(Y) - H(Y|X) = H(X) + H(Y) - H(X, Y) , \quad (184)$$

where  $H(X)$  is the entropy per second of the transmitted signal  $X$ , defined in Section 5.3.3,  $H(X|Y)$  is the entropy per second of the transmitted signal  $X$  given the received signal  $Y$ ,  $H(Y|X)$  is the entropy per second of the received signal  $Y$  given the transmitted signal  $X$ , and  $H(X, Y)$  is the joint entropy per second of the transmitted signal  $X$  and the received signal  $Y$ .

In the following, we analytically compute the mutual information of a molecular communication system, as expressed by (183), by considering the transmitter, the channel and the receiver, defined through (317), (180) and (182), when evaluating (184). From the physical system defined at the beginning of Section 5.3, two phenomena play an important role in the quantification of the mutual information, namely, the channel memory and the molecular noise, as we highlight in Section 5.3.2. For this, we propose to divide the computation of the mutual information into two processes, namely, the Fick's diffusion, treated in Section 5.3.3, which captures solely the effects of the channel memory, and the particle location displacement, treated in Section 5.3.4, which isolates the effects of the molecular noise.

### 5.3.2 The Molecule Diffusion as Fick's Diffusion and Particle Location Displacement

The Langevin equation in (176) is the most general expression of the molecule diffusion due to the Brownian motion. In a MC system, it impacts on the communication performance (mutual information and capacity) through the following two phenomena:

- **Channel memory:** it is the effect of the persistent presence in the three-dimensional space of the particles from the moment they are emitted by the transmitter until infinite time. This is a consequence of the fact that in the physical system considered in this section of the Ph.D. thesis each emitted particle is subject to the Brownian motion. For this, each particle wanders randomly in the three-dimensional space without being destroyed. This is expressed through a positive probability of having any of the emitted particles at any time after the emission instant inside the receiver volume, expressed as

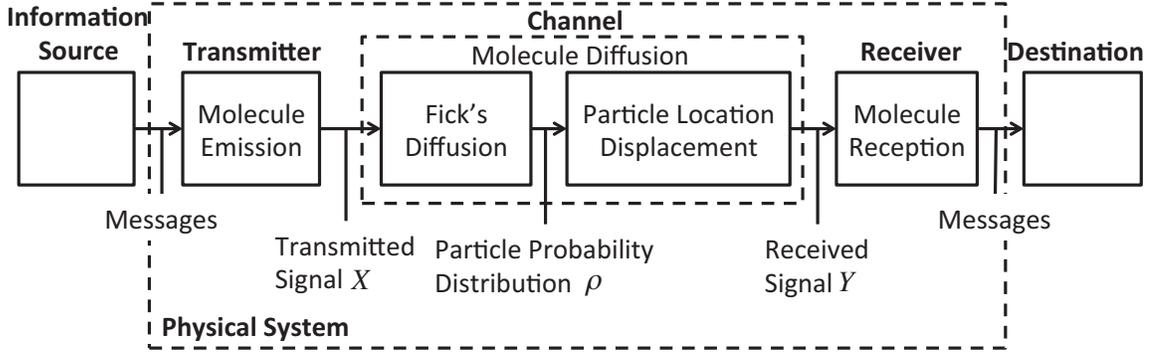
$$P(n \in \mathcal{N}_T(t) : \mathbf{p}^n(t) \in V_R) > 0 \quad \forall n, t > 0, \quad (185)$$

where  $\mathcal{N}_T(t)$  is given by (181),  $\mathbf{p}^n(t)$  is the vector with the location coordinates for the particle  $n$  at time  $t$  and  $V_R$  is the set containing all the space coordinates included in the receiver volume.

- **Molecular noise:** it is the effect of the randomness of the particle locations in the three-dimensional space, which results in random fluctuations of the received signal. This is a consequence of the random process  $f_i(t)$  of the particle locations expressed in (176). This is expressed by considering the received signal  $Y$  as a random variable with a generic distribution  $F$ . Its expected value  $E[Y]$  is the integral of the expected particle distribution, denoted by  $\rho(\mathbf{p}, t)$ , integrated in the receiver volume  $V_R$ , as

$$Y \sim F, \quad E[Y] = \int_{V_R} \rho(\mathbf{p}, t) d\mathbf{p}, \quad t > 0, \quad (186)$$

where  $\rho(\mathbf{p}, t)$  is the particle distribution at location  $\mathbf{p} = (p_X, p_Y, p_Z)$  and time  $t$ , whose equation will be defined in the following.



**Figure 56. Diagram of the diffusion-based MC system with the Fick's diffusion and the particle location displacement contributions to the molecule diffusion.**

In this section of the Ph.D. thesis, we propose to analyze the impact of the aforementioned phenomena on the mutual information (184) by separating the molecule diffusion from the Langevin equation (176) into two processes, namely, the **Fick's diffusion** and the **particle location displacement**, as shown in Figure 56. This is possible since the molecule diffusion expressed by the stochastic differential equation in (176) and having  $X$  (317) as input and  $Y$  (182) as output can be equivalently expressed by the deterministic Fick's equation [27] followed by a stochastic process which results in the assignment of the particle locations in the three-dimensional space.

The Fick's equation is a parabolic partial differential equation [27] in the variable  $\rho(\mathbf{p}, t)$ , which is the particle distribution at location  $\mathbf{p} = (p_X, p_Y, p_Z)$  and time  $t$ . The expression of this equation for the diffusion-based MC system accounts for the transmitter as a source of particles at location  $\mathbf{T}$ . This translates into an additional term, namely,  $n_T(t)\delta(|\mathbf{p} - \mathbf{T}|)$ , which corresponds to the number of particles  $n_T(t)$  emitted into the space as a function of the time  $t$  at the location  $\mathbf{T}$ , where the Dirac delta  $\delta(|\mathbf{p} - \mathbf{T}|)$  is non-zero. We express the Fick's equation as follows:

$$\frac{\partial \rho(\mathbf{p}, t)}{\partial t} = D\nabla^2 \rho(\mathbf{p}, t) + n_T(t)\delta(|\mathbf{p} - \mathbf{T}|), \quad t > 0, \quad (187)$$

where  $\partial./\partial t$  is the time derivative operator of the particle distribution  $\rho(\mathbf{p}, t)$ , which corresponds to the expected number of particles at location  $\mathbf{p}$  and time  $t$ , and  $\nabla^2$  is the Laplacian

operator.  $D$  is the particle diffusion coefficient, whose expression is as follows:

$$D = \frac{K_b T}{6\pi\mu r}, \quad (188)$$

where  $K_b$  is the Boltzmann constant,  $T$  is the absolute temperature of the system,  $\mu$  is the viscosity of the fluid and  $r$  is the particle radius.

The particle location displacement is expressed through the stochastic process that randomly assigns the location to each transmitted particle according to the particle distribution  $\rho(\mathbf{p}, t)$  at each time instant  $t$ , as

$$\mathbf{p}^n(t) \sim \rho(\mathbf{p}, t), \forall n \in \mathcal{N}_T(t), \quad (189)$$

where  $\mathcal{N}_T(t)$  is given by (181).

The channel memory phenomenon of the molecule diffusion introduced above is fully captured by the Fick's diffusion contribution. This is expressed by stating that the probability that the location of a particle is inside the receiver volume is never zero from the time instant of the particle emission until infinite time. This is detailed through the following relation:

$$\int_{V_R} \rho_{t'}(\mathbf{v}, t) d\mathbf{v} > 0, \quad \forall t, t' : t > t' > 0, \quad (190)$$

where the integral is performed by spanning the set  $V_R$  containing all the space coordinates included in the receiver volume.  $\rho_{t'}(\mathbf{v}, t)$  is the particle distribution. This is the solution of the Fick's equation in case the particles are emitted only at the time instant  $t'$

$$\frac{\partial \rho_{t'}(\mathbf{p}, t)}{\partial t} = D \nabla^2 \rho_{t'}(\mathbf{p}, t) + n_T(t') \delta(|\mathbf{p} - \mathbf{T}|) \delta(t - t'), \quad t > 0. \quad (191)$$

The molecular noise phenomenon is isolated into the particle location displacement contribution, since it contains the stochastic process which contributes to the Langevin equation (176). This is expressed by noting that the number of the particles whose location  $\mathbf{p}^n(t)$  is within the receiver volume at time  $t$  is a realization of the particle location displacement, as expressed in (189).

The cascade of the Fick's diffusion and the particle location displacement contributions, as shown in Figure 56, define a Markov chain [77] in the variables  $X$ ,  $\rho$  and  $Y$  following the order  $X \rightarrow \rho \rightarrow Y$ . This is justified by the property that  $X$  and  $Y$  are conditionally independent given  $\rho$ , which is expressed as follows:

$$f_{X,Y|\rho}(x, y) = f_{X|\rho}(x) f_{Y|\rho}(y), \quad (192)$$

since  $\rho$  is function of  $X$  from (317) and (324), and the distribution of  $Y$  is a function of  $\rho$  from (182) and (189). The chain rule applied to the joint entropy of  $X$ ,  $\rho$  and  $Y$  states the following [77]:

$$H(X, \rho, Y) = H(X, Y|\rho) + H(\rho) = H(X|\rho) + H(Y|\rho) + H(\rho). \quad (193)$$

Since  $\rho$  is a deterministic function of  $X$  through the Fick's equation from (324), then the joint entropy per second of  $X$ ,  $\rho$  and  $Y$  is equal to the joint entropy per second of  $X$  and  $Y$ :

$$H(X, \rho, Y) = H(X, Y). \quad (194)$$

By applying (193) and (194) to the third expression in (184), we obtain that the mutual information  $I(X; Y)$  of the transmitted signal  $X$  and the received signal  $Y$  as the sum of the mutual information of a communication system which includes only the Fick's diffusion (mutual information  $I(X; \rho)$  of the transmitted signal and the particle distribution) and the mutual information of a system which includes only the particle location displacement (mutual information  $I(Y; \rho)$  of the received signal and the particle distribution), respectively, with the subtraction of the entropy per second  $H(\rho)$  of the particle distribution. This is expressed as follows:

$$I(X; Y) = H(X) + H(Y) - H(X|\rho) - H(Y|\rho) - H(\rho) = I(X; \rho) + I(Y; \rho) - H(\rho), \quad (195)$$

where we applied the first two definitions of mutual information from (184) to obtain the last expression.

We provide closed-form solutions to the mutual information of the Fick's diffusion and to the mutual information of the particle location displacement in Section 5.3.3 and Section 5.3.4, respectively. In Section 5.3.5 we apply (183) and (195) to the results of the previous sections to obtain a closed-form expression of a lower bound to the mutual information of the diffusion-based MC system and, ultimately, to obtain a lower bound to its capacity.

### 5.3.3 The Fick's Diffusion Mutual Information

The closed-form expression for the mutual information  $I(X; \rho)$  in  $[bit/sec]$  of the Fick's diffusion is computed by applying the following relation:

$$I(X; \rho) = H(X) - H(X|\rho), \quad (196)$$

where  $H(X)$  is the entropy per second of the transmitted signal  $X$  and  $H(X|\rho)$  is the conditional entropy per second of the transmitted signal  $X$  given the particle distribution  $\rho$ .

The entropy per second  $H(X)$  of the transmitted signal is computed as the entropy measured in  $[bit/symbol]$ , multiplied by twice the bandwidth  $W$ , which corresponds here to the rate of the symbol transmission in  $[symbol/sec]$ . This results from considering the transmitted signal  $X$  defined in (317) as a band-limited ensemble of functions [69] within a bandwidth  $W$ . The ensemble has the following expression:

$$X = \sum_{k=0}^{\infty} n_T \left( \frac{k}{2W} \right) \frac{\sin [\pi(2Wt - k)]}{\pi(2Wt - k)}, \quad k \in \mathbb{N}, \quad (197)$$

where the bandwidth  $W$  is here defined as the maximum frequency contained in the time-continuous signal  $n_T(t)$  (317), which corresponds to the number of emitted particles as function of the time  $t$ . The Shannon-Hartley theorem [77] assures the equivalence of the expressions in (197) and (317), respectively. As proven in [69], we can express the entropy per second  $H(X)$  of the transmitted signal  $X$  as the entropy of the ensemble per degree of freedom in  $[bit/sample]$  multiplied by twice the bandwidth  $W$  in  $[sample/sec]$ . The entropy of the ensemble per degree of freedom corresponds to the entropy  $H(\hat{n}_T)$  of a

sample  $n_T(k/2W)$  of the time-continuous signal  $n_T(t)$ , expressed as follows:

$$H(X) = 2W H(\hat{n}_T). \quad (198)$$

The distribution of the stochastic process model for the sampled signal  $\hat{n}_T$ , which allows to compute a lower-bound expression of the capacity  $C$  through (183), is assigned in Section 5.3.5 as the distribution leading to the maximum possible mutual information for the MC system constrained by the average power consumption for particle emission at the transmitter.

The conditional entropy per second  $H(X|\rho)$  of the transmitted signal  $X$  given the particle distribution  $\rho$  is computed as a result of the two following properties of the Fick's diffusion from (324):

- Its linearity, which allows to interpret the Fick's diffusion block in Figure 56 as a linear filter having the transmitted signal  $X$  as input and the particle distribution  $\rho$  as output. As a consequence, the formula of the entropy loss in linear filters [85] can be applied to compute the entropy per second  $H(\rho)$  of the particle distribution as the sum of the entropy per second  $H(X)$  of the transmitted signal and the integral of the transfer function Fourier transform [51] of the Green's function [86] of the Fick's diffusion in the portion  $W$  of its frequency spectrum that is excited by the transmitted signal  $X$ , expressed as

$$H(\rho) = H(X) + \frac{1}{W} \int_w \log_2 |G_d(f)|^2 df. \quad (199)$$

- Its deterministic nature, since (324), in contrast to the expression in (176), does not contain any stochastic term. For this, given the transmitted signal  $X$  as the input of the Fick's diffusion, the output particle distribution  $\rho$  is completely known. As a consequence, the conditional entropy per second  $H(\rho|X)$  of the particle distribution given the transmitted signal is equal to zero, expressed as

$$H(\rho|X) = 0. \quad (200)$$

Given the aforementioned properties, the conditional entropy per second  $H(X|\rho)$  of the transmitted signal  $X$  given the particle distribution  $\rho$  is computed by applying (199) and (200) to the following relation [77]:

$$H(\rho) = H(X) + H(\rho|X) - H(X|\rho), \quad (201)$$

which results in the following expression:

$$H(X|\rho) = -\frac{1}{W} \int_w \log_2 |G_d(f)|^2 df, \quad (202)$$

where  $G_d(f)$  is the transfer function Fourier transform [51] as function of the frequency  $f$  of the Green's function [86] of the Fick's diffusion, expressed by (324).  $W$  is the bandwidth of the transmitted signal  $X$ , which corresponds to the portion of the frequency spectrum of the transfer function  $G_d(f)$  that is excited by the transmitted signal  $X$ .

The transfer function Fourier transform [51] as function of the frequency  $f$  of the Green's function [86] of the Fick's diffusion from (324) has the following expression:

$$G_d(f) = \frac{e^{-(1+j)\sqrt{\frac{2\pi f}{2D}}d}}{\pi Dd}, \quad (203)$$

where  $d$  is the distance between the transmitter and the receiver and  $D$  is the diffusion coefficient expressed by (188). By applying (203) to (202) we obtain the closed-form expression of the conditional entropy per second  $H(X|\rho)$  of the transmitted signal  $X$  given the particle distribution  $\rho$  as follows:

$$H(X|\rho) = 2 \log_2(\pi Dd) + \frac{4d}{3 \ln 2} \sqrt{\frac{\pi W}{D}}. \quad (204)$$

The closed-form expression for the mutual information  $I(X; \rho)$  of the Fick's diffusion is finally computed by subtracting the conditional entropy per second  $H(X|\rho)$  of the transmitted signal  $X$  (204) given the particle distribution  $\rho$  from the entropy per second  $H(X)$  (198) of the transmitted signal  $X$ . This results in the following expression:

$$I(X; \rho) = 2WH(\hat{n}_T) - 2 \log_2(\pi Dd) - \frac{4d}{3 \ln 2} \sqrt{\frac{\pi W}{D}}, \quad (205)$$

where  $H(\hat{n}_T)$  is the entropy of a time sample of  $n_T(t)$ .

### 5.3.4 The Particle Location Displacement Mutual Information

The mutual information  $I(Y; \rho)$  in  $[bit/sec]$  of the particle location displacement is computed from the following expression:

$$I(Y; \rho) = H(\rho) - H(\rho|Y), \quad (206)$$

where  $H(\rho)$  is the entropy per second of the particle distribution  $\rho$  and  $H(\rho|Y)$  is the conditional entropy per second of the particle distribution  $\rho$  given the received signal  $Y$ .

The entropy  $H(\rho)$  per second of the particle distribution  $\rho$  is computed from (199) by substituting the expression for the entropy per second  $H(X)$  of the transmitted signal  $X$  from (198) and by applying (202) and (204) as the solution of the integral in (199). As a result, the entropy per second  $H(\rho)$  of the particle distribution  $\rho$  has the following expression:

$$H(\rho) = 2WH(\hat{n}_T) - \log_2 [(\pi Dd)^2] - \frac{4d}{3 \ln 2} \sqrt{\frac{\pi W}{D}}. \quad (207)$$

The conditional entropy per second  $H(\rho|Y)$  of the particle distribution  $\rho$  given the received signal  $Y$  is computed similarly to (198) in Section 5.3.3. Under the assumption that the realizations of the stochastic process  $\rho|Y$  are independent [66] for different time instants, and band limited within a bandwidth  $W$ , we express the entropy per second  $H(\rho|Y)$  as the entropy of  $H(\rho|\mathcal{Y})$  in  $[bit]$ , where  $\mathcal{Y}$  is the received signal per time sample, multiplied by the maximum time sample rate  $2W$  in  $[1/sec]$  given by the Shannon-Hartley theorem [51], expressed as

$$H(\rho|Y) = 2WH(\rho|\mathcal{Y}). \quad (208)$$

The received signal  $\mathcal{Y}$  per time sample is defined as

$$\mathcal{Y} = \sum_{i=1}^{1/(2W\tau_p)} y_i, \quad (209)$$

where  $1/(2W\tau_p)$  is the number of independent measures of the number of particles inside the receiver volume that can be performed within a time sample, for which we consider a quasi-constant particle distribution.  $W$  is the bandwidth of the transmitted signal  $X$ . We

assume independent measures when they are taken at time instants spaced by an interval  $\tau_p$ , as we considered in Section 4.4.2. The time interval  $\tau_p$  is equal to the squared linear dimension of the receiver volume  $R_{V_R}$  divided by the diffusion coefficient  $D$ , as derived in Section 4.4.2, expressed as

$$\tau_p = \frac{R_{V_R}^2}{D}. \quad (210)$$

The conditional entropy  $H(\rho|\mathcal{Y})$  of the particle distribution  $\rho$  given the received signal  $\mathcal{Y}$  per time sample is defined as

$$H(\rho|\mathcal{Y}) = \int H(\rho|\mathcal{Y} = y)p_{\mathcal{Y}}(y)dy = E_y [H(\rho|\mathcal{Y} = y)], \quad (211)$$

where  $H(\rho|\mathcal{Y} = y)$  is the entropy of the particle distribution  $\rho$  given a value  $y$  for the received signal per time sample  $\mathcal{Y}$ ,  $p_{\mathcal{Y}}(y)$  is the probability density of the received signal  $\mathcal{Y}$  per time sample and  $E_y[.]$  is the average value operator with respect to the probability density of the value  $y$ .

The entropy  $H(\rho|\mathcal{Y} = y)$  is based on the probability density  $p_{\rho|\mathcal{Y}}(r|y)$  of the possible values  $y$  of the particle distribution  $\rho$  at the receiver given a value  $y$  for the received signal per time sample  $\mathcal{Y}$  through the formula

$$H(\rho|\mathcal{Y} = y) = - \int p_{\rho|\mathcal{Y}}(r|y) \log_2 p_{\rho|\mathcal{Y}}(r|y) dr. \quad (212)$$

With the goal of having a closed-form expression for this probability density, we use the following assumptions, with reference to our previous work on the particle counting noise in Section 4.4:

- The actual number of particles  $y_i$  inside the receiver volume for every measurement is a random process whose average value is the average particle distribution at the receiver  $\bar{\rho}$  within a time sample multiplied by the size  $size(V_R)$  of the receiver volume, expressed as

$$E[y_i] = \bar{\rho} \text{ size}(V_R). \quad (213)$$

Since the particle distribution  $\rho$  is the output of the Fick's diffusion, whose input is the stationary stochastic process of the number of the emitted particles  $\hat{n}_T$  for every time sample, by applying the theory of the random processes through linear filters [66] we obtain

$$\bar{\rho} = \text{E} [\hat{n}_T] G_d(f)|_{f=0} = \frac{\text{E} [\hat{n}_T]}{\pi D d}, \quad (214)$$

where  $G_d(f)$  is from (203) and  $\text{E} [.]$  is the average operator.

- It is unlikely to have two particles occupying the same location in space at the same time instant. In other words, the probability of having a distance equal to zero between two particles is zero, expressed as

$$\text{Pr} [\|P_p(t) - P_q(t)\| = 0] = 0 \quad p \neq q, \quad p, q \in \mathcal{N}_T(t), \quad (215)$$

where  $\mathcal{N}_T(t)$  is given by (181),  $\|.\|$  is the Euclidian distance operator and  $p$  and  $q$  are two particles previously emitted by the transmitter, which are subject to the Brownian motion. This assumption is justified by the independence of the Brownian components in the movement of different particles in the space.

- An event concerning a particle which occupies a location in space is independent from any event of the same kind occurring at any other space location. This assumption is justified by the property of the Wiener process [27] underlying the particle Brownian motion of having independent realizations. This implies that the location of a particle is independent from the location of any other particle. As a consequence, the events concerning the location of particles in the space have the property of *memorylessness*.
- The occurrence rate of particle locations in the space is proportional to the particle distribution at the receiver location  $\rho$ .

Under these assumptions, the resulting single measurement  $y_i$ , which corresponds to the

number of particles inside the receiver volume, is a volumetric Poisson counting process [66], whose rate of occurrence corresponds to the average value of the particle distribution  $\bar{\rho}$  within a time sample, expressed as

$$y_i \sim \text{Poiss}(\bar{\rho}), \quad (216)$$

According to the theory of Bayesian inference [87], the estimator of the rate of occurrence  $\hat{\rho}$  of a Poisson counting process given a series of measurements of the output of the process, which corresponds to a value  $y$  of the received signal per time sample  $\mathcal{Y}$  defined in (209), follows a Gamma distribution with parameters  $y$  and  $1/(2W\tau_p)$ , expressed as

$$\hat{\rho} \sim \text{Gamma}(y, 1/(2W\tau_p)), \quad (217)$$

where  $W$  is the bandwidth of the transmitted signal  $X$  and  $\tau_p$  is the time interval in which we consider a quasi-constant particle distribution. The probability density of the estimator  $\hat{\rho}$  corresponds to the probability density  $p_{\rho|\mathcal{Y}}(r|y)$  of the possible values  $r$  of the particle distribution  $\rho$  at the receiver given a value  $y$  for the received signal in a time sample  $\mathcal{Y}$  [87], expressed as

$$p_{\rho|\mathcal{Y}}(r|y) = r^{y-1} \frac{e^{-rW\tau_p}}{(W\tau_p)^y \Gamma(y)}, \quad r \geq 0, y > 0, \quad (218)$$

where  $\Gamma(y)$  is the gamma function [88], defined as follows:

$$\Gamma(y) = (y - 1)! . \quad (219)$$

The entropy  $H(\rho|\mathcal{Y} = y)$  of the particle distribution  $\rho$  given a value  $y$  for the received signal per time sample  $\mathcal{Y}$  corresponds to the computation of the formula in (212) by using the expression of the probability density  $p_{\rho|\mathcal{Y}}(r|y)$  from (218), thus obtaining the following expression [87]:

$$H(\rho|\mathcal{Y} = y) = y + \ln(2W\tau_p) + \ln(\Gamma(y)) + (1 - y)\psi(y), \quad (220)$$

where  $\psi(y)$  is the digamma function.

For the final computation of the conditional entropy  $H(\rho|\mathcal{Y})$ , expressed in (211), we apply a formulation of the Jensen's inequality [66], which is based on the consideration that  $H(\rho|\mathcal{Y} = y)$  is a concave function of  $y$ , since its expression in (220) is a sum of concave or linear components:

- The first term  $y$  is linear.
- The second term  $\ln(\Gamma(y))$  is concave in  $y$  since the gamma function  $\Gamma(y)$  has the property of being log-concave [88].
- The third term  $(1 - y)\psi(y)$  is concave in  $y$  when the value of  $y$  is sufficiently high ( $y \gg 1$ ). This can be proven from the decomposition of the digamma function as follows [89]:

$$\psi(y) = \ln y - \frac{1}{2y} - \frac{1}{12y^2} + \frac{1}{120y^4} - \frac{1}{252y^6} + O\left(\frac{1}{y^8}\right), \quad (221)$$

By taking the limit of  $(1 - y)\psi(y)$  as  $y \rightarrow \infty$  we obtain

$$(1 - y)\psi(y) \rightarrow -y \ln y, \quad (222)$$

which is a concave function of  $y$ .

For the aforementioned considerations, the Jensen's inequality [66] applied to (211) states that the average value  $E_y [H(\rho|\mathcal{Y} = y)]$  of the entropy  $H(\rho|\mathcal{Y} = y)$  as function of the value  $y$  for the received signal per time sample  $\mathcal{Y}$  is less or equal than the entropy  $H(\rho|E[\mathcal{Y}])$  as function of the average value  $E[\mathcal{Y}]$  of the received signal per time sample  $\mathcal{Y}$ , as

$$H(\rho|\mathcal{Y}) = E_y [H(\rho|\mathcal{Y} = y)] \leq H(\rho|E[\mathcal{Y}]), \quad (223)$$

As a consequence, by substituting the left hand side of (223) for the computation (211) of the conditional entropy  $H(\rho|\mathcal{Y})$  of the particle distribution  $\rho$  given the received signal  $Y$  per time sample with the right hand side of (223), we provide a higher bound to the real value of  $H(\rho|\mathcal{Y})$ . This results in a higher bound to the conditional entropy  $H(\rho|Y)$  of the

particle distribution  $\rho$  given the received signal  $Y$  in (208) and, consequently, in a lower bound to the mutual information  $I(\rho; Y)$  of the transmitted signal and the particle distribution in (206). **Since the capacity is the maximum mutual information  $I(X; Y)$  between the transmitted signal  $X$  and the received signal  $Y$ , as expressed in (183), the substitution of  $I(\rho; Y)$  with its lower bound in the computation of the mutual information  $I(X; Y)$  expressed in (195) results in a lower bound to the capacity  $C$  of a molecular communication system.** We consider this in agreement with the purpose of this section of the Ph.D. thesis, since it allows expressing achievable performance of a molecular communication system with a closed-form mathematical expression, even if it is an underestimate of the theoretical capacity  $C$ .

The average value  $E[\mathcal{Y}]$  of the received signal per time sample  $\mathcal{Y}$  is given by (213) and (214). Since the distribution of particles  $\rho$  is a deterministic function of the transmitted signal  $X$ , whose average value per time sample is  $E[\hat{n}_T]$ , the value of  $E[\mathcal{Y}]$  becomes

$$E[\mathcal{Y}] = \sum_{i=1}^{1/(2W\tau_p)} E[y_i] = \frac{1}{2W\tau_p} \bar{\rho} size(V_R) = \frac{D}{2WR_{V_R}^2} \frac{E[\hat{n}_T] 4}{\pi D d} \frac{4}{3} \pi R_{V_R}^3 = \frac{2}{3} \frac{E[\hat{n}_T] R_{V_R}}{Wd}. \quad (224)$$

As stated previously, the expression for the conditional entropy  $H(\rho|\mathcal{Y})$  can be approximated with the right hand side of (223), whose expression is found by applying the average value  $E[\mathcal{Y}]$  of the received signal per time sample  $\mathcal{Y}$  to (220) as follows:

$$H(\rho|\mathcal{Y}) \cong H(\rho|E[\mathcal{Y}]) = E[\mathcal{Y}] + \ln(2W\tau_p) + \ln(\Gamma(E[\mathcal{Y}])) + (1 - E[\mathcal{Y}])\psi(E[\mathcal{Y}]). \quad (225)$$

The final approximated expression for the conditional entropy  $H(\rho|\mathcal{Y})$  is found by substituting the expression in (224). This becomes

$$H(\rho|\mathcal{Y}) \cong \frac{2}{3} \frac{E[\hat{n}_T] R_{V_R}}{Wd} + \ln(2W\tau_p) + \ln\left(\Gamma\left(\frac{2}{3} \frac{E[\hat{n}_T] R_{V_R}}{Wd}\right)\right) + \left(1 - \frac{2}{3} \frac{E[\hat{n}_T] R_{V_R}}{Wd}\right) \psi\left(\frac{2}{3} \frac{E[\hat{n}_T] R_{V_R}}{Wd}\right), \quad (226)$$

where  $W$  is the bandwidth of the transmitted signal  $X$ ,  $\tau_p$  is the time interval in which we consider a quasi-constant particle distribution,  $\psi(\cdot)$  is the digamma function,  $D$  is the diffusion coefficient,  $d$  is the distance between the transmitter and the receiver, and  $R_{V_R}$  is the radius of the spherical receiver volume  $V_R$ .

The closed-form expression for the mutual information  $I(Y; \rho)$  of the particle location displacement is finally computed by subtracting the conditional entropy  $H(\rho|\mathcal{Y})$  of the particle distribution  $\rho$  given the received signal  $\mathcal{Y}$  per time sample from (226) multiplied by two times the bandwidth  $W$  of the transmitted signal  $X$  from the entropy  $H(\rho)$  (207) of the particle distribution  $\rho$ . This results in the following expression:

$$I(Y; \rho) = 2WH(\hat{n}_T) - \log_2 [(\pi Dd)^2] - \frac{4d}{3 \ln 2} \sqrt{\frac{\pi W}{D}} - 2W \frac{2 \text{E}[\hat{n}_T] R_{V_R}}{3 Wd} - 2W \ln(2W\tau_p) + \\ - 2W \ln \left( \Gamma \left( \frac{2 \text{E}[\hat{n}_T] R_{V_R}}{3 Wd} \right) \right) - 2W \left( 1 - \frac{2 \text{E}[\hat{n}_T] R_{V_R}}{3 Wd} \right) \psi \left( \frac{2 \text{E}[\hat{n}_T] R_{V_R}}{3 Wd} \right). \quad (227)$$

### 5.3.5 The Capacity

The capacity of the diffusion-based MC system is computed from (183), by maximizing the mutual information  $I(X; Y)$ , expressed in Section 5.3.5.1 with respect to the probability density function  $f_X(x)$  of the transmitted signal  $X$ . It is common in information theory [69] to compute the maximum probability density function  $f_X(x)$  subject to a constraint on the average power of the transmitted signal  $X$  defined in (317). As explained in Section 5.3.5.2, the expression for this average power is here related to the thermodynamic energy spent for the emission of particles in the MC signal transmission. Finally, in Section 5.3.5.3 we obtain the closed-form expression of the lower bound to the capacity.

#### 5.3.5.1 The Mutual Information

The expression for the mutual information  $I(X; Y)$  of the diffusion-based MC system is obtained by applying the expression of the mutual information  $I(X; \rho)$  of the Fick's diffusion from (205), the mutual information  $I(Y; \rho)$  of the particle location displacement from (227) and the entropy  $H(\rho)$  of the particle distribution  $\rho$  from (207) to the formula in (195). We obtain the following expression:

$$I(X; Y) = 2WH(\hat{n}_T(t)) - \log_2 [(\pi Dd)^2] - \frac{4d}{3 \ln 2} \sqrt{\frac{\pi W}{D}} - 2W \frac{2 \text{E}[\hat{n}_T] R_{V_R}}{3 Wd} - 2W \ln(W\tau_p) + \\ - 2W \ln \left( \Gamma \left( \frac{2 \text{E}[\hat{n}_T] R_{V_R}}{3 Wd} \right) \right) - 2W \left( 1 - \frac{2 \text{E}[\hat{n}_T] R_{V_R}}{3 Wd} \right) \psi \left( \frac{2 \text{E}[\hat{n}_T] R_{V_R}}{3 Wd} \right), \quad (228)$$

where  $W$  is the bandwidth of the transmitted signal  $X$ ,  $\tau_p$  is the time interval in which we consider a quasi-constant particle distribution,  $\psi(\cdot)$  is the digamma function,  $D$  is the diffusion coefficient,  $d$  is the distance between the transmitter and the receiver and  $R_{V_R}$  is the radius of the spherical receiver volume  $V_R$ .

### 5.3.5.2 The Average Thermodynamic Power

Given the physical system considered at the beginning of Section 5.3, the average power necessary for signal transmission corresponds to the energy necessary to emit the average number  $E[\hat{n}_T]$  of particles per time sample, divided by the duration of a time sample. In thermodynamics, this energy is defined as enthalpy.

**Definition 5.3.1** *The enthalpy  $\mathcal{H}$  [90] is the energy necessary to emit  $N$  particles in the physical system and to heat these particles up to a temperature  $T$  when the system has the pressure  $P$  and the volume  $V$ . The considerations detailed at the beginning of Section 5.3 of having spherical particles with radius  $r \ll d$  independently and randomly moving in the space are in agreement with the approximation of the system as an ideal gas. According to the ideal gas theory [91], the enthalpy is expressed through the following formula:*

$$\mathcal{H} = PV + \frac{3}{2}K_bTN, \quad (229)$$

where  $P$  and  $V$  are the pressure and the volume and  $T$  is the absolute temperature of the physical system,  $K_b$  is the Boltzmann constant.

When associated to the transmitter of the molecular communication system, the enthalpy is the energy necessary for communication when  $N$  particles are emitted into the space.

In this section of the Ph.D. thesis, we define the **average thermodynamic power**  $\bar{P}_{\mathcal{H}}$  as the enthalpy variation  $\Delta\mathcal{H}$  in a time sample divided by the time sample duration  $1/2W$ . As a consequence, the average thermodynamic power  $\bar{P}_{\mathcal{H}}$  quantifies the energy necessary to emit  $E[\hat{n}_T]$  particles per time sample divided by the time sample duration  $1/2W$ , at a temperature  $T$ . This is given by the following expression:

$$\bar{P}_{\mathcal{H}} = \frac{\Delta\mathcal{H}}{1/2W} = \frac{3}{2}K_bTE[\hat{n}_T]2W, \quad (230)$$

where the enthalpy variation  $\Delta\mathcal{H}$  is computed from (230) by taking into account that no variations in the pressure  $P$  and the volume  $V$  occur in the physical system, and the absolute temperature  $T$  is considered a constant parameter with respect to the time  $t$ .

As a consequence, a constraint on the average thermodynamic power  $\bar{P}_{\mathcal{H}}$  spent by the transmitter corresponds to a constraint in the average number  $E[\hat{n}_T]$  of emitted particles according to the following expression:

$$E[\hat{n}_T] = \frac{\bar{P}_{\mathcal{H}}}{3WK_bT}. \quad (231)$$

### 5.3.5.3 The Lower-bound Expression of the Capacity

In the expression of  $I(X; Y)$  (228), only the term  $H(\hat{n}_T(t))$  depends on the probability density function  $f_{\hat{n}_T}(n)$ . Therefore, the lower-bound expression of the capacity  $C$  is achieved (183) for a probability density function  $f_{\hat{n}_T}(n)$  leading to the maximum entropy  $H(\hat{n}_T(t))$ .

The distribution  $f_{\hat{n}_T}(n)$  with the maximum possible entropy  $H(\hat{n}_T)$  in the number of emitted particles per time sample constrained on its average value  $E[\hat{n}_T]$ , as expressed in (231), is the Exponential distribution [77] whose rate corresponds to  $E[\hat{n}_T]$  as follows

$$f_{H(\hat{n}_T)}(n) = \frac{e^{-\frac{n}{E[\hat{n}_T]}}}{E[\hat{n}_T]}. \quad (232)$$

The entropy of the number  $H(\hat{n}_T)$  of emitted particles per time sample is therefore [77]

$$H(\hat{n}_T) = 1 + \log_2 E[\hat{n}_T]. \quad (233)$$

By applying (231) and (233) to the expression of the mutual information  $I(X; Y)$  from (228), we obtain the lower-bound expression of the capacity  $C$  of the diffusion-based MC system as follows:

$$\begin{aligned} C = & 2W \left( 1 + \log_2 \frac{\bar{P}_{\mathcal{H}}}{3WK_bT} \right) - 2 \log_2 (\pi D d) - \frac{4d}{3 \ln 2} \sqrt{\frac{\pi W}{D}} - 2W \frac{2\bar{P}_{\mathcal{H}}R_{V_R}}{9W^2 d K_b T} - 2W \ln(W\tau_p) + \\ & - 2W \ln \left( \Gamma \left( \frac{2\bar{P}_{\mathcal{H}}R_{V_R}}{9W^2 d K_b T} \right) \right) - 2W \left( 1 - \frac{2\bar{P}_{\mathcal{H}}R_{V_R}}{9W^2 d K_b T} \right) \psi \left( \frac{2\bar{P}_{\mathcal{H}}R_{V_R}}{9W^2 d K_b T} \right), \end{aligned} \quad (234)$$

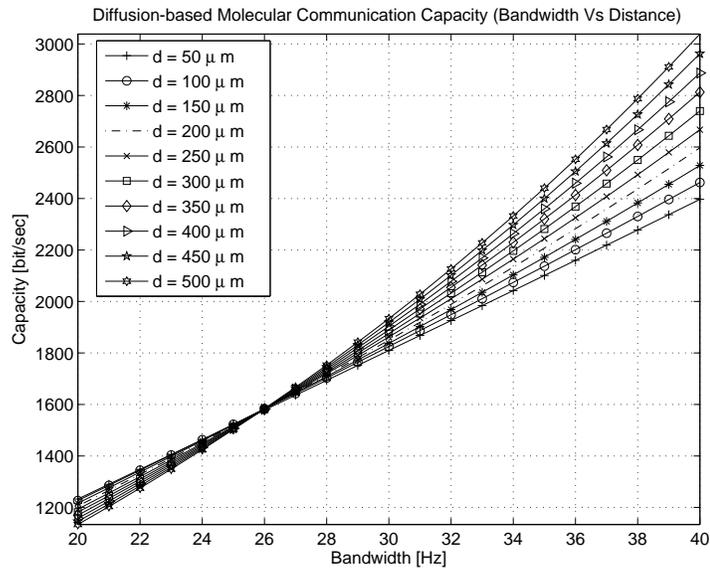
where  $\bar{P}_{\mathcal{H}}$  is the average thermodynamic power spent by the transmitter,  $K_b$  is the Boltzmann constant,  $T$  is the absolute temperature of the system,  $W$  is the bandwidth of the

transmitted signal  $X$ ,  $\tau_p$  is the time interval in which we consider a quasi-constant particle distribution,  $\psi(\cdot)$  is the digamma function,  $D$  is the diffusion coefficient,  $d$  is the distance between the transmitter and the receiver and  $R_{V_R}$  is the radius of the spherical receiver volume  $V_R$ .

### 5.3.6 Numerical Results

In this section, we provide a numerical evaluation of the closed-form expression of the lower bound to the capacity of a diffusion-based MC obtained in Section 5.3.5. All the results are computed for a common set of parameters, whose values are assigned as follows. The radius  $R_{V_R}$  of the receiver volume  $V_R$ , which we assume to be spherical, is set to 10 nm. The temperature  $T$  of the system is set to a standard room temperature of 25 °C and the diffusion coefficient  $D$  is set to  $10^{-9}$  [m<sup>2</sup>/sec], as explained in Chapter 3. The Boltzmann constant [81] is  $K_b = 1.380650424 \times 10^{-23}$  [Joule/K].

#### 5.3.6.1 Capacity Vs Bandwidth



**Figure 57.** Capacity in relation to the bandwidth and for different values of the transmitter-receiver distance  $d$ .

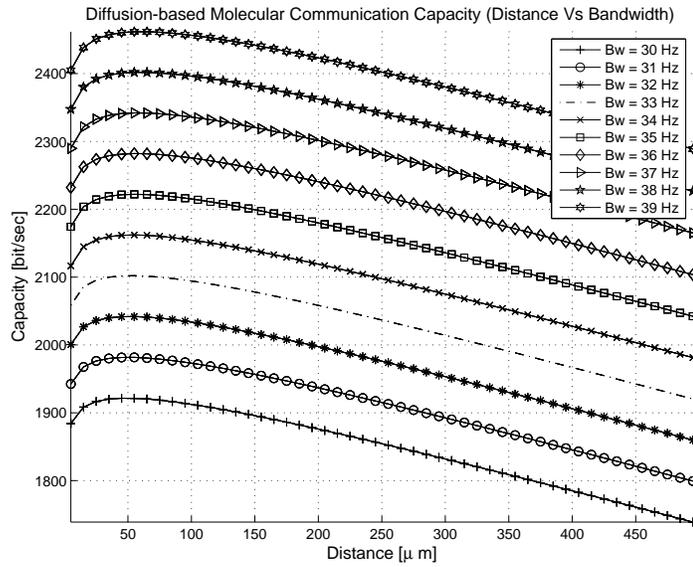
The values of the lower bound to the capacity of a diffusion-based MC system are shown in Figure 57 in relation to the bandwidth  $W$  and different values of the transmitter-receiver

distance  $d$ . We evaluate the capacity in [bit/sec] for a bandwidth  $W$  ranging from 20 Hz to 40 Hz and different lines refer to different distance values, from 50  $\mu\text{m}$  to 500  $\mu\text{m}$ . The choice of the values for the bandwidth can be justified from a biological viewpoint, since, according to biochemical studies [2], the neurons in our brain communicate through the exchange of molecules (and their diffusion between the synapses) at a frequency of around 20 Hz for the processing of general information and around 60 Hz for the processing of visual images. We restricted our range to a maximum of 40 Hz in order to visualize better the intersection of the curves around 26 Hz. The average transmitter power  $P_{\mathcal{H}}$  is set to  $10^{-12}$  W, equivalent to 1 pW. This value should not be compared to the transmitted power values used for electrical devices, since the average transmitted power is a thermodynamic quantity. Note also that this is only a reference value, since so far we do not know how much average thermodynamic power a transmitter nanomachine will be able to provide. According to the obtained results, the capacity of a molecular communication system with the chosen parameters can achieve a value close to 3 [Kbit/sec] at a distance of 500  $\mu\text{m}$  and for a bandwidth of 40 Hz. This is a theoretically achievable maximum value, which reveals the maximum potential of molecular communication. Further investigation on information coding schemes is required in order to provide achievable bit rates related to specific molecular communication implementations.

Figure 57 shows the trend of the MC capacity, which is monotonically increasing as the bandwidth increases from 20 Hz to 40 Hz for all the given values of the transmitter-receiver distance  $d$ . The capacity values range from 1.2 [Kbit/sec] to 2.4 [Kbit/sec] for a distance of 50  $\mu\text{m}$  and between a few [bits/sec] and 3 [Kbit/sec] for a distance of 500  $\mu\text{m}$ . For a bandwidth value within 20 and 26 Hz, the MC capacity values corresponding to the lowest transmitter-receiver distance  $d$  are higher than the values corresponding to other transmitter-receiver distances and, as this distance increases, the MC capacity values decrease monotonically. For a bandwidth value higher than 26 Hz, higher MC capacity values correspond to higher transmitter-receiver distances  $d$ . This behavior, which is apparently

counterintuitive, can be explained as a consequence of the interactions between the channel memory and the molecular noise contributions in the second and third term of the first line and in the second, third and fourth lines of (234), respectively. For low bandwidth values (lower than 26 Hz), the channel memory terms tend to outperform the molecular noise terms and the MC capacity values tend to be proportional to the transmitter-received distance (higher MC capacity when lower transmitter-receiver distance). For high bandwidth values (higher than 26 Hz), the molecular noise terms outperform the channel memory terms and the MC capacity values become inversely proportional to the transmitter-receiver distance (higher MC capacity when higher transmitter-receiver distance).

### 5.3.6.2 Capacity Vs Distance



**Figure 58.** Capacity in relation to the transmitter-receiver distance and for different values of the bandwidth  $W$ .

In Figure 58 we show the values of the lower bound to the capacity in relation to the distance  $d$  between the transmitter and the receiver locations and different values of the bandwidth  $W$ . We evaluate the capacity in  $[bit/sec]$  for a distance ranging from  $1 \mu m$  to  $500 \mu m$ . The different lines refer to different bandwidth  $W$  values, from 30 Hz to 39 Hz. We restricted these numerical results to this narrow bandwidth interval in order to

better visualize the differences in the MC capacity for the considered values of the distance between the transmitter and the receiver. The average transmitted power  $P_{\mathcal{H}}$  is set to 1 pW.

The curves in Figure 58 show a monotonically increasing trend of the capacity as function of the transmitter-receiver distance ranging from 1  $\mu\text{m}$  to 50  $\mu\text{m}$ , while they show a monotonically decreasing value for a distance ranging from 50  $\mu\text{m}$  to 500  $\mu\text{m}$ . The capacity values range from a value around 1.9 [Kbit/sec] and a 1.85 [bits/sec] for a bandwidth of 30 Hz and between 2.45 [Kbit/sec] and 2.3 [Kbit/sec] for a bandwidth of 39 Hz. The different behavior when the distance ranges from 1  $\mu\text{m}$  to 50  $\mu\text{m}$  with respect to when the distance ranges from 50  $\mu\text{m}$  to 500  $\mu\text{m}$  can be explained as a consequence of the interactions between the channel memory and the molecular noise contributions, similarly to Figure 57: as the distance increases from 1  $\mu\text{m}$  to 50  $\mu\text{m}$ , the contribution coming from the channel memory gets lower and the capacity values tend to increase, until reaching a distance of 50  $\mu\text{m}$ , where the contribution coming from the molecular noise becomes relevant and decreases the capacity values as the distance is further increased.

### 5.3.6.3 Capacity and Average Transmitted Power

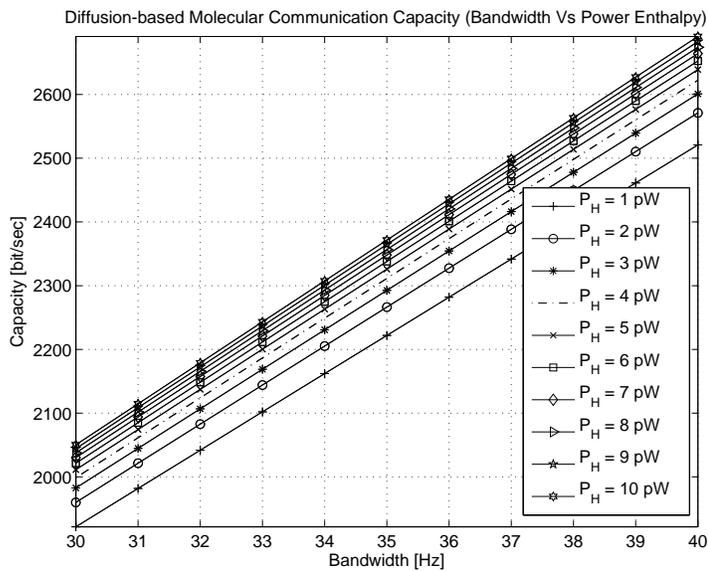


Figure 59. Capacity in relation to the bandwidth and for different values of the average transmitted power  $P_{\mathcal{H}}$ .

In Figure 59 we show the values of the lower bound to capacity as a function of the bandwidth  $W$  ranging from 30 Hz to 40 Hz and the average transmitted power  $P_{\mathcal{H}}$ . Similarly to Figure 58, we restricted these numerical results to this narrow bandwidth interval in order to better visualize the differences in the MC capacity for the values of the average transmitter power. Different lines refer to different average transmitted power  $P_{\mathcal{H}}$  values, from 1 pW to 10 pW. The transmitter-receiver distance  $d$  is here set to  $50 \mu\text{m}$ .

Figure 59 shows for all the curves a monotonic increasing trend as the bandwidth increases from 30 Hz to 40 Hz. The capacity values range from a value between 1.92 [Kbit/sec] and around 2.5 [kbits/sec] for an average transmitted power of 1 pW and between 2.05 [Kbit/sec] and 3.7 [Kbit/sec] for an average transmitted power of 10 pW. The MC capacity values are higher for higher values of the average transmitted power. Even if the average transmitted power is applied with constant increments of 1 pW from a value of 1 pW to 10 pW, the increment in the values of the capacity is not constant and it is higher for lower values of the average transmitted power. This behavior can be explained through the dependency of the molecular noise terms in the second, third and fourth lines of (234) with respect to the average transmitted power. As the average transmitted power increases, we have an increase in the first positive term in the first line of (234), but, at the same time, we have an increase in the aforementioned molecular noise terms.

### 5.3.7 Conclusion

In this section of the Ph.D. thesis, unlike previous contributions from the literature, an expression of the lower bound to the true diffusion-based MC capacity is provided by taking account the two main effects of the molecule diffusion channel, namely, the memory and the molecular noise. The capacity analysis in this section is also independent from any specific coding scheme by assuming that the transmitter can send in general any continuous time signal which complies to a constraint on the average transmitted power.

The closed-form expression is obtained here by combining two separate contributions,

namely, the Fick's diffusion and the particle location displacement, which separately capture the effects of the channel memory and the molecular noise, respectively. The obtained lower-bound expression of the capacity is a function of the medium diffusion coefficient, the system temperature, the distance between the transmitter and the receiver, and the bandwidth of the transmitted signal. The MC capacity is also expressed as a function of the average transmitted power, which corresponds to the thermodynamic power spent at the transmitter for molecule emission. Numerical results show interesting properties of the relationship between diffusion-based MC capacity and parameters such as the distance, the bandwidth and the average thermodynamic power.

According to the provided results, capacity values of a few [Kbit/sec] can be reached within a distance of tenth of  $\mu\text{m}$  between the transmitter and the receiver and for an average transmitted power around 1 pW (Note that this power value should not be compared to the transmitted power values used for electrical devices, since the transmitted power in a MC system is a thermodynamic quantity).

# CHAPTER 6

## INTERFERENCE ANALYSIS IN DIFFUSION-BASED MOLECULAR COMMUNICATION

### 6.1 Motivation and Related Work

The analysis of the interference produced at the reception of a transmitted signal, either by distortions of the signal itself or by other concurrent signals coming from different transmitters, is fundamental to design interference mitigation techniques and increase the performance of a communication systems. The focus of this chapter of the Ph.D. thesis is on the analysis of interference in diffusion-based Molecular Communication (MC), when multiple transmitters access the fluid medium simultaneously and emit molecular signals directed to a single receiver.

Several different methods have been used in the engineering literature to measure the impact of the interference. In this chapter of the thesis, three main methods are used, namely, the InterSymbol Interference (ISI), the Co-Channel Interference (CCI) and the statistical-physical modeling. The ISI is here defined as the overlap between two consecutively received signals in molecule concentration, which were transmitted from a single molecular transmitter, while the CCI is considered here as the overlap between a received molecule concentration signal, which was transmitted by a single transmitter, and all the received molecule concentration signals transmitted by the other concurrent transmitters. Both ISI and CCI depend greatly on the number and locations of the transmitters, how information is encoded in the signal modulation, and how the signals propagate through the channel from the moment they are transmitted until they combine at the receiver side. In most of the classical communication channels, this propagation is expressed through the so-called wave equation, while in diffusion-based MC it is expressed through the fundamentally different diffusion equation [26, 27]. While the ISI and CCI parameters provide a more practical evaluation of the interference in specific situations, the latter method to

measure the impact of the interference, the statistical-physical modeling, provides a more general probabilistic description of the received signal when multiple transmitters access the medium simultaneously, which is independent from the transmitter number, specific transmitter locations or coding schemes.

Previous literature has addressed the problem of diffusion-based MC interference. In [92] the effects of the ISI and CCI are analyzed in reference to two specific modulation techniques proposed by the same authors. In [45] the ISI is characterized in a unicast MC system with binary amplitude modulation. In [47], interference is studied for another specific modulation technique, based on the transmission order of different types of molecules.

In the first part of this chapter, the ISI and the CCI are jointly analyzed for a diffusion-based MC system having a limited number of transmitters in predetermined locations. Moreover, the information transmitted in this system is encoded through the modulation of Gaussian pulses in the molecule emission rate. An in-depth analysis of the propagation of signals through a diffusion-based channel is performed by studying two main parameters, namely, the attenuation and the dispersion. For this, the diffusion equation is interpreted in terms of diffusion wave propagation, which allows to apply the wave theory to the realm of the diffusion-based MC, and to find mathematical expressions for the attenuation and the dispersion in a diffusion-based channel. From these, simple closed-form formulas for both the ISI and the CCI are derived. Two different modulation schemes, namely, the baseband modulation and the diffusion wave modulation, are considered for the release of molecules in the diffusion-based MC, and are compared in terms of interference. The obtained analytical results for both ISI and CCI are compared and validated by simulation results. This ultimately allows to assess the validity of the simple closed-form formulas for the evaluation of the interference in a diffusion-based MC system.

The objective of the second part of this chapter is to provide a statistical-physical modeling of the interference in diffusion-based MC. Our method to characterize interference differentiates from the previous literature on diffusion-based MC, since the developed

statistical-physical model is independent from the transmitter number, specific transmitter locations or coding schemes. In particular, this model stems from the same general assumptions used in interference study for radio communication networks [93], namely, a spatial Poisson distribution of interfering transmitters having independent and identically distributed (i.i.d.) emissions. Moreover, the specific probability distribution used for the molecule emissions is in agreement with a chemical description of the transmitters in terms of Langevin equation [94], which models the randomness in the chemical reactions involved in the production of the molecules.

The statistical-physical modeling detailed the second part of this chapter is based on the property of the received molecular signal of being a stationary Gaussian Process (GP), which results from the molecule emission distribution and the diffusion-based molecule propagation. As a consequence, the statistical-physical modeling is operated on the received Power Spectral Density (PSD), for which it is possible to obtain an analytical expression of the log-characteristic function. The derivation of this log-characteristic function is inspired by the mathematical framework in [95], where the authors derive the log-characteristic function of the received signal in radio communication interference. The expression of the received PSD log-characteristic function ultimately leads to the estimation of the received PSD probability distribution. The received PSD probability distribution provides a complete description of the GP of the received molecular signal, which corresponds to the interference in diffusion-based molecular nanonetworks. By using the derived statistical-physical interference model, we also provide numerical results in terms of received PSD probability distribution and probability of interference for selected values of the physical parameters of the molecular nanonetwork, such as the diffusion coefficient, the transmitter density and the average power of molecule emissions, and we compare them with the outcomes of a simulation environment.

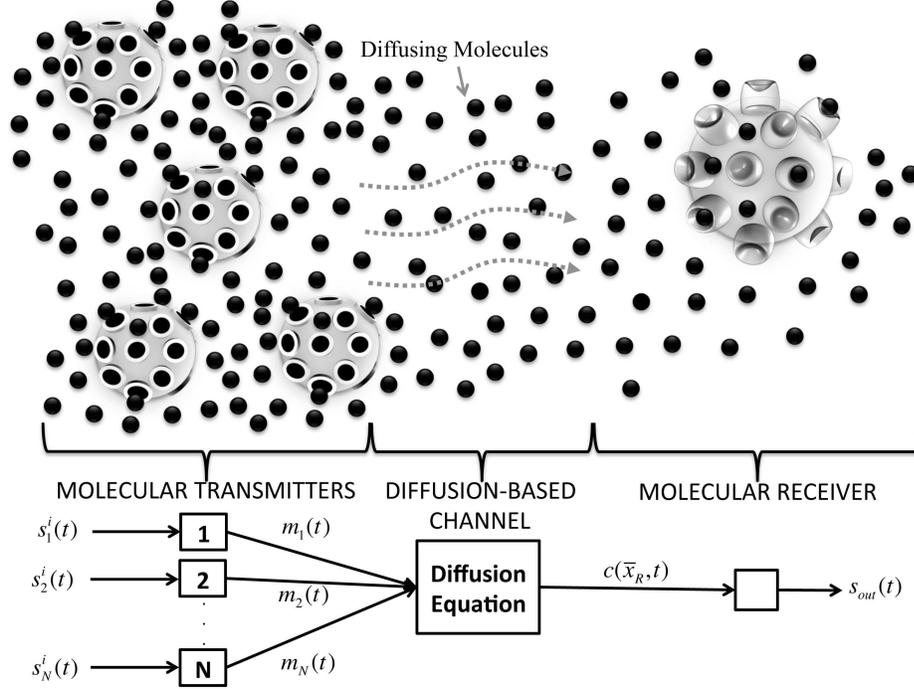


Figure 60. Block scheme of the diffusion-based MC system considered in this section of the Ph.D. thesis.

## 6.2 Intersymbol Interference and Co-Channel Interference

The MC system model considered in this section of the Ph.D. thesis, and shown in Figure 64, includes  $N$  **MOLECULAR TRANSMITTERS**. Each transmitter, denoted by  $n$  and located at  $\bar{x}_n$ , is responsible for the modulation of the number of molecules  $m_n(t)$  emitted into the space as a function of the time  $t$  according to input information signals, denoted as  $s_n^i(t)$ , where  $i = 1, 2, \dots$  is a sequential index. We assume the emitted molecules are identical and undistinguishable between each other. Two different modulation schemes that can be adopted by the molecular transmitters are studied and compared in this section of the Ph.D. thesis from the point of view of interference, namely, the *baseband modulation* and the *diffusion wave modulation*. For both modulation schemes, the transmitters produce a number of molecules  $m_n(t)$  emitted at location  $\bar{x}_n$  and time  $t$  corresponding to the amplitude modulation of an oscillation with angular frequency  $\omega_0$ , expressed as

$$m_n(t) = \sum_{i=0}^{\infty} s_n^i(t) e^{j\omega_0 t}, \quad (235)$$

where  $\omega_0 = 0$  in the baseband scheme and  $\omega_0 > 0$  in the diffusion wave scheme.

The MC system model includes a **DIFFUSION-BASED CHANNEL** which is based on the free diffusion of molecules between the transmitter and the receiver. Each molecular transmitter  $n$  emits a number of molecules  $m_n(t)$  in this space at location  $\bar{x}_n$  and time  $t$ . For this, the total number of emitted molecules  $m(\bar{x}, t)$ , which is the input of the molecular channel, is expressed as

$$m(\bar{x}, t) = \sum_{n=1}^N m_n(t) \delta(\bar{x} - \bar{x}_n) , \quad (236)$$

where  $\delta(\bar{x} - \bar{x}_n)$  is a Dirac delta defined in the three dimensional space and centered at the corresponding transmitter location  $\bar{x}_n$ . Once emitted, every molecule moves independently from the others and according to its Brownian motion in a fluidic medium. The output of the molecular channel is the molecule concentration  $c(\bar{x}, t)$  as function of the space location  $\bar{x}$  and the time  $t$ , whose relation with the input  $m(\bar{x}, t)$  is expressed by the **diffusion equation** [26, 27], as

$$\frac{\partial c(\bar{x}, t)}{\partial t} = D \nabla^2 c(\bar{x}, t) + m(\bar{x}, t) , \quad (237)$$

where  $D$  is the diffusion coefficient and it is considered a constant parameter within the scope of this section of the Ph.D. thesis.

The *linearity* of (324) gives the following results:

- Given a modulated number of emitted molecules  $m_n(t)$  from a single transmitter  $n$ , the output molecule concentration  $c(\bar{x}, t)$  at any location  $\bar{x}$  and time  $t$  is computed through the convolution integral with the Green's function [86]  $g(\bar{x}, t)$  (*linear channel*), expressed as follows:

$$c(\bar{x}, t) = m(\bar{x}, t) * g(\bar{x}, t) = \int_0^\infty m_n(t) g(\bar{x}_n - \bar{x}, t' - t) dt' , \quad (238)$$

where  $(. * .)$  denotes the convolution integral between the two arguments. The Green's function [86] is the solution of the diffusion equation (324) when the input  $m(\bar{x}, t)$  is a Dirac delta and it is expressed as follows:

$$g(\bar{x}, t) = \frac{1}{\sqrt{(4\pi Dt)^3}} e^{-\frac{|\bar{x}|^2}{4Dt}} . \quad (239)$$

- Given the modulated number of molecules  $m_n(t)$  emitted simultaneously from multiple transmitters, where  $n = 1, \dots, N$ , the output molecule concentration is the sum of the outputs of the diffusion-based channel applied independently to each single molecule concentration rate input (*additive channel*), expressed as

$$c(\bar{x}, t) = \sum_{n=1}^N (m_n(t)\delta(\bar{x} - \bar{x}_n) * g(\bar{x}, t)) . \quad (240)$$

The MC system model includes a single **MOLECULAR RECEIVER**, whose task is to read the incoming molecular concentration  $c(\bar{x}_R, t)$  at its location  $\bar{x}_R$  and to demodulate the output information signal  $s_{out}(t)$ . For this, the molecular receiver produces an output information signal  $s_{out}(t)$  equal to the real part of the molecule concentration signal  $c(\bar{x}_R, t)$  at the receiver location  $\bar{x}_R$ , multiplied by an oscillation with angular frequency  $-\omega_0$ , expressed as

$$s_{out}(t) = \Re \left\{ c(\bar{x}_R, t) e^{-j\omega_0 t} \right\} , \quad (241)$$

where  $\omega_0 = 0$  or  $\omega_0 > 0$  in case the transmitter adopted the baseband or the diffusion wave modulation scheme, respectively.  $\Re\{\cdot\}$  denotes the operator which extracts the real part from the complex operand.

### 6.2.1 Interference Formulas

The ISI is quantified as the time integral of the product of two output information signals which derive from two input information signals sent from a transmitter  $n$ , expressed as follows:

$$ISI = \int_{-\infty}^{\infty} s_{n,out}^i(t) s_{n,out}^{i+1}(t) dt , \quad (242)$$

where  $s_{n,out}^i(t)$  is the output information signal of the MC system when the input information signal  $s_n^i(t)$  is sent by the transmitter  $n$ .

The CCI is quantified as the time integral of the product of an output information signal which is sent in a modulated number of emitted molecules by a transmitter  $n$  with all the other received output information signals which are sent as modulated number of molecules

by all the other  $N - 1$  transmitters, expressed as follows:

$$CCI = \int_{-\infty}^{\infty} s_{n,\text{out}}^i(t) \sum_{k=1}^{N, k \neq n} \sum_{l=0}^{\infty} s_{k,\text{out}}^l(t) dt . \quad (243)$$

In case of *baseband modulation* scheme, the output information signal  $s_{n,\text{out}}^i(t)$ , which derives from the input information signal  $s_n^i(t)$  sent by the transmitter  $n$ , has the following expression:

$$s_{n,\text{out}}^i(t) = s_n^i(t) * g(\bar{x}, t) , \quad (244)$$

where  $(. * .)$  denotes the convolution integral between the two arguments and  $g(\bar{x}, t)$  has the expression from (239). In case of *diffusion wave modulation* scheme the same output information signal  $s_{n,\text{out}}^i(t)$  has the following expression:

$$s_{n,\text{out}}^i(t) = \Re \left\{ \left[ \left( s_n^i(t) e^{j\omega_0 t} \right) * g(\bar{x}, t) \right] e^{-j\omega_0 t} \right\} . \quad (245)$$

In order to evaluate the ISI through (242) and the CCI through (243), it is necessary to analyze how the shape of an information signal changes from its transmission as  $s_n^i(t)$  until its reception as  $s_{n,\text{out}}^i(t)$ . For this, we decompose an input information signal into its frequency components  $S_n^i(\omega)$  by applying the Fourier transform [51], expressed as

$$s_n^i(t) = \int_0^{\infty} S_n^i(\omega) e^{j\omega t} d\omega . \quad (246)$$

Each frequency component  $S_n^i(\omega) e^{j\omega t}$ , as it propagates in the diffusion-based channel defined by (324), is in general attenuated and it has a finite propagation velocity. As will be proved in the following section, this attenuation and velocity are functions of the angular frequency  $\omega$  of the frequency component  $S_n^i(\omega)$  itself. As a consequence, the output information signal  $s_{n,\text{out}}^i(t)$  will be composed by the same frequency components as the transmitted input information signal, each one attenuated by a different value and propagated with a different velocity. These two effects, identified as the **attenuation** and the **dispersion** of a signal, are at the basis of the changes in the information signal shape as it propagates through the diffusion-based channel. For this, in the following section we analyze these two parameters by using the wave theory [96].

### 6.2.2 Attenuation and Dispersion of Diffusion-Waves

This section deals with the analysis of the attenuation and the dispersion which affect any modulated total number of emitted molecules  $m(\bar{x}, t)$  in the diffusion-based channel defined at the beginning of Section 6.2 as it propagates from the transmitter to the receiver. For this, as suggested in [96, 86], we apply the wave theory to the diffusion equation from (324).

According to the wave theory, given an oscillatory input  $q(t)$  with angular frequency  $\omega$  of the following type:

$$q(\bar{x}, t) = Q(\bar{x}, \omega)e^{j\omega t} , \quad (247)$$

the propagation of a wave defined by the following expression:

$$u(\bar{x}, t) = U(\bar{x}, \omega)e^{j\omega t} , \quad (248)$$

stems from a differential equation that can be defined in the space  $\bar{x}$  and angular frequency  $\omega$  as follows:

$$\nabla^2 U(\bar{x}, \omega) - k^2(\omega)U(\bar{x}, \omega) = Q(\bar{x}, t) , \quad (249)$$

where  $Q(\bar{x}, t)$  and  $U(\bar{x}, \omega)$  are the input and the output respectively, as function of the space  $\bar{x}$  and the input angular frequency  $\omega$ .  $k(\omega)$  is the wavenumber, which is in general a function of  $\omega$ . We have the following definitions based on  $k(\omega)$ :

- The attenuation of a wave  $\alpha(\omega)$  is the imaginary part of the wavenumber  $k(\omega)$ , expressed as

$$\alpha(\omega) = \Im\{k(\omega)\} , \quad (250)$$

where  $\Im\{\cdot\}$  denotes the operator which extracts the imaginary part from the complex operand.

- The phase velocity  $v_p$  is equal to the angular frequency  $\omega$  divided by the real part of the wavenumber  $k(\omega)$  and it is defined as the propagation velocity of a point of constant phase (wavefront velocity), expressed as

$$v_p = \frac{\omega}{\Re\{k(\omega)\}} . \quad (251)$$

- The group velocity  $v_g$  is the time first derivative of the angular frequency  $\omega$  with respect to the real part of the wavenumber  $k(\omega)$  and it is defined as the propagation velocity of a group of waves having a narrow frequency range around  $\omega$  (wave-packet velocity), expressed as

$$v_g = \frac{\partial \omega}{\partial \Re \{k(\omega)\}} . \quad (252)$$

The wave propagation expressed through (249) is subject to dispersion if the expressions of the phase velocity (251) and the group velocity (252) are different. The resulting propagating wave from (248) can be written as function of the oscillatory input  $q(\bar{x}, t)$ , the attenuation  $\alpha(\omega)$  and the phase velocity  $v_p$  as follows:

$$u(\bar{x}, t) = q(\bar{x}, t) e^{-\alpha(\omega)|\bar{x}|} e^{j\frac{\omega}{v_p}|\bar{x}|} . \quad (253)$$

By taking the Fourier transform [51] of the diffusion equation from (324) and by rearranging the terms we obtain an expression of the same type as (249) defined in the space  $\bar{x}$  and angular frequency  $\omega$ , expressed as

$$\nabla^2 C(\bar{x}, \omega) - \frac{j\omega}{D} C(\bar{x}, \omega) = M(\bar{x}, \omega) , \quad (254)$$

where  $M(\bar{x}, \omega)$  and  $C(\bar{x}, \omega)$  are the Fourier transforms [51] of the modulated total number of emitted molecules  $m(\bar{x}, t)$  and the output molecule concentration signal  $c(\bar{x}, t)$ , respectively. The similarity with the wave equation in (249) suggests an interpretation of the diffusion equation in terms of waves, thus identifying the so-called diffusion waves [96]. Although the diffusion waves have different properties [86] if compared to the waves generated by the wave equation, also for the diffusion waves we can identify a wavenumber  $k(\omega)$ , this time equal to

$$k(\omega) = \sqrt{\frac{j\omega}{D}} = (1 + j) \sqrt{\frac{\omega}{2D}} . \quad (255)$$

As a consequence, the attenuation of a diffusion wave  $\alpha(\omega)$  is given by (250)

$$\alpha(\omega) = \sqrt{\frac{\omega}{2D}} . \quad (256)$$

The phase velocity  $v_p$  is given by applying (255) to (251)

$$v_p = \sqrt{2D\omega} , \quad (257)$$

and the group velocity  $v_g$  is computed through (255) and (252), expressed as

$$v_g = 2 \sqrt{2D\omega} . \quad (258)$$

Since the phase velocity in (257) and the group velocity in (258) are different, the wave propagation in the diffusion-based channel is affected by dispersion. This is a consequence of the frequency dependency of the phase velocity and the group velocity of the diffusion waves.

The resulting propagating diffusion wave can be written as function of an oscillatory total number of emitted molecules  $M(\bar{x}, \omega)e^{j\omega t}$ , the attenuation  $\alpha(\omega)$  and the phase velocity  $v_p$  as follows:

$$c(\bar{x}, t) = M(\bar{x}, \omega)e^{j\omega t} e^{-\sqrt{\frac{\omega}{2D}}|\bar{x}|} e^{j\sqrt{\frac{\omega}{2D}}|\bar{x}|} . \quad (259)$$

### 6.2.3 Interference Analysis

In this section, the  $i$ -th input information signal  $s_n^i(t)$  for the transmitter  $n$  is modeled as a Gaussian-shaped pulse with standard deviation  $\sigma$ , which is a user-defined parameter, and it is expressed as follows:

$$s_n^i(t) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(t-t_n-i\Delta t)^2}{2\sigma^2}} , \quad (260)$$

where  $t_n + i\Delta t$  is the time instant at which the receiver  $n$  transmits the maximum of the  $i$ -th pulse. Equation (260) allows to simplify the following interference analysis and to find closed-form expressions for the interference. Although these expressions depend on (260), this does not prevent from considering the general conclusions of this section of the Ph.D. thesis valid for any other input signal shape.

We describe the changes in the shape of the pulse sent by the transmitter  $n$  from its transmission as  $s_n^i(t)$  until its reception as  $s_{n,\text{out}}^i(t)$  by using two parameters, namely, the

amplitude  $A_n$  at the peak maximum and the broadening factor  $B_n$ , expressed as

$$s_{n,\text{out}}^i(t) = A_n s_n^i \left( \frac{t - t_d}{B_n} \right) = A_n \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(t-t_n-i\Delta t-t_d)^2}{2(B_n\sigma)^2}}, \quad (261)$$

where  $t_d$  is the pulse propagation delay and its value is not relevant for the following interference analysis since we account only for the time of pulse reception. The relation in (261) is a first approximation of the changes in the pulse shape and it allows a simplification of the expressions of the ISI and the CCI in (242) and (243), respectively. In light of (261) the ISI becomes

$$ISI \simeq 2A_n^2 \operatorname{erfc} \left( \frac{\Delta t/2}{\sqrt{2}B_n\sigma} \right) \left[ 1 - \operatorname{erfc} \left( \frac{\Delta t/2}{\sqrt{2}B_n\sigma} \right) \right]. \quad (262)$$

Similarly, the CCI becomes

$$CCI \simeq \sum_{k=1}^{N, k \neq n} 2A_n A_k \operatorname{erfc} \left( \frac{|t_n - t_k|/2}{\sqrt{2}B_n\sigma} \right) \left[ 1 - \operatorname{erfc} \left( \frac{|t_n - t_k|/2}{\sqrt{2}B_k\sigma} \right) \right], \quad (263)$$

where  $t_n$  and  $t_k$  are the time instants of transmission of the pulses and  $B_n$  and  $B_k$  are the broadening factors for the pulses transmitted by the transmitter  $n$  and the transmitter  $k$ , respectively.  $\operatorname{erfc}(x/\sqrt{2}B_n\sigma)$  denotes the complementary error function, which corresponds to the integral of the Gaussian pulse dilated by a factor  $B_n$  between  $x$  and  $\infty$ .

By stemming from the formulas discussed in Section 6.2.1 concerning the attenuation and the dispersion of the diffusion waves, we can derive closed-form formulas for the amplitude  $A_n$  at the peak maximum and the broadening factor  $B_n$ :

- The pulse *amplitude  $A_n$  at the peak maximum* after propagation from the transmitter  $n$  located at  $\bar{x}_n$  to the receiver located at  $\bar{x}_R$  is given by the double of the integral of the attenuation contribution  $e^{-\sqrt{\frac{\omega}{2D}}|\bar{x}_R - \bar{x}_n|}$  of each frequency component  $M(\bar{x}_n, \omega)$  of the transmitted signal, expressed as

$$A_n = 2 \int_0^\infty M(\bar{x}_n, \omega) e^{-\sqrt{\frac{\omega}{2D}}|\bar{x}_R - \bar{x}_n|} d\omega, \quad (264)$$

where  $M(\bar{x}_n, \omega)$  is the Fourier transform [51] of the total number of emitted molecules  $m(\bar{x}_n, t)$  when only a single pulse is transmitted from (241), having  $\omega_0 = 0$  in the case of baseband scheme and  $\omega_0 > 0$  in the case of diffusion wave modulation.

- The pulse *broadening* factor  $B_n$  is computed as the squared root of the sum of 1 with the squared integral of the delay contribution  $\frac{\partial}{\partial\omega}\left(\frac{1}{v_g}\right)|\bar{x}_R - \bar{x}_n|$  of each frequency component  $M(\bar{x}_n, \omega)$  of the total number of emitted molecules, expressed as

$$B_n = \sqrt{1 + \left( \int_0^\infty \frac{\partial}{\partial\omega} \left( \frac{1}{v_g} \right) |\bar{x}_R - \bar{x}_n| M(\bar{x}_n, \omega) d\omega \right)^2}, \quad (265)$$

where  $v_g$  is the group velocity expressed in (258). The first derivative of the inverse of the group velocity  $1/v_g$  with respect to the angular frequency  $\omega$  has the following expression:

$$\frac{\partial}{\partial\omega} \left( \frac{1}{v_g} \right) = \sqrt{\frac{1}{2D\omega^3}}. \quad (266)$$

In order to compare the ISI and CCI results for the two modulation schemes, we simplify further (264) and (265) by approximation. For the **baseband modulation** scheme, the amplitude  $A_n^{base}$  at the peak maximum becomes

$$A_n^{base} = \frac{2}{\sigma^2} e^{-\sqrt{\frac{\omega_c}{2D}} |\bar{x}_R - \bar{x}_n|}, \quad (267)$$

where  $\omega_c$  is the cut-off frequency of the Gaussian pulse (260). The cut-off frequency is the angular frequency of the pulse spectrum component whose amplitude value is half the amplitude of the maximum. The pulse broadening factor  $B_n^{base}$  for the baseband modulation scheme can be approximated with

$$B_n^{base} = \sqrt{1 + \left( \frac{1}{\sigma^2} \sqrt{\frac{1}{2D\omega_c^3}} |\bar{x}_R - \bar{x}_n| \right)^2}. \quad (268)$$

In case of **diffusion wave modulation** scheme, the amplitude  $A_n^{wave}$  at the peak maximum becomes

$$A_n^{wave} = \frac{2}{\sigma^2} e^{-\sqrt{\frac{\omega_0}{2D}} |\bar{x}_R - \bar{x}_n|}, \quad (269)$$

where  $\omega_0$  is the frequency of the modulating oscillation, as expressed in (235). The pulse broadening factor  $B_n^{wave}$  for the diffusion wave modulation scheme can be approximated with

$$B_n^{wave} = \sqrt{1 + \left( \frac{1}{\sigma^2} \sqrt{\frac{1}{2D\omega_0^3}} |\bar{x}_R - \bar{x}_n| \right)^2}, \quad (270)$$

where for both (269) and (270) we assumed to have a frequency  $\omega_0$  much higher than the the cut-off frequency  $\omega_c$  of the Gaussian pulse (260).

The amplitude at the peak maximum for the baseband modulation scheme  $A_n^{base}$  and for the diffusion wave modulation scheme  $A_n^{wave}$ , if compared, guide to the following result:

$$A_n^{base} > A_n^{wave} \quad \forall \omega_0 > \omega_c . \quad (271)$$

When comparing the pulse broadening in case of baseband modulation scheme  $B_n^{base}$  and in case of diffusion wave modulation scheme  $B_n^{wave}$ , we can conclude the following result:

$$B_n^{base} > B_n^{wave} \quad \forall \omega_0 > \omega_c . \quad (272)$$

As a conclusion, for a diffusion-based channel model as defined by (324), the intersymbol interference  $ISI_{\bar{x}_R}^{wave}$  in the case of diffusion wave modulation scheme is lower with respect to the intersymbol interference  $ISI_{\bar{x}_R}^{base}$  in the case of baseband modulation scheme, expressed as

$$ISI_{wave} < ISI_{base} \quad \forall \omega_0 > \omega_c . \quad (273)$$

In addition, we deduce that the higher is the wave modulation frequency  $\omega_0$ , the lower is the value of the intersymbol interference  $ISI_{wave}$ , expressed as

$$ISI_{wave}|_{\omega_0=\omega_1} < ISI_{wave}|_{\omega_0=\omega_2} \quad \forall \omega_1 > \omega_2 . \quad (274)$$

Similarly, we compare the co-channel interference for the two modulation schemes applied to systems having the same values for the locations  $\bar{x}_n$  and the time instants  $t_n$ , which correspond to the maximum of the transmitted Gaussian pulses for all the  $N$  transmitters. We deduce also for the CCI the following result:

$$CCI_{wave} < CCI_{base} \quad \forall \omega_0 > \omega_c , \quad (275)$$

and

$$CCI_{wave}|_{\omega_0=\omega_1} < CCI_{wave}|_{\omega_0=\omega_2} \quad \forall \omega_1 > \omega_2 . \quad (276)$$

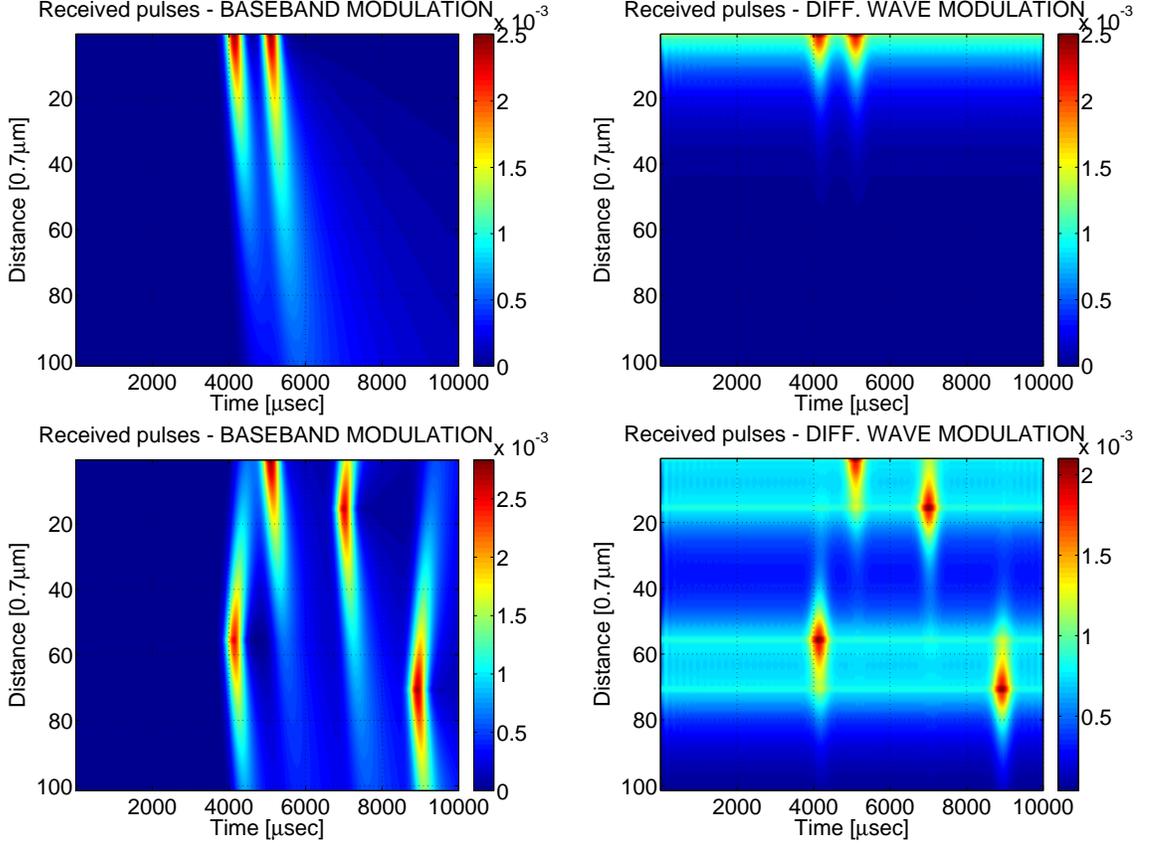
#### 6.2.4 Numerical Results

In this section we simulate the system detailed at the beginning of Section 6.2 and we compare the results in terms of ISI and CCI with the simple formulas resulting from the interference analysis of Section 6.2.3, which stems from the diffusion-wave attenuation and dispersion studied in Section 6.2.2. The goal of this comparison is to prove that the simple formulas for the ISI (262) and for the CCI (263) constitute a valid approximation for the evaluation of the interference in a diffusion-based MC system.

The simulations are based on (235), (238) and (241),  $\omega_0 = 0$  in the baseband scheme and  $\omega_0 > 0$  in the diffusion wave scheme. The diffusion coefficient  $D$  in (239) is set to  $\sim 10^{-6}cm^2sec^{-1}$  of calcium molecules diffusing in a biological environment (cellular cytoplasm, [58]) and the distance  $\bar{x} - \bar{x}_n$  is varied from 0 to  $70\mu m$ .

For the evaluation of the ISI, the simulations are performed by sending two Gaussian pulses of the type in (260) where  $\sigma$  is set to  $0.32sec$ ,  $\Delta t$  is set to  $0.96sec$  and  $i = 1, 2$ . The parameter  $t_n$  is set to  $4.14sec$  from the starting time of the simulation. The resulting amplitude of the received pulses as function of the time  $t$  ranging from 0 to  $10sec$  and the distance  $\bar{x} - \bar{x}_n$  ranging from 0 to  $70\mu m$  are shown in Figure 61 (upper), (left) for the baseband modulation and (right) for the diffusion wave modulation. The ISI is evaluated by applying (242) with the computed values of  $s_{n,out}^1(t)$  and  $s_{n,out}^2(t)$ .

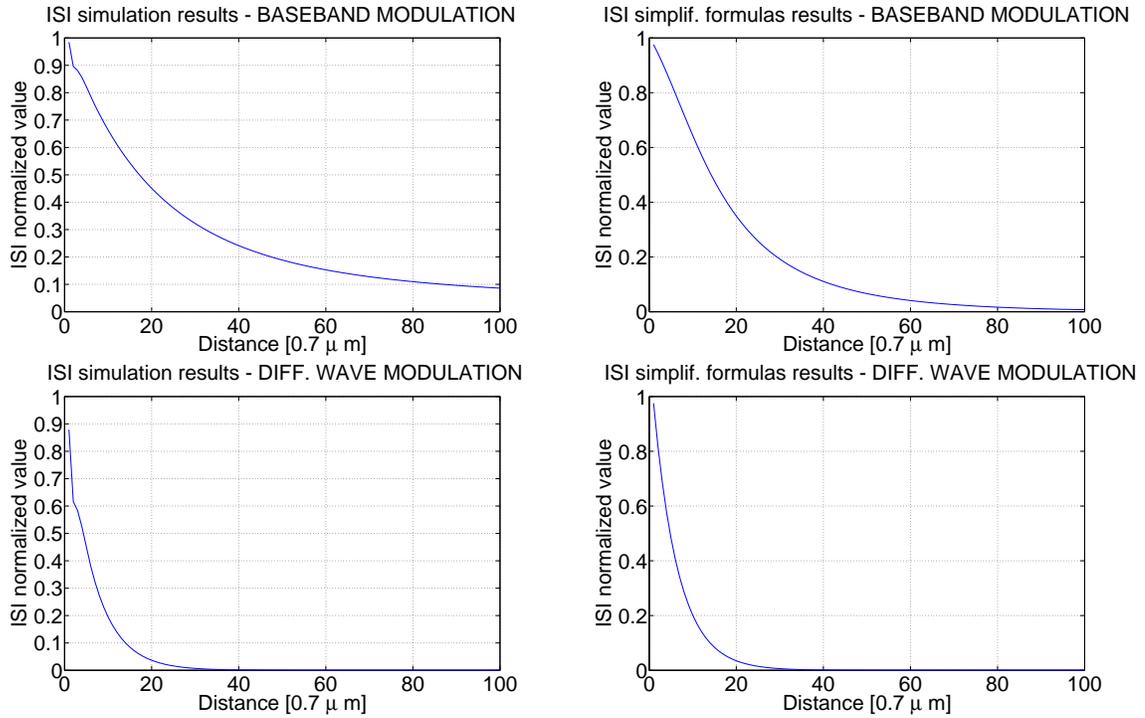
In Figure 62 (upper-right) and Figure 62 (lower-right) we show the results of the numerical evaluation of the formula (262) for the baseband modulation and the diffusion wave modulation, respectively. In case of baseband modulation, we apply (267) and (268), while for the diffusion wave modulation we used (269) and (270). The comparison of the ISI simulation results shown in Figure 62 (upper-left) and Figure 62 (lower-left) with the results of the simple formulas in Figure 62 (upper-right) and Figure 62 (lower-right) reveals strong similarities between the results of (242) and (262) in the case of baseband modulation and diffusion wave modulation and confirms the validity of the simple formulas for the ISI developed in Section 6.2.3. Moreover, both the results in terms of ISI in the simulation and



**Figure 61.** The received pulses as function of the time and the distance in the case of baseband modulation (left) and diffusion wave modulation (right) for the simulation-based for the simulation-based ISI evaluation (upper) and CCI evaluation (lower).

in the numerical evaluation confirm the relation in (273).

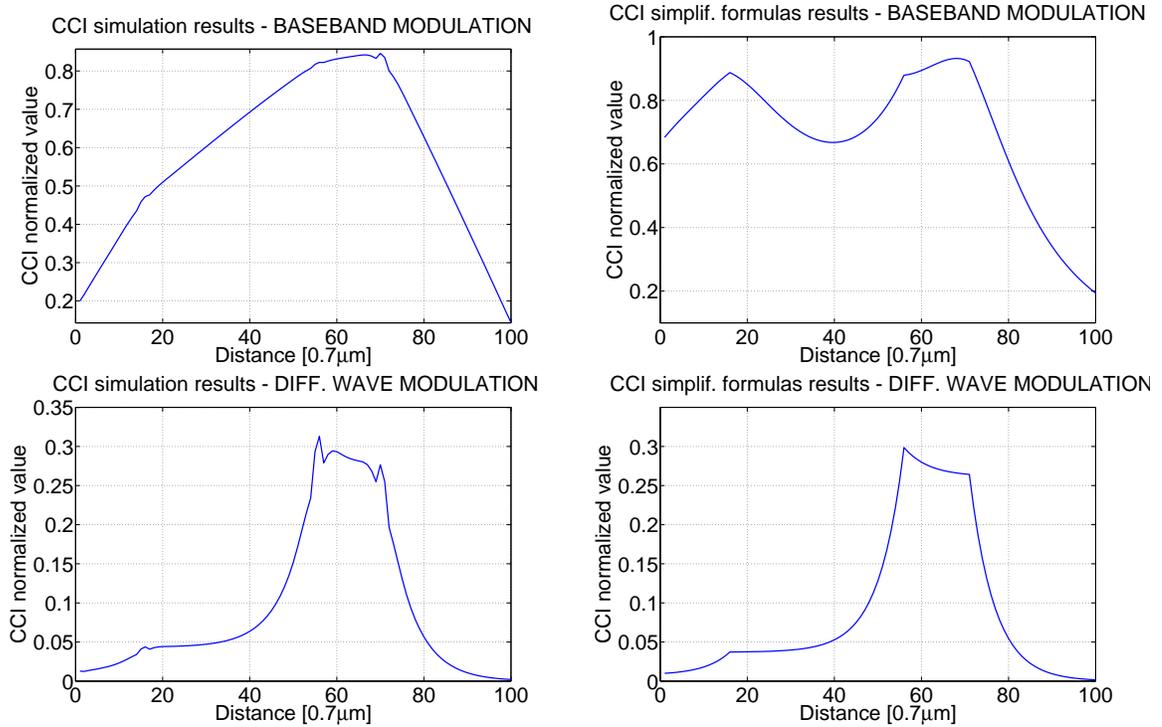
For the evaluation of the CCI, the simulations are performed by sending a Gaussian pulse of the type in (260) from each one of  $N = 4$  transmitters placed at distances from the second transmitter  $\|\bar{x}_n - \bar{x}_2\|$  equal to  $39.2\mu m$ ,  $0\mu m$ ,  $11.2\mu m$  and  $49.7\mu m$  respectively. In (260)  $\sigma$  is set to  $0.32sec$  and the index  $i$  is set to 0. We set  $t_1 = 4.14sec$ ,  $t_2 = 5.09sec$ ,  $t_3 = 7sec$  and  $t_4 = 8.9sec$ . The resulting amplitude of the received pulses  $s_{n,out}^i(t)$  as function of the time  $t$  ranging from 0 to  $10sec$  and the distance  $\bar{x}_R - \bar{x}_2$  ranging from 0 to  $70\mu m$  are shown in Figure 61 (lower-left) and Figure 61 (lower-right) for the baseband modulation and the diffusion wave modulation, respectively. The CCI is evaluated through (243) with the values of  $s_{n,out}^i(t)$ .



**Figure 62.** The ISI values for the baseband modulation scheme (upper) and the diffusion wave modulation scheme (lower) from the simulation (left) and from the simple formulas (right).

In Figure 63 (upper-right) and Figure 63 (lower-right) we show the results of the numerical evaluation of the formula (263) for the baseband modulation and the diffusion wave modulation, respectively. In case of baseband modulation, we apply (267) and (268), while for the diffusion wave modulation we used (269) and (270). The comparison of the CCI simulation results shown in Figure 63 (upper-left) with results of the simple formulas in Figure 63 (upper-right) clearly show the limit of the simple formulas to properly capture the real CCI curve as function of the distance. This is explained by the fact that the shape of the received pulse in case of baseband modulation is subject to a high value of dispersion and it is distorted with respect to the Gaussian shape assumed for the simple formula (263). This phenomenon is more clearly visible for the CCI computation (243) since its value is the result of the contributions of more than two received pulses, as in the case of the ISI (242). On the contrary, the results in Figure 63 (lower) guide to the same conclusion as mentioned above for the ISI and confirm the validity of the simple formulas for the CCI

developed in Section 6.2.3. Moreover, the result from (275) is supported by both Figure 63 (upper) and Figure 63 (lower).



**Figure 63. The CCI values for the the baseband modulation scheme (upper) and the diffusion wave modulation scheme (lower) from the simulation (left) and from the simple formulas (right).**

### 6.2.5 Conclusion

In this section of the Ph.D. thesis, the effects of the InterSymbol Interference (ISI) and the Co-Channel Interference (CCI) in a diffusion-based molecular communication system are analyzed. For this, characterization of the diffusion channel in terms of signal propagation is provided by studying two main parameters, namely, the attenuation and the dispersion, and simple closed-form formulas for the evaluation of the ISI and the CCI are derived for the baseband modulation and the diffusion wave modulation schemes. According to the ISI and CCI formulas, the diffusion wave modulation scheme shows lower values of interference with respect to the baseband modulation scheme. This is also confirmed by numerical results obtained through the simulation of the MC system, which also assess the validity of the derived simple closed-form formulas. The interference analysis presented

in this section of the Ph.D. thesis will contribute to the general understanding of molecular communication systems based on diffusion and it will support the design of nanonetwork architectures based on this paradigm.

## 6.3 Statistical-physical Model of Interference

### 6.3.1 Reference Models, Assumptions, and Definitions

In the following, we describe the main reference models, assumptions, and definitions used in this section of the Ph.D. thesis for the statistical-physical modeling of the interference in molecular nanonetworks.

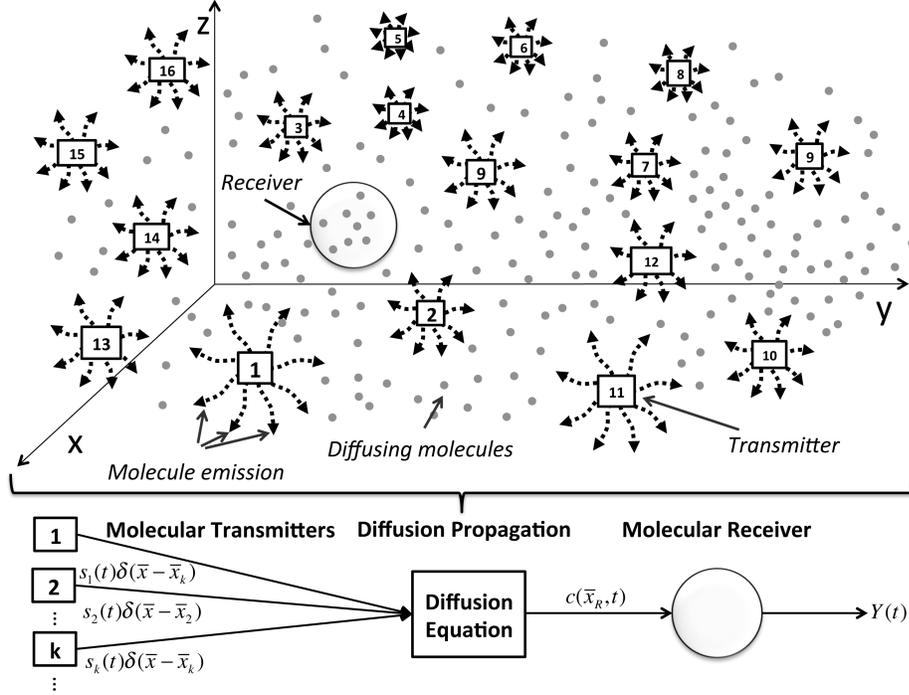
#### 6.3.1.1 Reference Molecular Nanonetwork

In the following, we detail the main elements of the reference molecular nanonetwork considered in this section of the Ph.D. thesis. As sketched in Figure 64, these elements are the molecular transmitters, responsible for the emission of molecular signals, the diffusion-based propagation, which broadcasts the molecular signals in the space by means of free molecule diffusion, and the molecular receiver, which senses the incoming molecular signals.

A **Molecular Transmitter**, identified by a number  $i$  and located at  $\bar{x}_k$ , is responsible for the emission of molecules in the space according to a molecular signal  $s_k(t)$  as function of the time  $t$ . We assume that all the transmitters emit molecules of the same species  $n$  within an equal definite volume  $V_T$ , whose size is negligible with respect to the distance between each transmitter and the receiver. Upon this emission of molecules, identified with the time derivative  $dX_n(t)/dt$  in the number  $X_n$  of molecules of species  $n$  inside the volume  $V_T$  at the transmitter  $i$ , each molecular transmitter  $i$  causes a change in the molecule concentration  $c(\bar{x}, t)$  at its location  $\bar{x}_k$ , which is expressed through the following relation:

$$\frac{\partial c(\bar{x}, t)}{\partial t} = \frac{1}{V_T} \frac{dX_n(t)}{dt} \delta(\bar{x} - \bar{x}_k) = s_k(t) \delta(\bar{x} - \bar{x}_k), \quad (277)$$

where  $\partial c(\bar{x}_k, t)/\partial t$  is the time derivative in the molecule concentration at the location  $\bar{x}_k$  and time  $t$ , and  $\delta(\cdot)$  is the Dirac delta. Moreover, we assume that a transmitter is able to produce



**Figure 64.** Reference diffusion-based molecular nanonetwork considered for the interference modeling.

molecules, thus resulting in a positive time first derivative  $dX_n(t)/dt > 0$  and a positive transmitted signal  $s_k(t) > 0$ , or to subtract molecules, thus resulting in a negative time first derivative  $dX_n(t)/dt < 0$  and in a negative transmitted signal  $s_k(t) < 0$ .

The **Diffusion Propagation** broadcasts the emitted molecular signal  $s_k(t)$  from each transmitter location  $\bar{x}_k$  to any other location  $\bar{x}$  in the space. In this section of the Ph.D. thesis, we rely on the assumption to have a 3-dimensional space, which contains a fluidic medium and has infinite extent in all the three dimensions. Moreover, the molecules of species  $n$  are all identical and undistinguishable, and they move independently from each other according to the Brownian motion. We define the total molecule concentration in the space as  $c_{base} + c(\bar{x}, t)$ , where  $c_{base}$  is a component of the molecule concentration that is homogeneous in the space and constant in time, while  $c(\bar{x}, t)$  is the varying component of the molecule concentration as a function of the space  $\bar{x}$  and time  $t$ . We assume that the component  $c_{base}$  has a value sufficient to keep the total molecule concentration positive throughout the space even when  $c(\bar{x}, t) < 0$  due to transmitters subtracting molecules from the space,

$s_k(t) < 0$ . The diffusion propagation is based on the following **Diffusion Equation** [26, 27] in the variable  $c(\bar{x}, t)$ :

$$\frac{\partial c(\bar{x}, t)}{\partial t} = D \nabla^2 c(\bar{x}, t), \quad (278)$$

where  $\partial(\cdot)/\partial t$  and  $\nabla^2(\cdot)$  are the time first derivative and the Laplacian operator (sum of the 3-dimensional spatial second derivatives), respectively.  $D$  is the diffusion coefficient and it is considered a constant parameter within the scope of this section of the Ph.D. thesis. This is in agreement with the assumption of having independent Brownian motion for every molecule in the space.

The **Molecular Receiver** senses the total incoming molecular concentration  $c_{base} + c(\bar{x}_R, t)$  at its location  $\bar{x}_R$  and recovers the received signal  $Y(t)$  from the varying component  $c(\bar{x}_R, t)$ . This is expressed by the following relation:

$$Y(t) = c(\bar{x}_R, t). \quad (279)$$

As a consequence, when no transmitter is emitting molecular signals (creating or subtracting molecules), the total molecule concentration is constant and equal to  $c_{base}$ , and the received signal  $Y(t)$  is equal to zero.

### 6.3.1.2 Assumptions on Interferers

For our interference study, we consider multiple molecular transmitters (interferers), each one emitting a molecular signal from a different location. We apply the following assumptions:

- The molecular transmitters are assumed to be infinite in number and distributed in the 3-dimensional space according to a spatial homogeneous Poisson process whose rate is equal to the transmitter density  $\lambda$ , which corresponds to the average number of transmitters per unit volume. For this, the probability to find a number  $k$  of transmitters in a region  $V$  of the space is expressed as follows:

$$P(k \text{ transmitters in } V) = \frac{[\lambda V]^k e^{-\lambda V}}{k!}. \quad (280)$$

- The molecular transmitters emit independent and identically distributed (i.i.d.) molecular signals  $s_k(t)$ . Each  $s_k(t)$  is a white Gaussian signal [66], whose values at each time instant  $t$  have zero mean and variance equal to  $\sigma^2$ , expressed as

$$s_k(t) \sim \mathcal{N}(0, \sigma^2) \quad \forall t. \quad (281)$$

The expression in (281) models the variability of the transmitter emissions according to the variance parameter  $\sigma^2$ , which corresponds to the average power of the molecular signals emitted by the transmitters.

The white Gaussian model for the molecular signals  $s_k(t)$  expressed in (281) is in agreement with a chemical description of the molecule emission at the molecular transmitters. Without loss of generality, we assume that the total molecule concentration  $c_{base} + c(\bar{x}_k, t)$  at each transmitter  $i$  is a function of  $M$  different chemical reactions involving  $N$  different chemical species (molecule types) within the transmitter definite volume  $V_T$ . According to the chemical Langevin equation approximation [94], the time derivative  $dX_n(t)/dt$  from (277) in the number  $X_n$  of species- $n$  molecules, and function of the time  $t$ , is given by the following expression:

$$\frac{dX_n(t)}{dt} = \sum_{m=1}^M \nu_{mn} a_m(\mathbf{X}(t)) + \sum_{m=1}^M \nu_{mn} \sqrt{a_m(\mathbf{X}(t))} \Gamma_m(t) \quad (282)$$

where  $\mathbf{X}(t) = [X_1(t), X_2(t), \dots, X_N(t)]'$  is a vector that contains the number of molecules of each reacting species,  $\nu_{mn}$  corresponds to the change in the number of molecules of the chemical species  $n$  produced by the chemical reaction  $m$ ,  $a_m(\mathbf{X}(t))$ , which is called *propensity function*, is the probability that the chemical reaction  $m$  will occur within the transmitter volume as function of the vector  $\mathbf{X}(t)$ , and  $\Gamma_m(t)$  are i.i.d. white Gaussian signals. Under the assumption to have the chemical reactions at equilibrium within every transmitter volume, which is expressed as  $\sum_{m=1}^M \nu_{mn} a_m(\mathbf{X}(t)) = 0$ , and given (282), the molecular signal  $s_k(t)$  in (277) is equal to a sum of i.i.d. white Gaussian signals as follows:

$$s_k(t) = \frac{1}{V_T} \sum_{m=1}^M \nu_{mn} \sqrt{a_m(\mathbf{X}(t))} \Gamma_m(t). \quad (283)$$

As a consequence of the property of a linear combination of i.i.d. Gaussian random variables [66],  $s_k(t)$ , as expressed in (281), is a white Gaussian signal with zero mean and variance  $\sigma^2$  equal to:

$$\sigma^2 = \frac{1}{V_T^2} \sum_{m=1}^M v_{mn}^2 a_m(\mathbf{X}(t)) \quad (284)$$

where  $\mathbf{X}(t)$  is the vector that contains the number of molecules of each reacting species,  $v_{mn}$  corresponds to the change in the number of molecules of the chemical species  $n$  produced by the chemical reaction  $m$ , and the propensity function  $a_m(\mathbf{X}(t))$  is the probability that the chemical reaction  $m$  will occur within the transmitter volume as function of the vector  $\mathbf{X}(t)$ .

### 6.3.1.3 Definition of Interference

We define as interference the received signal  $Y(t)$  expressed as the propagation function  $f_d(\cdot)$  of the multiple transmitted molecular signals  $s_k$ , where  $k = 0, \dots, \infty$ , as follows:

$$Y(t) = f_d \left( \sum_{k=0}^{\infty} s_k(t) \delta(\bar{x} - \bar{x}_k) \right), \quad (285)$$

where  $\delta(\cdot)$  is the Dirac delta,  $f_d(\cdot)$  is the diffusion propagation function that transforms the sum of transmitted molecular signals  $s_k$  into the incoming molecular concentration  $c(\bar{x}_R, t)$  at the receiver location  $\bar{x}_R$  through the diffusion equation (324) and, according to (279), into the received signal  $Y(t)$ .

Due to the linearity of (324) [26, 27], given multiple molecular signals transmitted simultaneously from multiple transmitters, the resulting varying component  $c(\bar{x}, t)$  of the molecule concentration is the sum of the varying components of the molecule concentration resulting from the emission of each molecular transmitter, computed as if each transmitter were emitting alone (*additive channel*). As a consequence, we can express the received signal  $Y(t)$  as the sum of the propagation functions applied separately to each transmitted molecular signal, which results into

$$Y(t) = \sum_{k=0}^{\infty} f_d(s_k(t) \delta(\bar{x} - \bar{x}_k)), \quad (286)$$

where the propagation function  $f_d(\cdot)$  is computed as the solution of the diffusion equation (324) when a single transmitter is emitting a molecular signal  $s_k(t)$ . For this, we consider the following expression:

$$\frac{\partial c(\bar{x}, t)}{\partial t} = D\nabla^2 c(\bar{x}, t) + s_k(t)\delta(\bar{x} - \bar{x}_k) . \quad (287)$$

The solution of (287) in terms of  $c(\bar{x}, t)$  corresponds to the following propagation function  $f_d(\cdot)$ :

$$f_d(s_k(t)\delta(\bar{x} - \bar{x}_k)) = c(\bar{x}, t) = g_d(r_k, t) * s_k(t) = \int_0^\infty g_d(r_k, \tau) s_k(\tau - t) d\tau , \quad (288)$$

where  $(. * .)$  is the convolution operator [51], and  $g_d(r_k, t)$  is the Green's function of the diffusion equation [86], equal to

$$g_d(r_k, t) = \frac{e^{-\frac{r_k^2}{4Dt}}}{(4\pi Dt)^{3/2}} , \quad (289)$$

where  $r_k$  is the Euclidian distance between the transmitter  $k$  location and the receiver location,  $r_k = \|\bar{x}_k - \bar{x}_R\|$ , and  $D$  is the diffusion coefficient. As a result, we can express the received signal  $Y(t)$  as

$$Y(t) = \sum_{k=0}^{\infty} g_d(r_k, t) * s_k(t) , \quad (290)$$

where  $g_d(r_k, t)$  is expressed in (289),  $(. * .)$  is the convolution operator [51], and  $s_k(t)$  is the molecular signal transmitted from each transmitter  $k$ , whose distribution is given by (281).

### 6.3.2 Statistical-physical Interference Modeling

The goal of the statistical-physical interference modeling is to find a **probabilistic description of the received signal  $Y(t)$**  expressed in (290), **as function of the transmitter density  $\lambda$ , the diffusion coefficient  $D$ , and the average power  $\sigma^2$**  of the molecular signals emitted by the transmitters.

In standard statistical-physical modeling of the interference for radio communication networks [93], since the propagation function corresponds to a multiplication of each transmitted signal (uncorrelated random process with zero mean value) by the radio propagation

amplitude loss, independent with respect to the time variable, the received signal  $Y(t)$  is an uncorrelated stochastic process with zero mean value. As a consequence, the received signal  $Y(t)$  can be probabilistically described with the Probability Density Function (PDF)  $P_Y(y)$  of a time sample, for which analytical expressions are usually provided in terms of log-characteristic functions [95].

### 6.3.2.1 Probabilistic Description of the Received Signal

In the context of diffusion-based molecular nanonetworks, as a consequence of the expression of the propagation function in (289) as function of the time variable  $t$ , the received signal  $Y(t)$  is in general a correlated stochastic process, which cannot be described by the PDF  $P_Y(y)$  of a single time sample. A probabilistic description of the received signal  $Y(t)$  can be provided upon the following considerations:

- Consider a realization of the spatial homogeneous Poisson process of the transmitter locations, expressed in (280), which results into a set of values  $R = \{r_k\}_{k=1,2,\dots,\infty}$  for the distances  $r_k$  between each transmitter  $k = 1, 2, \dots, \infty$  and the receiver.
- Given the previous consideration, each term of the sum in (290) is a convolution of a deterministic function  $g_d(r_k, t)$  of the time  $t$  with a zero-mean Gaussian white random signal  $s_k(t)$  with zero mean and variance equal to  $\sigma^2$ . The result of this convolution is a zero-mean stationary Gaussian process  $y_k|r_k$  with autocorrelation function  $\mathcal{R}_{y_k|r_k}(t)$  equal to  $\sigma^2$  multiplied by the correlation of  $g_d(r_k, t)$  with itself [66]. This is expressed as follows:

$$\mathcal{R}_{y_k|r_k}(t) = \sigma^2 \int_0^\infty g_d(r_k, \tau)g_d(r_k, \tau + t)d\tau . \quad (291)$$

- The autocorrelation of the sum of two uncorrelated random processes is a random process whose autocorrelations is the sum of their autocorrelations [66].

As a consequence of the aforementioned considerations, the received signal  $Y|R$ , given a realization of the transmitter locations  $R$ , is a **zero-mean stationary Gaussian Process**

(GP), probabilistically described as follows:

$$Y|R \sim \mathbf{GP}(0, \mathcal{R}_{Y|R}(t)) , \quad (292)$$

whose autocorrelation function  $\mathcal{R}_{Y|R}(t)$  is equal to the sum for each transmitter  $k = 1, 2, \dots, \infty$  of the autocorrelation function  $\mathcal{R}_{y_k|r_k}(t)$  in (291), expressed as

$$\mathcal{R}_{Y|R}(t) = \sum_{k=0}^{\infty} \mathcal{R}_{y_k|r_k}(t) . \quad (293)$$

Since  $Y|R$  is a continuous time stationary random process, according to the Wiener-Khintchine theorem [66] it can be equivalently described in terms of Power Spectral Density (PSD), which corresponds to the Fourier transform [51] of the autocorrelation function  $\mathcal{R}_{Y|R}(t)$ . Given the expressions in (293) and (291), the PSD  $\mathcal{S}_{Y|R}(\omega)$  results in the following:

$$\mathcal{S}_{Y|R}(\omega) = \sigma^2 \sum_{k=0}^{\infty} |G_d(r_k, \omega)|^2 , \quad (294)$$

where  $|\cdot|^2$  denotes the squared absolute value operator, and  $G_d(r_k, \omega)$  is the Fourier transform [51] of  $g_d(r_k, t)$  in (289), whose expression is

$$G_d(r_k, \omega) = \frac{e^{-(1+j)\sqrt{\frac{\omega}{2D}}r_k}}{\pi D r_k} , \quad (295)$$

where  $r_k$  is the Euclidian distance between the transmitter  $k$  and the receiver, and  $D$  is the diffusion coefficient.

### 6.3.2.2 Statistical-physical Modeling of the Received Power Spectral Density

The received PSD  $\mathcal{S}_Y(\omega)$  is defined as the distribution of the power of the received signal  $Y$  over each frequency  $\omega$ . Given the presence of multiple transmitters, and the probabilistic assumptions described in Section 6.3.1.2, the received PSD  $\mathcal{S}_Y(\omega)$  is a measure of the power of the interference which affects the communication system in each received frequency  $\omega$ . As a consequence, we aim at the statistical-physical modeling of the received PSD  $\mathcal{S}_Y(\omega)$  through the expression of its PDF  $P_{\mathcal{S}_Y(\omega)}(s)$  as a function  $f(\cdot)$  of the PSD value  $s$ , the frequency  $\omega$ , the transmitter density  $\lambda$ , the diffusion coefficient  $D$ , and the average power

$\sigma^2$  of the molecular signals emitted by the transmitters. This is expressed as follows:

$$P_{\mathcal{S}_Y(\omega)}(s) = f(s, \omega, \lambda, D, \sigma^2) . \quad (296)$$

As detailed in the following, the PDF  $P_{\mathcal{S}_Y(\omega)}(s)$  of the PSD  $\mathcal{S}_Y(\omega)$  is computed from the PSD  $\mathcal{S}_{Y|R}(\omega)$  in (294) by taking into account the spatial homogeneous Poisson process of the transmitter locations in (280).

The PDF  $P_{\mathcal{S}_Y(\omega)}(s)$ , as happens in standard statistical-physical modeling for the PDF  $P_Y(y)$  of the interference for radio communication networks [93], does not have a closed-form mathematical expression. As a consequence, by following the standard statistical-physical modeling of the interference for radio communication networks [93], we aim at the expression of the log-characteristic function  $\psi_{\mathcal{S}_Y(\omega)}(\Omega)$  of the received PSD  $\mathcal{S}_Y(\omega)$ , which is defined as the natural logarithm of the characteristic function  $\phi_{\mathcal{S}_Y(\omega)}(\Omega)$ , as

$$\psi_{\mathcal{S}_Y(\omega)}(\Omega) = \ln [\phi_{\mathcal{S}_Y(\omega)}(\Omega)] . \quad (297)$$

The characteristic function  $\phi_{\mathcal{S}_Y(\omega)}(\Omega)$  of the received PSD  $\mathcal{S}_Y(\omega)$  is defined as the expected value of the function  $e^{j\Omega s}$  of the PSD value  $s$ :

$$\phi_{\mathcal{S}_Y(\omega)}(\Omega) = E_{\mathcal{S}_Y(\omega)} [e^{j\Omega s}] = \int P_{\mathcal{S}_Y(\omega)}(s) e^{j\Omega s} ds , \quad (298)$$

The PDF  $P_{\mathcal{S}_Y(\omega)}(s)$  of the PSD  $\mathcal{S}_Y(\omega)$  is computed through the Fourier transform [51] of the exponential with the log-characteristic function  $\psi_{\mathcal{S}_Y(\omega)}(\Omega)$  as argument. This is expressed as follows:

$$P_{\mathcal{S}_Y(\omega)}(s) = \int e^{\psi_{\mathcal{S}_Y(\omega)}(\Omega)} e^{-j\Omega s} d\Omega , \quad (299)$$

As mentioned above, the formula in (299) does not in general result in a closed-form expression, and it is computed through numerical methods.

In the following, we derive the log-characteristic function  $\psi_{\mathcal{S}_Y(\omega)}(\Omega)$ , which admits an analytical expression as a function  $\Psi(\cdot)$  of the transmitter density  $\lambda$ , the diffusion coefficient  $D$  and the average power  $\sigma^2$  of the molecular signals emitted by the transmitters,

PSD frequency variable  $\omega$ , the characteristic function frequency variable  $\Omega$ , expressed as follows:

$$\psi_{S_Y(\omega)}(\Omega) = \Psi(\lambda, D, \sigma^2, \omega, \Omega) . \quad (300)$$

### 6.3.3 Log-Characteristic Function and PDF of the Received Power Spectral Density

In this section, we analytically derive the log-characteristic function  $\psi_{S_Y(\omega)}(\Omega)$  of the received PSD  $S_Y(\omega)$ . Through the derivation detailed in Section 6.3.3.1, we obtain the following analytical expression:

$$\psi_{S_Y(\omega)}(\Omega) = j \frac{16 \sqrt{2} \lambda \sigma^2 \Omega}{3\pi \sqrt{D^3} \omega} \int_0^\infty (x+1) e^{-2x} e^{-j \frac{e^{-2x}}{x^2} \frac{\sigma^2 \omega \Omega}{2\pi^2 D^3}} dx , \quad (301)$$

where  $\lambda$  is the transmitter density (number of transmitters per unit volume),  $D$  is the diffusion coefficient,  $\sigma^2$  is the average power of the molecular signals emitted by the transmitters,  $\omega$  is the PSD frequency variable, and  $\Omega$  is the frequency variable of the characteristic function. Subsequently, we derive the PDF  $P_{S_Y(\omega)}(s)$  of the received PSD  $S_Y(\omega)$  by numerically computing the expression in (299).

#### 6.3.3.1 Derivation of the Log-characteristic Function $\psi_{S_Y(\omega)}(\Omega)$

In the following, we adapt the analytical computation of the log-characteristic function of the received signal in case of radio communication networks [95] to derive the log-characteristic function  $\psi_{S_Y(\omega)}(\Omega)$  of the received PSD  $S_Y(\omega)$  in diffusion-based molecular nanonetworks. By applying the rule of the iterated expectations [66], we can perform the expectation in (298) with respect to the transmitter locations  $R = \{r_k\}_{k=1,2,\dots,\infty}$ , where  $r_k$  are the random distances between each transmitter  $k = 1, 2, \dots, \infty$  and the receiver, and substitute the PSD value  $s$  with the PSD  $S_{Y|R}(\omega)$  of the received signal given a realization of the transmitter locations. As a consequence, we obtain the following expression:

$$\phi_{S_Y(\omega)}(\Omega) = E_R \left[ e^{j\Omega S_{Y|R}(\omega)} \right] , \quad (302)$$

where the PSD  $S_{Y|R}(\omega)$  is computed through (294) and (295).

Since the transmitter locations are resulting from a spatial homogeneous Poisson process, as described in Section 6.3.1.2, the distances  $r_k$  are i.i.d. random variables, and the distribution in the number  $k$  of molecular transmitters in a space region  $V$  is given by (280). As a consequence, for an infinite space region, represented by a sphere centered at the receiver with infinite radius, namely,  $V = \lim_{\rho \rightarrow \infty} (4/3)\pi\rho^3$ , we derive the following expression from (294) applied to (302):

$$\phi_{S_Y(\omega)}(\Omega) = \lim_{\rho \rightarrow \infty} \sum_{k=0}^{\infty} \left( E_{r_k} \left[ e^{j\Omega\sigma^2 |G_d(r_k, \omega)|^2} \right] \right)^k \cdot \frac{[\lambda(4/3)\pi\rho^3]^k e^{-\lambda(4/3)\pi\rho^3}}{k!}, \quad (303)$$

where the summation from (294) is substituted with the power  $k$  operator  $(\cdot)^k$ , and the average operator  $E_R[\cdot]$  is written in terms of summation in  $k$  of the average operator  $E_{r_k}[\cdot]$  of the  $k$ -th transmitter distance, weighted by the probability density from (280).

By applying the following Taylor series expansion [97] substitution to (303):

$$\sum_{k=0}^{\infty} \frac{x^k}{k!} = e^x, \quad (304)$$

and by applying the definition of log-characteristic function  $\psi_{S_Y(\omega)}(\Omega)$  from (297), we obtain the following expression:

$$\psi_{S_Y(\omega)}(\Omega) = \lim_{\rho \rightarrow \infty} \frac{4}{3}\pi\rho^3 \lambda \left( E_{r_k} \left[ e^{j\Omega\sigma^2 |G_d(r_k, \omega)|^2} \right] - 1 \right). \quad (305)$$

Since the transmitters are distributed according to a Poisson process (280), the distribution of the distance between the transmitter and the receiver, given a space region  $V = (4/3)\pi\rho^3$ , has the following expression:

$$P_{r_k}(r) = \frac{3r^2}{\rho^3}, \quad 0 \leq r \leq \rho. \quad (306)$$

If we express in (305) the average operator  $E_{r_k}[\cdot]$  of the distance  $r_k$  between the transmitter and the receiver by using the distribution of this distance in (306), we obtain the following:

$$\psi_{S_Y(\omega)}(\Omega) = \lim_{\rho \rightarrow \infty} \frac{4}{3}\pi\rho^3 \lambda \left( \int_0^\rho e^{j\Omega\sigma^2 |G_d(r, \omega)|^2} \frac{3r^2}{\rho^3} dr - 1 \right). \quad (307)$$

By using the formula of the integration by parts [97] for the integral in (307), we obtain the following expression:

$$\psi_{S_Y(\omega)}(\Omega) = \lim_{\rho \rightarrow \infty} \frac{4}{3} \pi \rho^3 \lambda \left( e^{j\Omega\sigma^2 \frac{e^{-2\sqrt{\frac{\omega}{2D}}\rho}}{(\pi D \rho)^2}} - \frac{4j\Omega\sigma^2}{\rho^3(\pi D)^2} \int_0^\rho \left( \sqrt{\frac{\omega}{2D}} r + 1 \right) e^{-2\sqrt{\frac{\omega}{2D}} r} e^{j\Omega\sigma^2 \frac{e^{-2\sqrt{\frac{\omega}{2D}} r}}{(\pi D r)^2}} dr - 1 \right). \quad (308)$$

We note the following result:

$$\lim_{\rho \rightarrow \infty} \rho^3 \left( e^{j\Omega\sigma^2 \frac{e^{-2\sqrt{\frac{\omega}{2D}}\rho}}{(\pi D \rho)^2}} - 1 \right) = 0, \quad (309)$$

which is demonstrated by considering the following inequality:

$$\frac{e^{-2\sqrt{\frac{\omega}{2D}}\rho}}{(\pi D \rho)^2} < \frac{1}{\rho}, \quad \text{for } \rho \rightarrow \infty, \quad (310)$$

and by repeatedly applying L'Hôpital's rule [97] to the following limit:

$$\lim_{\rho \rightarrow \infty} \rho^3 (e^{1/\rho} - 1) = 0. \quad (311)$$

By applying (309) to (308), we obtain the following expression:

$$\psi_{S_Y(\omega)}(\Omega) = j \frac{16}{3} \frac{\lambda \sigma^2 \Omega}{\pi D^2} \int_0^\infty \left( \sqrt{\frac{\omega}{2D}} r + 1 \right) e^{-2\sqrt{\frac{\omega}{2D}} r} e^{j\Omega\sigma^2 \frac{e^{-2\sqrt{\frac{\omega}{2D}} r}}{(\pi D r)^2}} \frac{r^3}{R^3} dr. \quad (312)$$

By operating in the integral of (312) the following variable substitution:

$$x = \sqrt{\frac{\omega}{2D}} r, \quad (313)$$

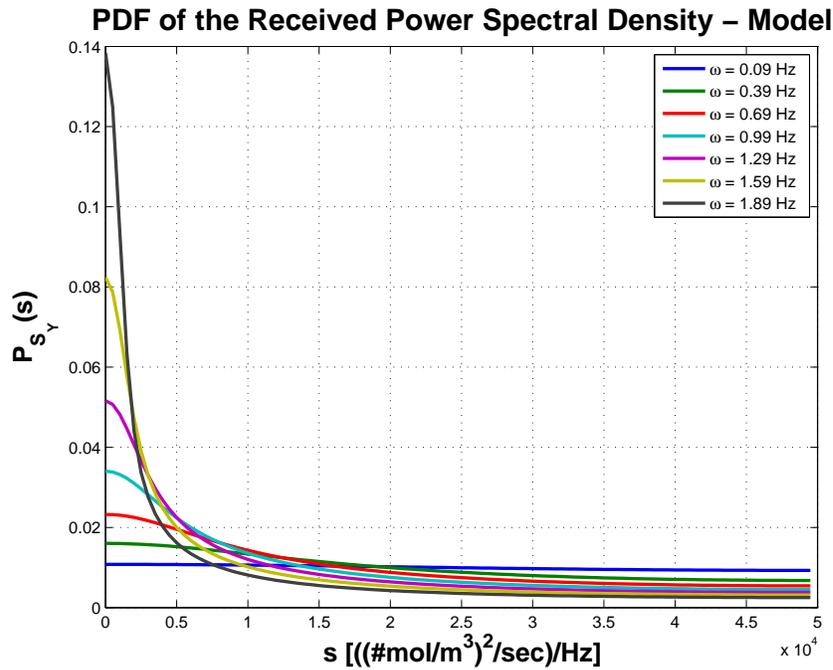
we obtain the final expression of the log-characteristic function  $\psi_{S_Y(\omega)}(\Omega)$  of the received PSD  $S_Y(\omega)$ , which is as follows:

$$\psi_{S_Y(\omega)}(\Omega) = j \frac{16\sqrt{2}\lambda\sigma^2\Omega}{3\pi\sqrt{D^3\omega}} \int_0^\infty (x+1) e^{-2x} e^{-j\frac{e^{-2x}}{x^2} \frac{\sigma^2\omega\Omega}{2\pi^2 D^3}} dx, \quad (314)$$

where  $\lambda$  is the transmitter density (number of transmitters per unit volume),  $D$  is the diffusion coefficient,  $\sigma^2$  is the average power of the molecular signals emitted by the transmitters,  $\omega$  is the PSD frequency variable, and  $\Omega$  is the frequency variable of the characteristic function.

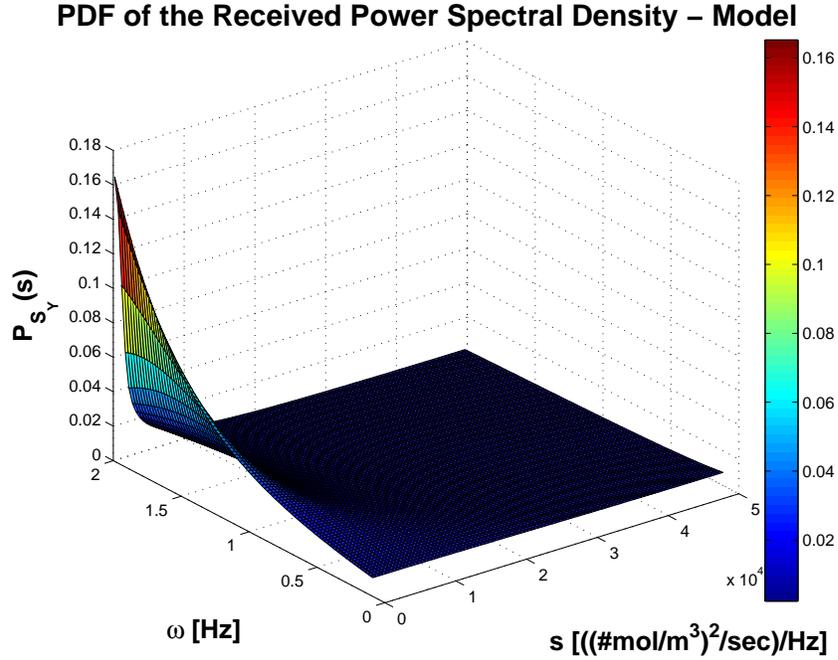
### 6.3.3.2 Derivation of the PDF $P_{S_Y(\omega)}(s)$

In this section, we derive the PDF  $P_{S_Y(\omega)}(s)$  of the received PSD  $S_Y(\omega)$ . In general, the log-characteristic function expressed in (314) does not have an expression which can be recognized as from a known probability distribution. For this, we numerically compute the formula in (299) by using the MATLAB © fft function applied to the values of the expression in (314). We also numerically compute the infinite integral in (314) by using the MATLAB © numerical integration.



**Figure 65.** PDF  $P_{S_Y(\omega)}(s)$  of the received PSD  $S_Y(\omega)$ . Different curves refer to different values of the frequency  $\omega$ .

The numerical results in terms of PDF  $P_{S_Y(\omega)}(s)$  of the received PSD  $S_Y(\omega)$  are shown in Figure 65 and Fig 66. The values of the PDF  $P_{S_Y(\omega)}(s)$  are computed for a transmitter density  $\lambda$  equal to  $10^9$  [transmitters  $\text{m}^{-3}$ ], an average power  $\sigma^2$  of the molecular signals equal to  $10^6$  [molecules<sup>2</sup>  $\text{m}^{-6}$   $\text{sec}^{-3}$ ], and for values of the PSD value  $s$  ranging from 0 to  $5 \cdot 10^4$  [molecules<sup>3</sup>  $\text{m}^{-6}$   $\text{sec}$   $\text{Hz}^{-1}$ ]. The diffusion coefficient  $D \sim 10^9$  [ $\text{m}^2\text{sec}^{-1}$ ] is set to the diffusion coefficient of molecules diffusing in a biological environment (cellular cytoplasm, [58]). Different curves in Figure 65 refer to different values of the frequency



**Figure 66.** PDF  $P_{S_Y(\omega)}(s)$  of the received PSD  $S_Y(\omega)$  for a range of frequencies  $\omega$  from 0 to 2 Hz.

$\omega$ , from 0.09Hz to 1.89Hz, while Figure 66 shows the PDF  $P_{S_Y(\omega)}(s)$  values for a range of frequencies  $\omega$  from 0 to 2 Hz.

As apparent from Figure 65 and Fig 66, the curves of the PDF  $P_{S_Y(\omega)}(s)$  as function of the PSD value  $s$  tend to horizontal lines for low values of the frequency  $\omega$ , while they tend to concentrate the higher values around  $s = 0$  as the frequency  $\omega$  increases. This is an expected behavior since, according to the absolute value of the expression of the Fourier transform of the propagation function  $G_d(r_k, \omega)$  in (295), which is a negative exponential function of the square root of the frequency  $\omega$ , lower frequencies are subject to lower attenuation than higher frequencies in the diffusion propagation. As a consequence, for lower frequencies the received PSD tends to have a shape similar to the PSD of the white transmitted signals  $s_k(t)$  in (281), equally distributed among all the possible PSD values  $s$  with a probability value around 0.01. On the contrary, since higher frequencies are more attenuated, for a high  $\omega$  lower values of the received PSD are more probable, which is more likely distributed around  $s = 0$ , with the highest value around 0.16 for  $\omega$  close to 2 Hz.

### 6.3.4 Numerical Results

In this section, we provide a simulation environment to evaluate the statistical-physical interference model presented in this section of the Ph.D. thesis (Section 6.3.4.1). In addition, we study the probability of interference, defined as the probability for a single molecular signal sent by a transmitter to suffer interference at the receiver, by using both the statistical-physical interference model and the simulation environment (Section 6.3.4.2).

#### 6.3.4.1 Simulation-based Evaluation

The simulation environment is based on the following additional assumptions:

- The space where the transmitters are distributed is confined within a sphere with radius  $\rho$  around the receiver location. This is motivated by the need to have in the simulation environment a finite number of transmitters, which is equal to  $K = \lfloor \lambda(4/3)\pi\rho^3 \rfloor$ , where  $\lfloor \cdot \rfloor$  denotes the rounding to the nearest lower integer.
- The transmitted signal  $s_k(n/f_s)$  from each transmitter is discrete, sampled with a frequency  $f_s$ , and composed by  $N_s$  samples.
- The simulation is repeated for a number of iterations  $Iter$ , where each iteration is based on i) a different realization of the spatial Poisson process with density  $\lambda$  of the molecular transmitter distribution expressed in (280), ii) a different realization of the Gaussian process in (281) with variance equal to  $\sigma^2$  for each transmitter  $k$  and for each sample  $s_k(n/f_s)$ .

The PDF  $P_{S_Y(\omega)}(s)$  of the received PSD  $S_Y(\omega)$ , where  $\omega = qf_s/N_s$ , and  $q = 1, \dots, N_s$ , is computed though the following expression:

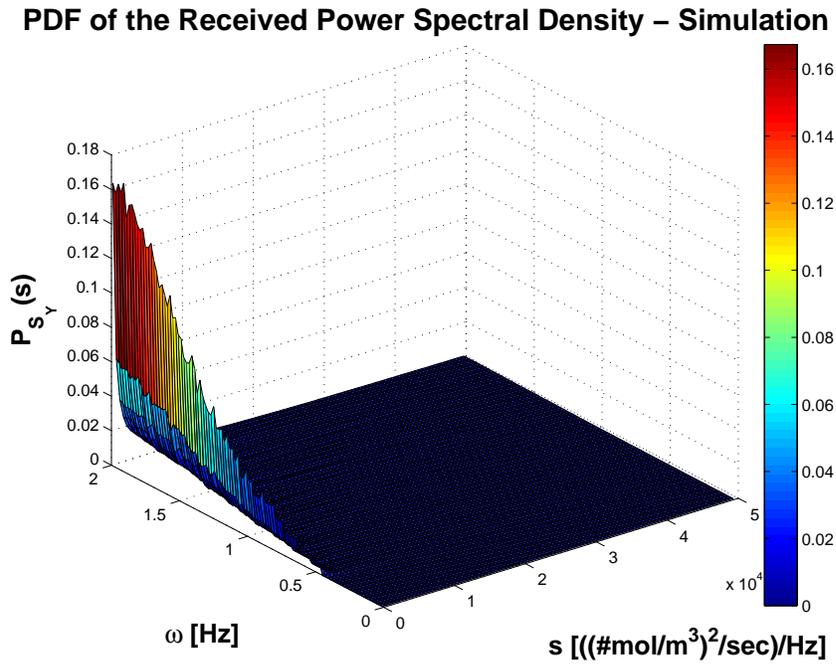
$$P_{S_Y(\omega)}(s)|_{\omega=qf_s/N_s} = \frac{1}{Iter} \sum_{l=1}^{Iter} \mathbf{1}_{S_{Y_l}(qf_s/N_s)=s}, \quad (315)$$

where  $\mathbf{1}_{S_{Y_l}(qf_s/N_s)=s}$  is non-zero and equal to 1 only when the PSD  $S_{Y_l}(qf_s/N_s)$  from the  $l$ -th iteration is equal to the value  $s$  at frequency  $qf_s/N_s$ , and  $f_s$  and  $N_s$  are the sampling

frequency and the number of samples for the transmitted molecular signals, respectively. The PSD  $S_{Y_i}(qf_s/N_s)$  results from the following formula:

$$S_{Y_i}(qf_s/N_s) = \left( \sum_{k=1}^K S_k(qf_s/N_s) G_d(r_k, qf_s/N_s) \right)^2, \quad (316)$$

where  $S_k(qf_s/N_s)$  is the discrete Fourier transform of  $s_k(n/f_s)$ , computed through the MATLAB © fft function, and  $G_d(r_k, qf_s/N_s)$  is the Fourier transform of the propagation function in (295) computed at the frequency value  $qf_s/N_s$ .



**Figure 67.** Simulation-based PDF  $P_{S_Y(\omega)}(s)$  of the received PSD  $S_Y(\omega)$  for a range of frequencies  $\omega$  from 0 to 2 Hz.

In Figure 67 we show the values of  $P_{S_Y(\omega)}(s)$  computed for the same parameters as for the results in Figure 66, namely, a transmitter density  $\lambda$  equal to  $10^9$  [transmitters  $\text{m}^{-3}$ ], an average power  $\sigma^2$  of the molecular signals equal to  $10^6$  [molecules<sup>2</sup>  $\text{m}^{-6}$   $\text{sec}^{-3}$ ] a diffusion coefficient  $D \sim 10^9$  [ $\text{m}^2\text{sec}^{-1}$ ], and for PSD values  $s$  ranging from 0 to  $5 \cdot 10^4$  [molecules<sup>3</sup>  $\text{m}^{-6}$   $\text{sec Hz}^{-1}$ ]. Moreover, the simulation is run with the following parameters: a spherical space radius  $\rho = 19\mu\text{m}$ , a sampling frequency  $f_s = 100\text{Hz}$ , a number of samples  $N_s = 10^4$ , and a number of iterations  $Iter = 50$ . The curves in Figure 67 have been also post-processed

through the use of a moving average filter [98] along the dimension of the PSD value  $s$  to reduce the noise given by the limited dataset.

The simulation-based results in terms of  $P_{S_Y(\omega)}(s)$  in Figure 67 show a high degree of similarity with the values computed through the statistical-physical model in Fig 66. Also in the simulation-based results, the curves of  $P_{S_Y(\omega)}(s)$  as function of the PSD value  $s$  tend to horizontal lines for low values of the frequency  $\omega = qf_s/N_s$ , while they tend to concentrate the higher values around  $s = 0$  as the frequency  $\omega$  increases. While for high frequencies  $\omega$  around 2 Hz the simulation-based PDF has a value around  $s = 0$  of 0.16, very close to the results of the statistical-physical model, for lower frequencies the values of the model-based PDF are overall lower than the values from the statistical-physical model. We believe that these differences between the values in Figure 66 and Figure 67 are due to the limited number of transmitters and the sampling of the molecular signals  $s_k$  considered for the simulation environment.

#### 6.3.4.2 Probability of Interference

We define here the probability of interference  $P_{Interf}(\omega)$  as the probability of having at the receiver a contribution from the interference whose PSD at frequency  $\omega$  exceeds the PSD of a contribution coming from a single transmitter. This single transmitter is placed at a distance  $r_{Tx}$  from the receiver, and it transmits a signal  $s_{Tx}(t)$  with power equal to  $\sigma_{tx}^2$ , expressed as

$$s_{Tx}(t) = \sigma_{tx} \frac{\sin[t(\omega_b - \omega_a)]}{t} e^{j\omega_a t}, \quad (317)$$

The PSD of the signal  $s_{Tx}(t)$  is then constant over the frequency range defined by  $\omega_a$  and  $\omega_b$ , and it is expressed as follows:

$$\mathcal{S}_{Tx}(\omega) = \sigma_{tx}^2 \text{rect}\left(\frac{\omega - \omega_a}{\omega_b - \omega_a}\right), \quad (318)$$

where  $\text{rect}(\cdot)$  is the rectangular function, and  $\sigma_{tx}^2$  is the constant PSD value. The contribution  $\mathcal{S}_{Rx}(\omega)$  to the PSD of the received signal coming from the transmitted signal  $s_{Tx}(t)$  is

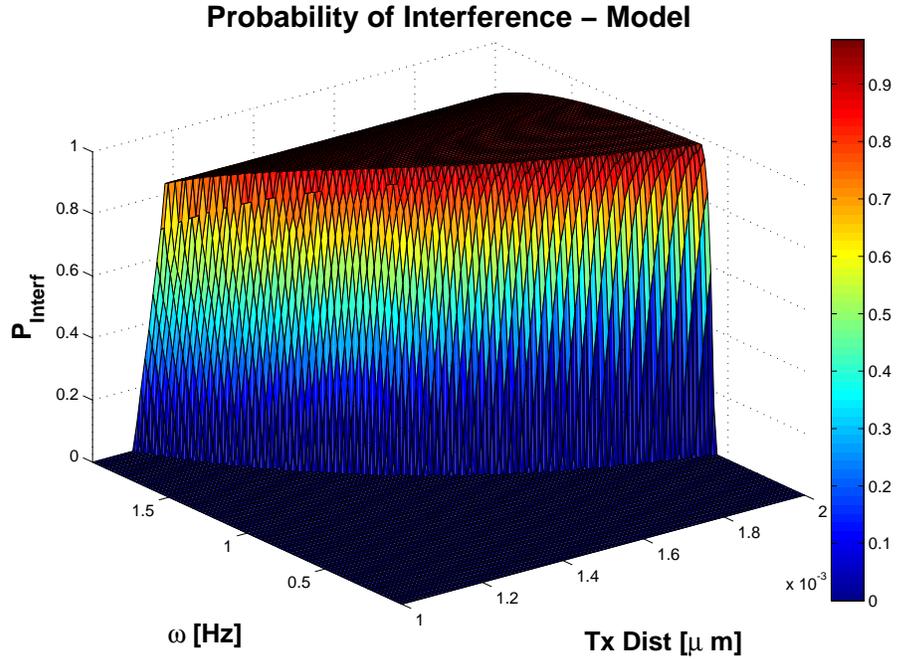
given as

$$\mathcal{S}_{R_x}(\omega) = \mathcal{S}_{T_x}(\omega) |G_d(r_{T_x}, \omega)|^2, \quad (319)$$

where  $G_d(r_{T_x}, \omega)$  is the Fourier transform [51] of the Green's function of the diffusion equation expressed in (295). The probability of interference  $P_{Interf}(\omega)$  is expressed as follows:

$$P_{Interf}(\omega) = \int_{\mathcal{S}_{R_x}(\omega)}^{\infty} P_{S_Y(\omega)}(s) ds, \quad (320)$$

where  $\mathcal{S}_{R_x}(\omega)$  is the PSD of the signal  $s_{T_x}(t)$  emitted by the single transmitter, given in (318), and  $P_{S_Y(\omega)}(s)$  is the PSD of the received PSD  $S_Y(\omega)$  computed above with either the statistical-physical model, given by (314), or the simulation environment, given by the numerical results of (315).



**Figure 68.** Probability of interference according to the statistical-physical model.

In Figure 68 and Figure 69 we show the probability of interference  $P_{Interf}(\omega)$  according to the statistical-physical model and the simulation environment, respectively, for a range of frequencies  $\omega$  from  $\omega_a = 0$  Hz to  $\omega_b = 2$  Hz and for a distance  $r_{T_x}$  between the single transmitter and the receiver ranging from  $1 \mu\text{m}$  to  $2 \mu\text{m}$ . The values in Figure 68 are derived

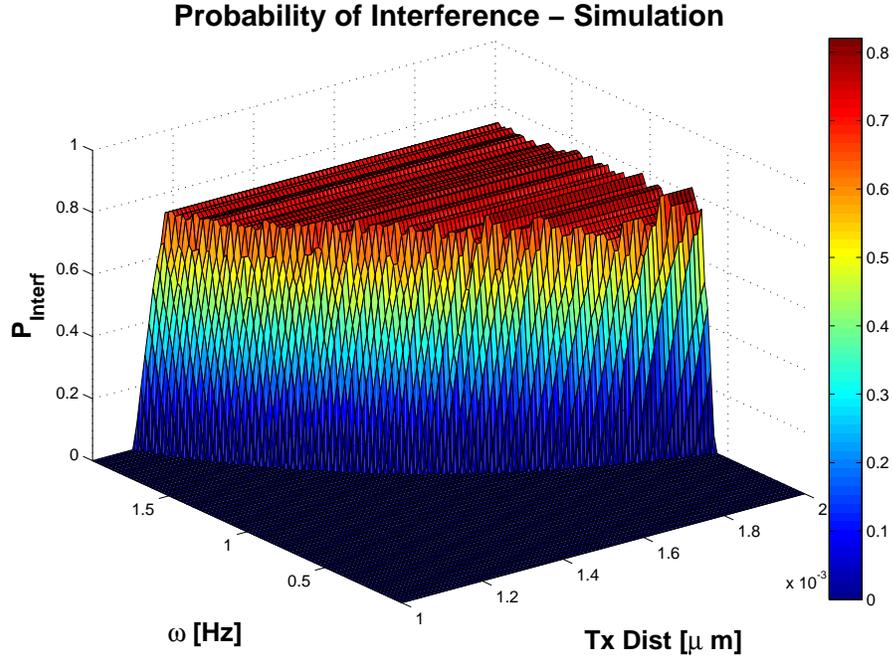


Figure 69. Probability of interference according to the simulation environment.

from the expression in (320) by using the PDF  $P_{S_Y(\omega)}(\omega)$  computed in Section 6.3.3.2, while for Figure 69 we applied the values of the PDF  $P_{S_Y(qf_s/N_s)}(s)$  computed through the simulation detailed in Section 6.3.4.1. The constant PSD of the signal  $s_{T_x}(t)$  is here set to two orders of magnitude higher than the average power of the molecular signals emitted by the interfering transmitters, namely,  $\sigma_{tx}^2 = 10^2 \sigma^2$ .

In both Figure 68 and Figure 69 we observe an almost zero probability of interference  $P_{Interf}(\omega)$  for low values of the frequency  $\omega$  and low values for the transmitter distance  $r_{T_x}$  from the receiver. As the frequency  $\omega$  and the distance  $r_{T_x}$  increase, also the probability of interference  $P_{Interf}(\omega)$  increases from zero to a maximum value. In both Figure 68 and Figure 69, values of the probability of interference higher than zero occur only for a frequency  $\omega$  higher than 0.59 Hz and a distance  $r_{T_x}$  higher than 1.1  $\mu\text{m}$ . In Figure 68 the maximum value of the probability of interference  $P_{Interf}(\omega)$  is 0.98 and it occurs for the range frequencies  $\omega$  between 0.67 Hz and 0.89 Hz and for a distance  $r_{T_x}$  higher than 1.4  $\mu\text{m}$ . The maximum value of the probability of interference  $P_{Interf}(\omega)$  in Figure 69 is around 0.82 and

it occurs for a frequency  $\omega$  around 0.73 Hz and a distance  $r_{Tx}$  higher than  $1.9 \mu\text{m}$ . The overall lower values of the simulation-based probability of interference  $P_{Interf}(\omega)$  in Figure 69 compared to the values in Figure 68 from the statistical-physical model are likely due to the limited number of interfering transmitters and iterations of the transmitter distribution realizations considered in the simulation environment, as explained in Section 6.3.4.1, while the statistical-physical model considers an infinite number of transmitters and it is based on their distribution PDF.

Different behaviors of the probability of interference  $P_{Interf}(\omega)$  for high frequencies  $\omega$  and high distances  $r_{Tx}$  are shown in Figure 68 and Figure 69. In the former, the  $P_{Interf}(\omega)$  reaches a plateau, corresponding to the aforementioned maximum value of 0.98, and then decreases as the frequency value increases from 0.89 Hz to 2 Hz, where it has a PDF value of 0.83. In the latter, after a maximum value at 0.82, and as the frequency increases from 0.73 Hz to 2 Hz, the  $P_{Interf}(\omega)$  oscillates between 0.74 and 0.72. Again, this oscillatory behavior is likely due to the limited data used in the simulation environment to compute the PDF  $P_{S_Y(\omega)}(s)$ , where we considered a limited number of interferers, within a spherical space of radius  $\rho = 19 \mu\text{m}$ , and a limited number of iterations for the realization of their location distribution.

### 6.3.5 Conclusion

In this section of the Ph.D. thesis, a statistical-physical modeling is provided for the interference in diffusion-based molecular nanonetworks when multiple transmitting nanomachines emit molecules simultaneously. Our method to characterize the interference differentiates from the previous literature since we develop a general model independent from the transmitter number, specific transmitter locations, or coding schemes. As a result of the property of the received molecular signal of being a stationary Gaussian Process (GP), the statistical-physical modeling is operated on its Power Spectral Density (PSD), for which it is possible to obtain an analytical expression of the log-characteristic function. This log-characteristic function expression ultimately leads to the estimation of the received PSD

probability distribution, which provides a complete model of the interference in diffusion-based molecular nanonetworks.

The numerical derivation of the PDF from the log-characteristic function expression of the received PSD is performed for selected values of the physical parameters of the molecular nanonetwork, such as the diffusion coefficient, the transmitter density, and the average power of molecule emissions. As apparent from the PDF of the received PSD, for low frequencies the power of the received signal tends to a uniform distribution over the range of considered values, while for higher frequencies the power tends with more probability to lower values.

The similarities of the results from the statistical-physical model with the outcomes from simulations are evaluated first in terms of received PSD, and then in terms of probability of interference. The latter comparison is based on the probability of having at the receiver a contribution from the interference whose PSD exceeds the PSD of a contribution coming from a single transmitter. In both cases, the probability of interference has very low values for frequencies lower than 0.59 Hz and a distance range lower than 1.1  $\mu\text{m}$ , while it assumes very high values otherwise. The statistical-physical model of the interference presented in this section of the Ph.D. thesis has the potential to greatly help to realization of the future diffusion-based molecular nanonetworks, in particular as a support to the design of interference mitigation techniques.

# CHAPTER 7

## A MOLECULAR COMMUNICATION SYSTEM DESIGN VIA BIOLOGICAL CIRCUITS

### 7.1 Motivation and Related Work

The design of a molecular communication system has to necessarily take into account the underlying type of nanomachines for which the system is intended. In this direction, a particularly promising type of nanomachines has been presented in Chapter 1 based on key synthetic biology techniques [12], which allow to genetically engineer cells, such as bacteria, and program functionalities in the biological environment. In particular, a synthetic biological circuit [99], or simply biological circuit, allows to program logical functions from simple controlled production of specific types of protein molecules, to complete engineered cell-to-cell interactions [100], such as a diffusion-based molecular communication system, in a similar way as it is done with electrical circuits. The focus of this chapter of the Ph.D. thesis is on the study from the communication engineering point of view of a diffusion-based molecular communication system design based on biological circuits, where both a deterministic and a stochastic modeling are applied to adapt the general results presented in Chapter 3 and Chapter 4 to this specific implementation.

A biological circuit is normally defined as a genetic regulatory network [99] embedded in a biological cell, where DNA genes are linked together by activation and repression mechanisms that regulate their expression into proteins, which are biological macromolecules. Each DNA gene contains coding sequences, which are chemical information for building proteins, and regulatory sequences, which are sites where proteins can bind and control the rate of the gene expression, either by increasing (activation) or decreasing (repression) the protein building rate. In biological circuits, genes are interconnected such as the proteins produced by one or more genes regulate the expression of one or more genes. In recent years, a great effort is being devoted to the standardization and the establishment

of catalogues of biological circuit parts [101]. By following the BioBrick™ standard [11], the units to measure the input and the output of a biological circuit are defined as Polymerases Per Second (PoPS), which correspond to the rates of the transcription process of the first and last biological circuit genes, respectively, proportional to their rate of expression. A biological circuit can process a PoPS signal as a function of the time in input by returning in output another PoPS signal as a function of the time through the aforementioned interconnection of gene regulations.

Some recent literature can be found on the analytical modeling of biological circuits, but with no specific mention to diffusion-based cell-to-cell communication through molecule exchange, for which only a biological description is provided in some specific works. Notable examples from this literature are given as follows. In [99] the genetic circuit design is introduced as an engineering discipline and the main mathematical framework for the modeling of biological circuit functions is introduced. The models of some important biological circuit patterns, called network motifs, are presented in a very complete theoretical framework in [102]. The standardization efforts of biological circuit parts are reviewed in [11], while the modeling techniques for biological circuits are discussed in [103]. The frequency domain analysis of biological circuits is presented in [104] both from a deterministic and a stochastic point of view, while the noise in biological circuit is discussed in [105]. In [106], the specific noise sources affecting cellular signaling pathways are described. Finally, the work in [12] treats engineering techniques to implement signals and sensors in bacteria through biological circuits.

In this section, deterministic and stochastic communication engineering models are presented for a biological circuit where a signal is transmitted from a PoPS input in a biological cell (transmitter cell) to a PoPS output in another biological cell (receiver cell), located at a predefined distance from the transmitter cell. This biological circuit, inspired by the cell-to-cell communication circuit sketched in [99], realizes a diffusion-based molecular communication system as defined in this Ph.D. thesis by encoding the signal to be transmitted

into signaling molecules, which propagate between the transmitter cell and the receiver cell through their diffusion in the intercellular space. In addition, the biological circuit detailed in this section is composed by the minimal subset of elements necessary to realize diffusion-based molecular communication between biological cells, and the resulting models are expected to have a general validity over other more complex implementations.

In the work presented in this section, first, a biological circuit for diffusion-based molecular communication is identified through a minimal subset of elements. Then, a mathematical model is detailed in terms of transfer functions, from which analytical expressions are derived for the attenuation and the delay experienced by an information signal through the biological circuit. Finally, the most significant noise sources within the biological circuit are identified, and statistical models for these sources are provided in terms of noise-generating random processes. For each statistical model, the impact of the generated noise is quantified through the Power Spectral Density (PSD) parameter at the output of the biological circuit. Numerical results are also provided in this chapter by applying to the developed models some biological parameters from the literature.

## **7.2 Biological Circuit Design for Diffusion-based MC**

### **7.2.1 Functional Blocks Description**

The main functional blocks of this biological circuit are shown in Figure 70, where a space is divided into the intracellular environments of a transmitter cell and a receiver cell, respectively, which are assumed chemically homogeneous, or well-stirred, and they are divided by an intercellular environment. As a consequence, in the intracellular environment the molecule concentrations are assumed homogeneous in the space, while in the intercellular environment there is in general a non-homogeneous concentration of signaling molecules, which is subject to propagation via diffusion. We assume that the intracellular space of the transmitter cell is a volume with size  $\Omega_{Tx}$ , while the intracellular space of the receiver cell is a volume with size  $\Omega_{Rx}$ . The main functional blocks of this biological circuit, shown in Figure 70, are detailed as follows:

- The **Signaling Enzyme Expression** takes place inside the transmitter cell, and it is initiated by a PoPS signal in input,  $PoPS_{in}$ , which promotes the transcription of an *enzyme coding sequence* and the translation of the contained information into a protein, denoted by  $E$  and called *enzyme* because of its specific chemical function, as explained in the following. The output of the signaling enzyme expression is the concentration of the produced enzymes, denoted by  $[E]$ .
- The **Signaling Molecule Production** is an enzymatic chemical reaction that occurs inside the transmitter cell, where the enzymes  $E$  catalyze the conversion of molecules present in the *intracellular environment*, called *substrates*, into other molecules, called products, by forming *enzyme-substrate complexes*. Among these products, the signaling molecules, denoted by  $S$ , are small organic molecules whose size allows them to cross the cell membrane and propagate through diffusion in the *intercellular environment*. The other products of the enzymatic reaction, denoted here as *subproducts*, remain in the intracellular environment and do not take part in the diffusion-based molecular communication. As a consequence, the input of the signaling molecule production is the concentration of enzymes  $[E]$ , while the output is the concentration of produced signaling molecules at the transmitter, denoted by  $[S]_{Tx}$ .
- The **Diffusion Process** realizes the propagation of the signaling molecules  $S$  in the intercellular environment, and it is the macroscopic effect of the random Brownian motion of the signaling molecules in the space. The diffusion process has the effect to propagate differences in the signaling molecule concentration from the transmitter cell to the receiver cell, where they cross the membrane and have access to the receiver intracellular environment. The input of the diffusion process is the concentration  $[S]_{Tx}$  of signaling molecules at the transmitter cell, while the output is the concentration  $[S]_{Rx}$  of signaling molecules at the receiver cell.

- The **Receptor Activator Expression** takes place inside the receiver cell, and it is initiated by an input PoPS auxiliary signal,  $PoPS_{aux}$ , which promotes the transcription of a *receptor coding sequence* and the translation of the contained information into proteins, called receptors, and denoted by  $R$ . The output of the receptor activator expression is the concentration of the produced receptors, denoted by  $[R]$ .
- The **Ligand-Receptor Binding** is a reaction that occurs inside the receiver cell, where the incoming signaling molecules  $S$  bind to the receptors  $R$  and form activator complexes, denoted by  $RS$ . The inputs of the ligand-receptor binding are the concentration of produced receptors  $[R]$  and the concentration  $[S]_{Rx}$  of signaling molecules at the receiver cell, and the output is the concentration  $[RS]$  of activator complexes.
- The **Output Transcription Activation** is initiated by the activator complexes  $RS$  upon binding to the *activator site*, where a PoPS output signal is produced according to the binding of RNA polymerase proteins, denoted as  $RNAP$ , to the promoter sequence. The inputs of the transcription activation are the concentration  $[RS]$  of activator complexes, the concentration  $[P_{Rx}]$  of promoter sequences, and the concentration  $[RNAP]$  of the RNA polymerase protein, respectively, while the output PoPS signal is denoted as  $PoPS_{out}$ .

### 7.2.2 Reaction-based Description

In the following, we provide a description of the biological circuit in terms of the chemical reactions undergoing in the aforementioned elements. This description serves to define all the chemical parameters of the biological circuit under study, and it sets the basis to build the deterministic model and the stochastic model, detailed in Section 7.3 and Section 7.4, respectively.

The **Signaling Enzyme Expression** is based on a *transcription and translation reaction*, which models the production of the  $np_E$  enzymes stimulated by the input signal

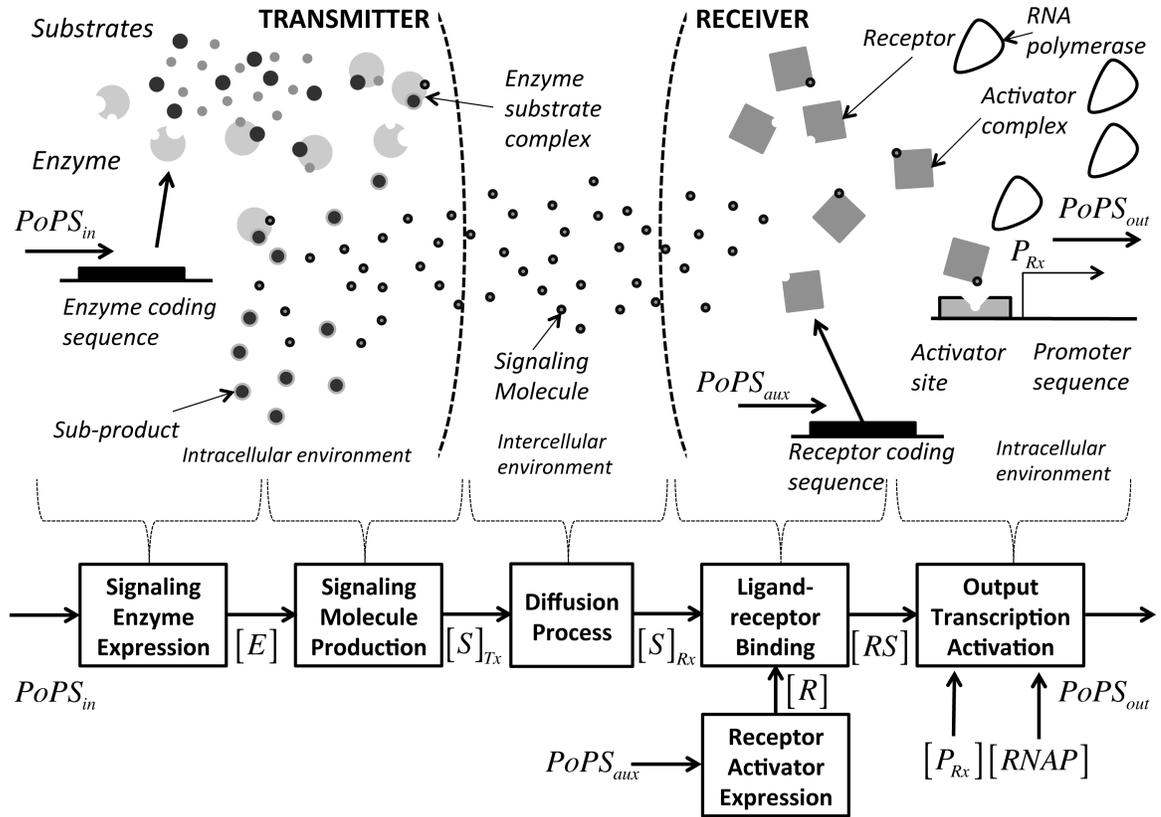


Figure 70. Main functional blocks of a biological circuit for diffusion-based molecular communication.

$PoPS_{in}$  with a rate  $k_E$ , expressed as follows:



The enzymes are also subject to degradation, with a degradation rate  $k_{dE}$ , expressed as:



The **Signaling Molecule Production** is based on an enzymatic reaction where the enzyme  $E$  and the substrates (one or more), here denoted as  $S_0$ , according to the rate  $k_{S_1}$  form a complex  $C_s$ , which can then either dissociate back into the enzyme  $E$  and the substrates  $S_0$ , with a rate  $k_{S_{-1}}$ , or evolve into the sum of the enzyme  $E$  and the signaling molecule  $S$  according to a rate  $k_{S_2}$ . This reaction is expressed as follows:

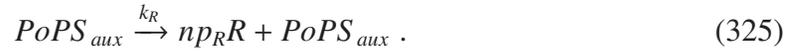


The **Diffusion Process** is based on the assumption to have a 3-dimensional intercellular space, which contains a fluidic medium and has infinite extent in all the three dimensions. The diffusion process is based on the following **Diffusion Equation** [26, 27] in the variable  $[S](r, t)$ , which is the concentration  $[S]$  of signaling molecules present at distance  $r$  from the transmitter and time instant  $t$ :

$$\frac{\partial[S](r, t)}{\partial t} = D\nabla^2[S](r, t) , \quad (324)$$

where  $\partial(\cdot)/\partial t$  and  $\nabla^2(\cdot)$  are the time first derivative and the Laplacian operator, respectively.  $D$  is the diffusion coefficient and it is considered a constant parameter within the scope of this thesis. This is in agreement with the assumption of having independent Brownian motion for every molecule in the space.

The **Receptor Activator Expression** is based on the *transcription and translation reaction* for the production of  $np_R$  receptors  $R$  stimulated by the signal  $PoPS_{aux}$  with a rate  $k_R$ , expressed as follows:



The degradation reaction of the receptors  $R$  is expressed as follows according to a degradation rate  $k_{dR}$ :

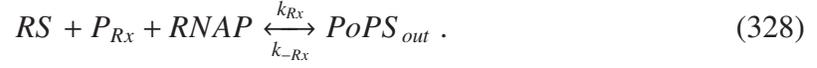


The **Ligand-Receptor Binding** is based on the binding and release reactions between receptors  $R$  and signaling molecules  $S$ . Upon binding, which occurs with a rate  $k_{RS}$ , a receptor  $R$  and a signaling molecule  $S$  form an activator complex  $RS$ , which will be the input of the next transcription activation reaction. A complex  $RS$  unbinds and releases a receptor  $R$  and a signaling molecule  $S$  according to a rate  $k_{-RS}$ . This is expressed as follows:



The **Output Transcription Activation** is based on an *open complex formation reaction*, where an activator complex  $RS$ , a promoter sequence  $P_{Rx}$ , and an RNA polymerase

*RNAP* trigger the open complex formation, quantified through the output signal  $PoPS_{out}$ , according to a rate  $k_{Rx}$ . The open complex can dissociate back into an activator complex *RS*, a promoter sequence  $P_{Rx}$ , and an RNA polymerase *RNAP* according to a rate  $k_{-Rx}$ . This has the following expression:



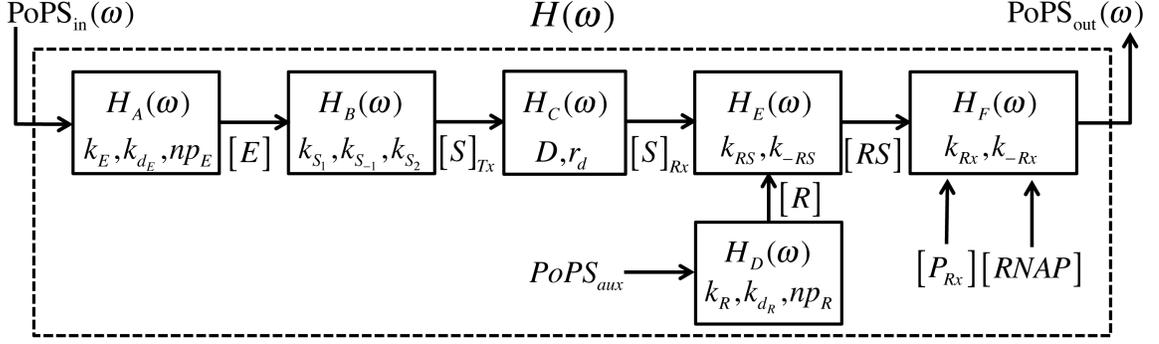
In the following, with reference to the aforementioned chemical reactions, we first detail the deterministic model of this biological circuit (Sec 7.3), which allows to derive the transfer function and, consequently, the attenuation and delay parameters for each functional block and for their overall end-to-end cascade. Subsequently, we derive a stochastic model of the biological circuit that takes into account the sources of randomness in each functional block, and we express the end-to-end signal-to-noise ratio (Sec 7.4).

### 7.3 Deterministic Model

The objective of the deterministic model is to derive the mathematical relation between the input signal  $PoPS_{in}(t)$  and the output signal  $PoPS_{out}(t)$  of the aforementioned biological circuit for diffusion-based molecular communication, where the input and output signals are function of the time  $t$ . As detailed in the following, we express this mathematical relation in terms of transfer function  $H(\omega)$ , where  $\omega$  corresponds to the frequency of the Fourier transforms [51] of the signals, namely,  $\mathbf{PoPS}_{in}(\omega)$  and  $\mathbf{PoPS}_{out}(\omega)$ , expressed as follows:

$$H(\omega) = \frac{\mathbf{PoPS}_{out}(\omega)}{\mathbf{PoPS}_{in}(\omega)} , \quad \mathbf{PoPS}_{\mathbf{i}}(\omega) = \int PoPS_{\mathbf{i}}(t)e^{-j\omega t} dt , \quad (329)$$

where  $\mathbf{i} \in \{in, out\}$ , and  $H(\omega)$  depends from all the chemical parameters defined in Section 7.2, namely, the transmitter cell volume  $\Omega_{Tx}$  and the receiver cell volume  $\Omega_{Rx}$ , the reaction rates  $k_E, k_{dE}, k_{S_1}, k_{S_{-1}}, k_{S_2}, k_R, k_{dR}, k_{RS}, k_{-RS}, k_{Rx}, k_{-Rx}$  and numbers of produced molecules  $np_E$  and  $np_R$ , the diffusion coefficient  $D$ , the auxiliary signal  $PoPS_{aux}$ , assumed constant in time, the concentration of substrates  $[S_0]$  at the transmitter cell, and the concentrations of promoter sequence  $[P_{Rx}]$  and RNA polymerase  $[RNAP]$  at the receiver cell.



**Figure 71. Decomposition of the transfer function of a biological circuit for diffusion-based molecular communication into the transfer functions of each functional block.**

As explained in the following and graphically shown in Figure 71, the linearity property of the mathematical expressions of the chemical reactions described in Section 7.2.2 allows the decomposition of the transfer function  $H(\omega)$  into the cascade of the transfer functions of each functional block, as shown in Figure 70. This decomposition has the following expression:

$$H(\omega) = H_A(\omega)H_B(\omega)H_C(\omega)H_D(\omega)H_E(\omega)H_F(\omega)[S_0]PoPS_{aux}[P_{Rx}][RNAP], \quad (330)$$

where  $H_X(\omega)$ ,  $X \in \{A, B, C, D, E, F\}$ , are the transfer functions of each functional block, as function of the frequency  $\omega$ , detailed in the following. The parameters  $[S_0]$ ,  $PoPS_{aux}$ ,  $[P_{Rx}]$  and  $[RNAP]$  are the concentration of substrates at the transmitter cell, and the auxiliary input signal, the concentration of promoter sequences and the concentration of RNA polymerase at the receiver cell, respectively, assumed constant in time for the scope of this thesis.

In the following, we analytically derive the transfer function of each functional block shown in Figure 71 from the reaction-based description provided in Section 7.2.2. Subsequently, we provide an approximation  $\hat{H}(\omega)$  of the transfer function of the biological circuit through considerations on the differences in the time scales of the chemical reactions of different functional blocks. Finally, starting from the expression of the approximated transfer function  $\hat{H}(\omega)$  of the biological circuit, we provide analytical expressions for the attenuation and delay experienced by an information signal through the biological circuit.

### 7.3.1 Functional Block Transfer Functions

The transfer function of each functional block, with reference to Figure 71, is analytically derived by applying the Classical Chemical Kinetic (CCK) modeling [107] to the reaction-based description provided for each block in Section 7.2.2.

The CCK model of the **Signaling Enzyme Expression** is expressed through the following Reaction-Rate Equation (RRE), which analytically models the chemical reactions in (321) and (322):

$$\frac{d[E](t)}{dt} = np_E k_E PoPS_{in}(t) - k_{d_E}[E](t), \quad (331)$$

where  $[E](t)$  and  $PoPS_{in}(t)$  are the concentration of produced enzymes inside the transmitter cell and the input signal, respectively, as functions of the time  $t$ . By applying the Fourier transform [51] to (331), we obtain the following:

$$j\omega[\mathbf{E}](\omega) = np_E k_E \mathbf{PoPS}_{in}(\omega) + k_{d_E}[\mathbf{E}](\omega), \quad (332)$$

As a consequence, the transfer function  $H_A(\omega)$  of the signaling enzyme expression functional block is derived by solving (332) with respect to the concentration of produced enzymes  $[\mathbf{E}](\omega)$  as function of the  $\mathbf{PoPS}_{in}(\omega)$ , expressed as

$$H_A(\omega) = \frac{np_E k_E}{j\omega + k_{d_E}}, \quad (333)$$

where  $np_E$ ,  $k_E$ , and  $k_{d_E}$  are the number of enzymes produced per reaction, the enzyme expression rate, and the enzyme degradation rate, respectively.

The **Signaling Molecule Production** is expressed through the following two RREs, which analytically model the chemical reactions in (323):

$$\begin{aligned} \frac{d[C_S](t)}{dt} &= k_{S_1}[E](t)[S_0] - k_{S_{-1}}[C_S](t) - k_{S_2}[C_S](t) \\ \frac{d[S]_{Tx}(t)}{dt} &= k_{S_2}[C_S](t), \end{aligned} \quad (334)$$

where  $[C_S](t)$ ,  $[E](t)$ , and  $[S]_{Tx}(t)$  are the concentration of formed complexes, produced enzymes and produced signaling molecules inside the transmitter cell, respectively, as functions of the time  $t$ , and  $[S_0]$  is the concentration of the substrates, assumed constant in time.

By applying the Fourier transform [51] to (334) and by substituting the first expression in the second expression, we obtain

$$j\omega[\mathbf{S}]_{\text{Tx}}(\omega) = \frac{k_{S_1}}{k_{S_{-1}} + k_{S_2} + j\omega} [S_0][\mathbf{E}](\omega) . \quad (335)$$

Starting from (335), by expressing the concentration of produced signaling molecules  $[\mathbf{S}]_{\text{Tx}}(\omega)$  as function of the produced enzymes  $[\mathbf{E}](\omega)$  and the frequency  $\omega$ , we derive the expression of the transfer function  $H_B(\omega)$  of the signaling molecule production functional block as follows:

$$H_B(\omega) = \frac{k_{S_1}}{\omega\{j(k_{S_{-1}} + k_{S_2}) - \omega\}} , \quad (336)$$

where  $k_{S_1}$ ,  $k_{S_{-1}}$ , and  $k_{S_2}$  are the complex formation rate, the complex dissociation rate, and the signaling molecule production rate, respectively.

The **Diffusion Process** functional block is expressed through the Inhomogeneous Diffusion Equation, which is based on the diffusion equation expression in (324), as follows:

$$\frac{\partial[S](r, t)}{\partial t} = D\nabla^2[S](r, t) + \frac{d[S]_{\text{Tx}}(t)}{dt}\delta(r) , \quad (337)$$

where  $[S](r, t)$  and  $\frac{d[S]_{\text{Tx}}(t)}{dt}$  are the concentration of signaling molecules present at distance  $r$  from the transmitter and the first time derivative of the concentration of signaling molecules at the transmitter, respectively, as function of the time  $t$ .  $\delta(r)$  is a Dirac delta centered at the transmitter location and  $D$  is the diffusion coefficient. The solution of (337) in terms of Fourier transform [51] of the concentration of signaling molecules  $[\mathbf{S}]_{\text{Rx}}(\omega)$  at the receiver, located at a distance  $r_{\text{Rx}}$  from the transmitter, as function of the produced signaling molecules  $[\mathbf{S}]_{\text{Tx}}(\omega)$ , is as follows [86]:

$$[\mathbf{S}]_{\text{Rx}}(\omega) = \frac{e^{-(1+j)\sqrt{\frac{\omega}{2D}}r_{\text{Rx}}}}{\pi D r_{\text{Rx}}} j\omega[\mathbf{S}]_{\text{Tx}}(\omega) . \quad (338)$$

As a consequence, the expression of the transfer function  $H_C(\omega)$  of the diffusion process functional block is as follows:

$$H_C(\omega) = j\omega \frac{e^{-(1+j)\sqrt{\frac{\omega}{2D}}r_{\text{Rx}}}}{\pi D r_{\text{Rx}}} , \quad (339)$$

where  $D$  and  $r_{Rx}$  are the diffusion coefficient and the distance of the receiver from the transmitter, respectively.

The CCK model of the **Receptor Activator Expression** is expressed through the following RRE, which analytically models the chemical reactions in (325) and (322):

$$\frac{d[R](t)}{dt} = np_R k_R PoPS_{aux} - k_{dR} [R](t) , \quad (340)$$

where  $[R](t)$  and  $PoPS_{aux}$  are the concentration of receptors inside the receiver cell as functions of the time  $t$  and the PoPS signal that controls the receptor expression, respectively. Since we assume that the auxiliary input signal  $PoPS_{aux}$  is constant in time, the resulting concentration of receptors inside the receiver cell is also constant in time. By solving (340), the expression of the transfer function  $H_D(\omega)$  of the diffusion process functional block is as follows:

$$H_D(\omega) = \frac{np_R k_R}{k_{dR}} , \quad (341)$$

where  $np_R$ ,  $k_R$ , and  $k_{dR}$  are the number of receptors produced per reaction, the receptor expression rate, and the receptor degradation rate, respectively.

The **Ligand-Receptor Binding** has a RRE CCK model which derives from the chemical reaction expression in (327), and it is as follows:

$$\frac{d[RS](t)}{dt} = k_{RS} [R](t)[S]_{Rx}(t) - k_{-RS} [RS](t) , \quad (342)$$

where  $[RS](t)$ ,  $[R](t)$  and  $[S]_{Rx}(t)$  are the concentration of activator complexes, receptors and signaling molecules inside the receiver cell, respectively, as functions of the time  $t$ . By applying the Fourier transform [51] to (342), we express the concentration of activator complexes  $[\mathbf{RS}](\omega)$  as function of  $[\mathbf{R}](\omega)$ ,  $[\mathbf{S}](\omega)$ , and the frequency  $\omega$  as follows:

$$j\omega[\mathbf{RS}](\omega) = ([\mathbf{R}](\omega) * [\mathbf{S}](\omega)) k_{RS} - k_{-RS} [\mathbf{RS}](\omega) , \quad (343)$$

where  $.*.$  is the convolution operator [51]. As explained above, we assume a constant auxiliary input signal  $PoPS_{aux}$  is constant in time, which results in a constant concentration of

receptors inside the receiver cell. As a consequence, the expression of the transfer function  $H_E(\omega)$  of the ligand-receptor binding functional block is as follows:

$$H_E(\omega) = \frac{k_{RS}}{j\omega + k_{-RS}}, \quad (344)$$

where  $k_{RS}$  and  $k_{-RS}$  are the ligand-receptor binding and release rates, respectively.

The CCK model of the **Output Transcription Activation** functional block is expressed through the RRE, which is derived from the description of the chemical reaction in (328). This RRE has the following expression:

$$\frac{dPoPS_{out}(t)}{dt} = k_{Rx}[P_{Rx}][RNAP][RS](t) - k_{-Rx}[RS](t), \quad (345)$$

where  $PoPS_{out}(t)$  and  $[RS](t)$  are the biological circuit output PoPS signal and the concentration of activator complexes inside the receiver cell, respectively, as functions of the time  $t$ , and  $[P_{Rx}]$  and  $[RNAP]$  are the concentrations of promoter sequences and RNA polymerase at the receiver cell, respectively, assumed constant in time. The expression in (342) is solved in the same way as done for the signaling enzyme expression functional block in (332). Finally, the expression of the transfer function  $H_F(\omega)$  of the output transcription activation functional block is as follows:

$$H_F(\omega) = \frac{k_{Rx}}{j\omega + k_{-Rx}}, \quad (346)$$

where  $k_{Rx}$ , and  $k_{-Rx}$  are the open complex formation and dissociation rates, respectively.

Since the RRE expression of the functional blocks in (331), (334), (337), (340), (342), and (342) are Ordinary Differential Equations (ODE), they represent Linear Time-Invariant systems, whose transfer function solutions can be combined through the formula in (330) to derive the transfer function  $H(\omega)$  of a biological circuit for diffusion-based molecular communication, expressed as

$$H(\omega) = \frac{np_E k_E}{j\omega + k_{dE}} \frac{k_{S_1} [S_0]}{\omega \{j(k_{S_{-1}} + k_{S_2}) - \omega\}} j\omega \frac{e^{-(1+j)\sqrt{\frac{\omega}{2D}}r_{Rx}}}{\pi D r_{Rx}} \frac{np_R k_R}{k_{dR}} PoPS_{aux} \cdot \frac{k_{RS}}{j\omega + k_{-RS}} \frac{k_{Rx}}{j\omega + k_{-Rx}} [P_{Rx}][RNAP], \quad (347)$$

where  $np_E$ ,  $k_E$ , and  $k_{d_E}$  are the number of enzymes produced per reaction, the enzyme expression rate, and the enzyme degradation rate, respectively,  $k_{S_1}$ ,  $k_{S_{-1}}$ , and  $k_{S_2}$  are the complex formation rate, the complex dissociation rate, and the signaling molecule production rate, respectively,  $D$  and  $r_{R_x}$  are the diffusion coefficient and the distance of the receiver from the transmitter, respectively,  $np_R$ ,  $k_R$ , and  $k_{d_R}$  are the number of receptors produced per reaction, the receptor expression rate, and the receptor degradation rate, respectively,  $PoPS_{aux}$  is the auxiliary input signal, assumed constant in time,  $k_{RS}$  and  $k_{-RS}$  are the ligand-receptor binding and release rates, respectively, and  $k_{R_x}$ , and  $k_{-R_x}$  are the open complex formation and dissociation rates, respectively.  $[S_0]$ ,  $[P_{R_x}]$  and  $[RNAP]$  are the concentrations of substrates at the transmitter cell, promoter sequences and RNA polymerase at the receiver cell, respectively, assumed constant in time.

### 7.3.2 Time Scale Approximation

According to [102], the chemical reactions involved in biological circuits have different time scales. In particular, the chemical reactions described in Section 7.2.2, and modeled through the transfer function expressions in Section 7.3.1, occur at significantly different speeds. As experimentally demonstrated in [102], the chemical reactions where a protein is expressed from the DNA coding sequence and accumulates/propagates in the space, such as the signaling enzyme expression, the receptor activator expression, and the diffusion process, are significantly slower than the reactions between two or more molecules for the formation of complexes, such as in the signaling molecule production, the ligand-receptor binding and the output transcription activation. Therefore, the former reactions dominate the dynamic behavior of the circuit, and the transfer functions of the latter reactions can be approximated with their steady state versions, as shown in Figure 72 and analytically derived in the following.

As a result, we define  $\hat{H}(\omega)$  as the approximate transfer function of the biological circuit, derived through considerations on the chemical reaction time scales. The transfer

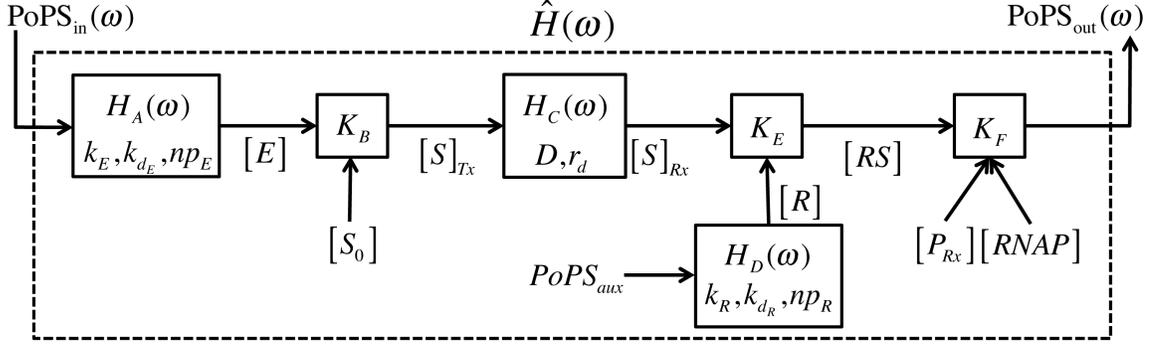


Figure 72. Approximation of the decomposition of the transfer function of a biological circuit for diffusion-based molecular communication.

function  $\hat{H}(\omega)$  is expressed as follows:

$$\hat{H}(\omega) = H_A(\omega)K_B(\omega)[S_0]H_C(\omega)H_D(\omega)PoPS_{aux}K_EK_F[P_{Rx}][RNAP], \quad (348)$$

where  $H_A(\omega)$ ,  $H_C(\omega)$ , and  $H_D(\omega)$  are the transfer functions of the signaling enzyme expression, the receptor activator expression, and the diffusion process, respectively,  $[S_0]$ ,  $PoPS_{aux}$ ,  $[P_{Rx}]$  and  $[RNAP]$  are the concentrations of substrates at the transmitter cell, the auxiliary input, the concentration promoter sequences and the concentration RNA polymerase at the receiver cell, respectively, assumed constant in time,  $K_B(\omega)$  is the steady state transfer function of the signaling molecule production,  $K_E$ , and  $K_F$  are the steady state approximations to constant values of the transfer functions of the ligand-receptor binding and the output transcription activation, respectively.

The steady state approximation  $K_B(\omega)$  of the **Signaling Molecule Production** functional block is computed by setting in (334) the first time derivative  $d[C_S](t)/dt$  in the concentration of formed complexes to 0. The solution to (334) becomes as follows:

$$\frac{d[C_S](t)}{dt} = 0 \rightarrow [S]_{Tx}(\omega) = \frac{k_{S_1}}{j\omega(k_{S_{-1}} + k_{S_2})}[S_0][E](\omega). \quad (349)$$

The steady state transfer function  $K_B(\omega)$  of the signaling molecule production is therefore given by

$$K_B(\omega) = \frac{k_{S_1}}{j\omega(k_{S_{-1}} + k_{S_2})}, \quad (350)$$

where  $k_{S_1}$ ,  $k_{S_{-1}}$ , and  $k_{S_2}$  are the complex formation rate, the complex dissociation rate, and the signaling molecule production rate, respectively.

The steady state approximations of the **Ligand-Receptor Binding** and the **Output Transcription Activation** functional blocks to the constant values  $K_E$  and  $K_F$  result from computing the transfer functions  $H_E(\omega)$  and  $H_F(\omega)$  for a value of the frequency  $\omega = 0$ , expressed as

$$K_E = \frac{k_{RS}}{k_{-RS}}, \quad K_F = \frac{k_{RX}}{k_{-RX}}, \quad (351)$$

which correspond to the solution of (342) and (345) when we set to 0 the time first derivative  $d[RS](t)/dt$  in the concentration of activator complexes and the time first derivative  $dPoPS_{out}(t)/dt$  in the biological circuit output PoPS signal.

The approximate transfer function  $\hat{H}(\omega)$  of the biological circuit, derived through the steady state approximations, has the following expression:

$$\hat{H}(\omega) = K_{\hat{H}} \frac{e^{-(1+j)\sqrt{\frac{\omega}{2D}}r_{Rx}}}{j\omega(j\omega + k_{d_E})}, \quad (352)$$

where the constant  $K_{\hat{H}}$  is as follows:

$$K_{\hat{H}} = \frac{np_E k_E k_{S_1} [S_0] np_R k_R PoPS_{aux} k_{RS} k_{Rx} [P_{Rx}] [RNAP]}{(k_{S_{-1}} + k_{S_2}) \pi D r_{Rx} k_{d_R} k_{-RS} k_{-Rx}}, \quad (353)$$

where all the parameters are the same as in (347).

### 7.3.3 Attenuation and Delay Expressions

The attenuation and delay experienced by a signal through the biological circuit are analytically derived from the approximate transfer function  $\hat{H}(\omega)$  expressed in (352),

The attenuation  $\alpha(\omega)$ , as function of the frequency  $\omega$ , is computed through the reciprocal of the absolute value of the approximate transfer function  $\hat{H}(\omega)$  in (352), which has the following expression:

$$\alpha(\omega) = \frac{1}{|\hat{H}(\omega)|} = \frac{\omega \sqrt{\omega^2 + k_{d_E}^2}}{K_{\hat{H}} e^{-\sqrt{\frac{\omega}{2D}}r_{Rx}}}, \quad (354)$$

where  $k_{d_E}$  is the enzyme degradation rate,  $D$  is the diffusion coefficient,  $r_{Rx}$  is the distance of the receiver from the transmitter, and  $K_{\hat{H}}$  is given in (353).

The delay  $\Delta(\omega)$ , as function of the frequency  $\omega$ , is computed as the frequency first derivative of the phase  $\phi_{\hat{H}}(\omega)$  of the approximate transfer function  $\hat{H}(\omega)$  in (352). The phase  $\phi_{\hat{H}}(\omega)$  has the following expression

$$\phi_{\hat{H}}(\omega) = \arctan\left(\frac{k_{dE}}{\omega}\right) - \sqrt{\frac{\omega}{2D}} r_{Rx}, \quad (355)$$

As a consequence, the delay  $\Delta(\omega)$  is expressed as

$$\Delta(\omega) = -\frac{d\phi_{\hat{H}}(\omega)}{d\omega} = \frac{r_{Rx}\omega^2 + 2k_{dE}\sqrt{2D\omega} + r_{Rx}k_{dE}^2}{2\sqrt{2D\omega}(\omega^2 + k_{dE}^2)}, \quad (356)$$

where  $r_{Rx}$ ,  $k_{dE}$ , and  $D$  are the distance of the receiver from the transmitter, the enzyme degradation rate, and the diffusion coefficient, respectively.

## 7.4 Stochastic Model

In this section, we identify the most significant noise sources within a biological circuit for diffusion-based molecular communication, and we provide a statistical model for each of these sources in terms of noise-generating random processes. For each statistical model, we quantify the impact of the generated noise on the biological circuit through the Power Spectral Density (PSD) parameter at the output of the biological circuit.

As we will prove next, the physical characteristics of the noise sources within the biological circuit allow to derive, given a constant value  $PoPS_{in}$  of the input signal, a frequency dependent PSD  $\mathcal{S}_{N_{out}}(\omega)$  of the noise generated at the output of the biological circuit. This is expressed as follows [66]:

$$\mathcal{S}_{N_{out}}(\omega) = \text{E} \left[ |\mathcal{F}\{n(t)\}|^2 \right], \quad (357)$$

where the  $n(t)$  is the random process as function of the time  $t$  that represent the noise at the output of the biological circuit

The computation of the PSD  $\mathcal{S}_{N_{out}}(\omega)$  is realized through the statistical modeling of the noise sources in the biological circuit for diffusion-based molecular communication, which arise from the discrete nature of the molecules and their interactions. In the following, we

provide in Section 7.4.1 an analysis of the most significant noise sources that stem from the randomness inherently present in the chemical reactions underlying a biological circuit, and we derive closed-form analytical expressions of upper bounds to the PSD  $\mathcal{S}_{N_{Tx}}(\omega)$  of the noise generated at the transmitter cell and the PSD  $\mathcal{S}_{N_{Rx}}(\omega)$  of the noise generated at the receiver cell, respectively. Second, in Section 7.4.2, we provide a closed-form analytical expression of an upper bound to the PSD  $\mathcal{S}_{N_{Diff}}(\omega)$  of the noise coming from the molecule diffusion process, by stemming from our previous results in Section 4.3. Finally, in Section 7.4.3, we derive the final expression of the PSD  $\mathcal{S}_{N_{out}}(\omega)$ .

### 7.4.1 Noise in Chemical Reactions

According to the  $\tau$ -leaping approximate stochastic method [103], a set of chemical reactions can be statistically modeled as follows. Let the vector  $\mathbf{X}(t) = [X_1(t), X_2(t), \dots, X_N(t)]'$  contain the number of molecules for each species  $n$  out of  $N$  present in a space, supposed to be well-stirred and with a volume  $\Omega$ . The value  $\mathbf{X}(t + \tau)$  at time  $t + \tau$  is given by the value  $\mathbf{X}(t)$  at time  $t$  plus the sum over the  $M$  reactions undergoing among the  $N$  molecule species of the parameter  $P_m$  multiplied by the vector  $\mathbf{V}_m$ , defined in the following. This is expressed as follows:

$$\mathbf{X}(t + \tau) = \mathbf{X}(t) + \sum_{m=1}^M P_m \mathbf{V}_m, \quad (358)$$

where  $P_m$  is a realization of the following Poisson counting process:

$$P_m \sim \text{Pois}(a_m(\mathbf{X}(t)), \tau), \quad (359)$$

where  $a_m(\mathbf{X}(t))$  is the rate parameter of the Poisson counting process, and  $\tau$  is the time interval in which the counting process realization is computed. The parameter  $a_m(\mathbf{X}(t))$  is called propensity function of the chemical reaction  $m$ , and it corresponds to the probability that the reaction  $m$  occurs in an infinitesimal time interval after time  $t$ , given the values in  $\mathbf{X}(t)$  of the number of molecules for each species at time  $t$ . The propensity function  $a_m(\mathbf{X}(t))$  for a chemical reaction  $m$  of the type considered in this chapter of the Ph.D. thesis

is computed as follows [99]:

$$a_m(\mathbf{X}(t)) = c_m \prod_{X_i \in \mathcal{R}_m} X_i(t) \quad (360)$$

where  $\mathcal{R}_m$  is the set of reactant species for the chemical reaction  $m$ , and  $c_m$  is computed as follows:

$$c_m = \frac{k_m}{\Omega^{|\mathcal{R}_m|-1}} \quad (361)$$

$k_m$  is the rate and  $|\mathcal{R}_m|$  is the number of reactant species of the chemical reaction  $m$ , and  $\Omega$  is the volume of the space that contains these reactants species.

The value of the time interval  $\tau$  is chosen so as it is long enough to allow each reaction from the set of  $M$  reactions to occur more than once, but short enough to consider the propensity functions  $a_m(\mathbf{X}(t))$  constant within  $\tau$  with negligible error [103]. The vector  $\mathbf{V}_m$ , termed stoichiometric vector, is equal to the changes in the number of molecules for each species that the reaction  $m$  operates when it occurs.

According to the Chemical Langevin Equation (CLE) [94] formulation of the stochastic model expressed in (358), we can rewrite the stochastic model through the RREs expressed in Section 7.3.1 by adding the noise contribution as a Gaussian Process [94], as detailed in the following. This latter approach ultimately leads to the expression of the PSD  $\mathcal{S}_{N_i}(\omega)$  of a noise source  $i$  in the biological circuit.

The CLE is expressed as follows [94]:

$$\frac{dX_i(t)}{dt} = \sum_{m=1}^M V_{im} a_m(\mathbf{X}(t)) + \sum_{m=1}^M V_{im} \sqrt{a_m(\mathbf{X}(t))} \Gamma_m(t), \quad (362)$$

where  $X_i(t)$  and  $V_{im}$  are the  $i$ -th components of the number of molecules vector  $\mathbf{X}(t)$  and the stoichiometric vector  $\mathbf{V}_m$ , respectively,  $a_m(\mathbf{X}(t))$  is the propensity function for the reaction  $m$ , computed through (360), and  $\Gamma_m(t)$  is a white noise process for the reaction  $m$ , statistically independent from the white noise processes of other reactions.

The PSD  $\mathcal{S}_{N_i}(\omega)$  of the noise source affecting the  $i$  component  $X_i(t)$  of the number of molecules vector  $\mathbf{X}(t)$  is computed through the expression in (357) where  $n(t)$  is substituted

with  $V_{im} \sqrt{a_m(\mathbf{X}(t))} \Gamma_m(t)$ :

$$\mathcal{S}_{N_i}(\omega) = \mathbb{E} \left[ \left| \mathcal{F} \left\{ V_{im} \sqrt{a_m(\mathbf{X}(t))} \Gamma_m(t) \right\} \right|^2 \right], \quad (363)$$

where  $\mathbb{E}[\cdot]$  and  $\mathcal{F}\{\cdot\}$  denote the average value and the Fourier transform [51] operators, respectively.

We apply the following assumptions which allow to derive an analytical expression of an upper bound  $\tilde{\mathcal{S}}_{N_i}(\omega)$  to the PSD  $\mathcal{S}_{N_i}(\omega)$  of the noise source  $i$ :

- We consider the following inequality

$$\left| \mathcal{F} \left\{ V_{im} \sqrt{a_m(\mathbf{X}(t))} \Gamma_m(t) \right\} \right|^2 \leq \left| \mathcal{F} \left\{ V_{im} \sqrt{a_m(\mathbf{X}(t))} \right\} \right|^2 * \left| \mathcal{F} \left\{ \Gamma_m(t) \right\} \right|^2, \quad (364)$$

which is derived by applying the formula of the Fourier transform of the product of two functions [51] and the Cauchy-Schwarz inequality [97] to the convolution integral, denoted by the operator  $*$ .

- We consider that  $\Gamma_m(t)$  is a white noise process. By definition, we have the following results:

$$\left| \mathcal{F} \left\{ \Gamma_m(t) \right\} \right|^2 = 1. \quad (365)$$

- We consider  $a_m(\mathbf{X}(t))$  to be a constant parameter in time. This is an approximation, since even in the presence of a constant input signal  $PoPS_{in}$  to the biological circuit, in general the number of molecules contained in the vector  $\mathbf{X}(t)$  fluctuates as function of the time as a consequence of the noise coming from previous sources with respect to the noise source  $i$  in the cascade of functional blocks in the biological circuit. Since we aim at the computation of an upper bound to the PSD  $\mathcal{S}_{N_i}(\omega)$  of the noise source  $i$ , we overestimate the noise contributions by considering all the noise sources within the biological circuit as uncorrelated through this aforementioned assumption. This assumption allows to write the following expression:

$$\left| \mathcal{F} \left\{ V_{im} \sqrt{a_m(\mathbf{X}(t))} \right\} \right|^2 = V_{im}^2 a_m(\mathbf{X}(t)) \delta(\omega), \quad (366)$$

where  $\delta(\omega)$  is a Dirac delta as function of the frequency  $\omega$ .

As a consequence of these assumptions, the upper bound  $\tilde{\mathcal{S}}_{N_i}(\omega)$  to PSD  $\mathcal{S}_{N_i}(\omega)$  of the noise source  $i$  has the following expression:

$$\tilde{\mathcal{S}}_{N_i}(\omega) = V_{im}^2 a_m(\mathbf{X}(t)) , \quad (367)$$

In the following, we express the  $\tau$ -leaping statistical model in (358) for the two separate sets of chemical reactions present in the biological circuit for diffusion-based MC, namely, the chemical reactions in the transmitter cell and the chemical reactions in the receiver cell, respectively. As detailed in the following, the statistical models of two noise sources within the biological circuit are in agreement with the particle counting noise and the ligand-receptor-binding noise that we detailed in our previous work included in Section 4.3 and Section 4.5, respectively. Subsequently, for each noise source we derive the CLE formulation and the expression of the upper bound to the source PSD. Finally, we compute the PSD  $\mathcal{S}_{N_{Tx}}(\omega)$  of the noise generated at the transmitter cell and the PSD  $\mathcal{S}_{N_{Rx}}(\omega)$  of the noise generated at the receiver cell, respectively.

#### 7.4.1.1 Transmitter Cell

In the transmitter cell, we take into account the chemical reactions in (321), (322), and (323), which involve the input signal  $PoPS_{in}$ , the number of signaling enzymes  $E$ , the number of complexes  $C_s$ , and the number of produced signaling molecules  $S$ , as functions of the time  $t$ . We exclude from this model the number of substrates  $S_0$  at the transmitter cell, since it is assumed constant in time. The  $\tau$ -leaping statistical model is expressed for the transmitter cell through (358) where:

$$\mathbf{X}(t) = \left[ S \quad C_s \quad E \quad PoPS_{in} \right]' (t) , \quad (368)$$

the number of reactions  $M$  is equal to 5, and the propensity functions  $a_m(\mathbf{X}(t))$  and the stoichiometric vectors  $\mathbf{V}_m$ ,  $m = 1, \dots, 5$ , are computed for the transmitter cell reactions through the expressions in (360) and (361) as follows:

- In the **Signaling Enzyme Expression** functional block, the propensity function  $a_1(\mathbf{X}(t))$  and the stoichiometric vector  $\mathbf{V}_1$  are computed for the transcription and translation

reaction in (321) as follows:

$$a_1(\mathbf{X}(t)) = k_E P_o P S_{in}(t), \quad \mathbf{V}_1 = \begin{bmatrix} 0 & 0 & n p_E & 0 \end{bmatrix}'. \quad (369)$$

For the enzyme degradation reaction in (322), the propensity function  $a_2(\mathbf{X}(t))$  and the stoichiometric vector  $\mathbf{V}_2$  are as follows:

$$a_2(\mathbf{X}(t)) = k_{d_E} E(t), \quad \mathbf{V}_2 = \begin{bmatrix} 0 & 0 & -1 & 0 \end{bmatrix}'. \quad (370)$$

- In the **Signaling Molecule Production** functional block, the propensity function  $a_3(\mathbf{X}(t))$  and the stoichiometric vector  $\mathbf{V}_3$  for the formation of complexes  $C_S$  in (323) are expressed as

$$a_3(\mathbf{X}(t)) = \frac{k_{S_1}}{\Omega_{T_x}} E(t) S_0, \quad \mathbf{V}_3 = \begin{bmatrix} 0 & 1 & -1 & 0 \end{bmatrix}'. \quad (371)$$

where  $\Omega_{T_x}$  is the volume of the transmitter cell. For the dissociation of complexes reaction in (323), the propensity function  $a_4(\mathbf{X}(t))$  and the stoichiometric vector  $\mathbf{V}_4$  are as follows:

$$a_4(\mathbf{X}(t)) = k_{S_{-1}} C_S(t), \quad \mathbf{V}_4 = \begin{bmatrix} 0 & -1 & 1 & 0 \end{bmatrix}'. \quad (372)$$

For the production of signaling molecules from complexes reaction in (323), the propensity function  $a_5(\mathbf{X}(t))$  and the stoichiometric vector  $\mathbf{V}_5$  are as follows:

$$a_5(\mathbf{X}(t)) = k_{S_2} C_S(t), \quad \mathbf{V}_5 = \begin{bmatrix} 1 & -1 & 0 & 0 \end{bmatrix}'. \quad (373)$$

By isolating the contribution of the noise coming from the complex  $C_S$  formation in the signaling molecule production functional block, we can derive the following expression:

$$\begin{aligned} C_S(t + \tau) &= C_S(t) + P_3 - P_4, \quad P_3 \sim \text{Poiss}(a_3(\mathbf{X}(t)), \tau) \\ &P_4 \sim \text{Poiss}(a_4(\mathbf{X}(t)), \tau), \end{aligned} \quad (374)$$

where, according to the RRE expression in (334), the number of signaling complexes  $C_S$  is proportional to the time first derivative  $dS_{T_x}(t)/dt$  in the number of produced signaling

molecules at the transmitter cell. The expression in (374), where the production rate of signaling molecules is modeled as a double inhomogeneous Poisson process, is **in agreement with the stochastic model of the particle sampling noise that we derived in our previous work included in Section 4.5.**

The CLE formulation of the noise generated by the **Signaling Enzyme Expression** functional block is derived from (362) and (370) as follows:

$$\frac{dE(t)}{dt} = np_E k_E PoPS_{in} - k_{dE} E(t) + 2np_E \sqrt{k_E PoPS_{in}} \Gamma_1(t), \quad (375)$$

where  $\Gamma_1(t)$  is a white Gaussian noise [66], and the coefficient 2 takes into account the steady state noise generation from both the transcription and translation reaction and the enzyme degradation reaction, in agreement with [105]. The upper-bound PSD  $\tilde{S}_{N_1}(\omega)$  of the noise source of the signaling enzyme expression, as defined in (367), is equal to the following:

$$\tilde{S}_{N_1}(\omega) = 4np_E^2 k_E PoPS_{in}, \quad (376)$$

where  $np_E$  is the number of enzymes produced per reaction,  $k_E$  is the enzyme expression rate and  $PoPS_{in}$  is the constant input signal, respectively. Inspired by the theoretical framework in [105], we compute the PSD  $S_{N_A}(\omega)$  of the noise at the output of the signaling enzyme expression functional block as follows:

$$S_{N_A}(\omega) = \frac{\tilde{S}_{N_1}(\omega)}{\omega^2 + k_{dE}^2} = 4 \frac{PoPS_{in}}{k_E} |H_A(\omega)|^2, \quad (377)$$

where we applied the expression of  $H_A(\omega)$  defined in (333).

The CLE formulation of the noise generated by the **Signaling Molecule Production** functional block is derived from (362) and (371), and through the Michaelis-Menten approximation of the RRE of the enzymatic reaction, as explained in [2]. The CLE expression is as follows:

$$\frac{dS_{Tx}(t)}{dt} = \frac{k_{S_1}[S_0]}{[S_0] + K_M} E(t) + 2 \sqrt{\frac{k_{S_1}[S_0]}{[S_0] + K_M}} E(t) \Gamma_2(t), \quad (378)$$

where  $[S_0] = S_0/\Omega_{Tx}$ ,  $\Gamma_2(t)$  is a white Gaussian noise [66], and the coefficient 2 takes into account the steady state noise generation from both the signaling complex formation and

the signaling complex dissociation reaction, respectively, in agreement with [105]. The coefficient  $K_M$  is the Michaelis-Menten constant [2], which has the following value:

$$K_M = \frac{k_{S_{-1}} + k_{S_2}}{k_{S_1}}, \quad (379)$$

where  $k_{S_1}$ ,  $k_{S_{-1}}$ , and  $k_{S_2}$  are the complex formation rate, the complex dissociation rate, and the signaling molecule production rate, respectively. The upper-bound PSD  $\tilde{\mathcal{S}}_{N_2}(\omega)$  of the noise source of the signaling enzyme expression, as defined in (367), is equal to the following:

$$\tilde{\mathcal{S}}_{N_2}(\omega) = 4 \frac{k_{S_1}[S_0]}{[S_0] + K_M} PoPS_{in} H_A(0), \quad (380)$$

where we computed the value of  $E(t)$  by multiplying the constant input signal  $PoPS_{in}$  by the transfer function  $H_A(\omega)$  of the signaling enzyme expression functional block computed at frequency  $\omega = 0$ . Inspired by the theoretical framework in [105], we compute the PSD  $\mathcal{S}_{N_B}(\omega)$  of the noise at the output of the signaling molecule production functional block as follows:

$$\mathcal{S}_{N_B}(\omega) = 4 \frac{k_{S_1}[S_0]}{[S_0] + K_M} PoPS_{in} [H_A(\omega)]_{\omega=0} \frac{1}{|\omega|^2}. \quad (381)$$

Finally, the upper bounds to the PSD  $\mathcal{S}_{N_{Tx}}(\omega)$  of the noise generated at the transmitter cell is equal to the sum of the PSD at the output of the signaling enzyme expression and signaling molecule production functional blocks, respectively, multiplied by the squared absolute value  $\|\cdot\|^2$  of the transfer functions of the functional blocks to reach the output of the transmitter cell. This results in the following:

$$\mathcal{S}_{N_{Tx}}(\omega) = \mathcal{S}_{N_A}(\omega) |H_B(\omega)|^2 + \mathcal{S}_{N_B}(\omega), \quad (382)$$

where  $\mathcal{S}_{N_A}(\omega)$ ,  $H_B(\omega)$ , and  $\mathcal{S}_{N_B}(\omega)$  are computed through (377), (336), and (381), respectively.

#### 7.4.1.2 Receiver Cell

In the receiver cell, we take into account the chemical reactions in (325), (327), and (328), which involve the signal  $PoPS_{aux}$ , the number of receptors  $R$ , the number of signaling

molecules at the receiver  $S_{Rx}$ , the number of activator complexes  $RS$ , and the output signal  $PoPS_{out}$ , as functions of the time  $t$ . We exclude from this model the number of RNAP molecules and the number of promoter sequences  $P_{Rx}$  at the receiver cells, since we assume them constant in time. The  $\tau$ -leaping statistical model is expressed for the transmitter cell through (358) where:

$$\mathbf{X}(t) = \left[ PoPS_{out} \quad RS \quad S \quad R \quad PoPS_{aux} \right]' (t), \quad (383)$$

the number of reactions  $M$  is equal to 5, and the propensity functions  $a_m(\mathbf{X}(t))$  and the stoichiometric vectors  $\mathbf{V}_m$ ,  $m = 6, \dots, 11$ , are computed for the transmitter cell reactions through the expressions in (360) and (361) as follows:

- In the **Receptor Activator Expression** functional block, the propensity function  $a_6(\mathbf{X}(t))$  and the stoichiometric vector  $\mathbf{V}_6$  are computed for the transcription and translation reaction in (325) as follows:

$$a_6(\mathbf{X}(t)) = k_R PoPS_{aux}(t), \quad \mathbf{V}_6 = \left[ 0 \quad 0 \quad 0 \quad n p_R \quad 0 \right]'. \quad (384)$$

For the receptor degradation reaction in (326), the propensity function  $a_7(\mathbf{X}(t))$  and the stoichiometric vector  $\mathbf{V}_7$  are as follows:

$$a_7(\mathbf{X}(t)) = k_{d_R} R(t), \quad \mathbf{V}_7 = \left[ 0 \quad 0 \quad 0 \quad -1 \quad 0 \right]'. \quad (385)$$

- In the **Ligand-Receptor Binding** functional block, the propensity function  $a_8(\mathbf{X}(t))$  and the stoichiometric vector  $\mathbf{V}_8$  for the formation of activator complexes  $RS$  in (323) are expressed as

$$a_8(\mathbf{X}(t)) = \frac{k_{RS}}{\Omega_{Tx}} R(t) S(t), \quad \mathbf{V}_8 = \left[ 0 \quad 1 \quad -1 \quad -1 \quad 0 \right]'. \quad (386)$$

where  $\Omega_{Tx}$  is the volume of the transmitter cell. For the dissociation of activator complexes reaction in (327), the propensity function  $a_9(\mathbf{X}(t))$  and the stoichiometric vector  $\mathbf{V}_9$  are as follows:

$$a_9(\mathbf{X}(t)) = k_{-RS} RS(t), \quad \mathbf{V}_9 = \left[ 0 \quad -1 \quad 1 \quad 1 \quad 0 \right]'. \quad (387)$$

- For the **Output Transcription Activation** functional block, the propensity function  $a_{10}(\mathbf{X}(t))$  and the stoichiometric vector  $\mathbf{V}_{10}$  are as follows:

$$a_{10}(\mathbf{X}(t)) = \frac{k_{Rx}}{\Omega_{Rx}^2} P_{Rx} RNAPRS(t), \quad \mathbf{V}_{10} = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 \end{bmatrix}'. \quad (388)$$

For the dissociation of open complexes dissociation reaction in (328), the propensity function  $a_{11}(\mathbf{X}(t))$  and the stoichiometric vector  $\mathbf{V}_{11}$  are as follows:

$$a_{11}(\mathbf{X}(t)) = k_{-Rx} POPS_{out}(t), \quad \mathbf{V}_{11} = \begin{bmatrix} -1 & 1 & 0 & 0 & 0 \end{bmatrix}'. \quad (389)$$

By isolating the contribution to the receiver cell noise coming from the reaction of formation and dissociation of activator complexes  $RS$  within the ligand-receptor binding functional block, we can derive the following expression:

$$\begin{aligned} RS(t + \tau) &= RS(t) + P_8 - P_9, \quad P_8 \sim \text{Poiss}(a_8(\mathbf{X}(t)), \tau) \\ P_9 &\sim \text{Poiss}(a_9(\mathbf{X}(t)), \tau), \end{aligned} \quad (390)$$

The expression in (390), whose general interpretation in terms of Chemical Master Equation (CME) [99] describes a Markov jump process [66], is **in agreement with the ligand-receptor-binding noise that we described in our previous work included in Section 4.5.**

The CLE formulation of the noise generated by the **Receptor Activator Expression** functional block is derived from (362) and (384) as follows:

$$\frac{dR(t)}{dt} = np_R k_R POPS_{aux} - k_{dR} R(t) + 2 \sqrt{np_R k_R POPS_{aux}} \Gamma_3(t), \quad (391)$$

where  $\Gamma_3(t)$  is a white Gaussian noise [66], and the coefficient 2 takes into account the steady state noise generation from both the transcription and translation reaction and the receptor degradation reaction, in agreement with [105]. The upper-bound PSD  $\tilde{S}_{N_3}(\omega)$  of the noise source of the signaling enzyme expression, as defined in (367), is equal to the following:

$$\tilde{S}_{N_3}(\omega) = 4np_R^2 k_R POPS_{aux}, \quad (392)$$

where  $np_R$  is the number of receptors produced per reaction,  $k_R$  is the receptor expression rate, and  $PoPS_{aux}$  is the auxiliary input signal, assumed constant in time, respectively. Inspired by the theoretical framework in [105], we compute the PSD  $\mathcal{S}_{N_D}(\omega)$  of the noise at the output of the receptor activator expression functional block as follows:

$$\mathcal{S}_{N_D}(\omega) = \frac{\tilde{\mathcal{S}}_{N_3}(\omega)}{k_{dR}^2} = 4 \frac{PoPS_{aux}}{k_R} |H_D(\omega)|^2, \quad (393)$$

where we applied the expression of  $H_D(\omega)$  defined in (341).

The CLE formulation of the noise generated by the **Ligand-Receptor Binding** functional block is derived from (362) and (386) as follows:

$$\frac{dRS(t)}{dt} = k_{RS}[R](t)S_{Rx}(t) - k_{-RS}RS(t) + 2\sqrt{k_{RS}[R](t)S_{Rx}(t)}\Gamma_4(t), \quad (394)$$

where  $[R](t) = R(t)/\Omega_{Rx}$ ,  $\Gamma_4(t)$  is a white Gaussian noise [66], and the coefficient 2 takes into account the steady state noise generation from both the formation and the dissociation of activator complexes  $RS$ , in agreement with [105]. The upper-bound PSD  $\tilde{\mathcal{S}}_{N_4}(\omega)$  of the noise source of the signaling enzyme expression, as defined in (367), is equal to the following:

$$\tilde{\mathcal{S}}_{N_4}(\omega) = 4k_{RS}PoPS_{aux}H_D(\omega)PoPS_{in} [H_A(\omega)H_B(\omega)H_C(\omega)]_{\omega=0}, \quad (395)$$

where  $k_{RS}$  is the number of receptors produced per reaction,  $PoPS_{aux}$  is the auxiliary input signal, assumed constant in time, and  $H_D(\omega)$  is expressed in (341), respectively. We computed the value of  $S_{Rx}(t)$  by multiplying the constant input signal  $PoPS_{in}$  by the product of the transfer function  $H_A(\omega)$ , and  $H_B(\omega)$  and  $H_C(\omega)$ , computed at frequency  $\omega = 0$ . Inspired by the theoretical framework in [105], we compute the PSD  $\mathcal{S}_{N_E}(\omega)$  of the noise at the output of the ligand-receptor binding functional block as follows:

$$\mathcal{S}_{N_E}(\omega) = \frac{\tilde{\mathcal{S}}_{N_4}(\omega)}{\omega^2 + k_{-RS}^2} = 4 \frac{PoPS_{in}PoPS_{aux}}{k_{RS}} [H_A(\omega)H_B(\omega)H_C(\omega)]_{\omega=0} H_D(\omega) |H_E(\omega)|^2, \quad (396)$$

where we applied the expression of  $H_E(\omega)$  defined in (351).

The CLE formulation of the noise generated by the **Output Transcription Activation** functional block is derived from (362) and (388) as follows:

$$\frac{dPoPS_{out}(t)}{dt} = k_{Rx}[P_{Rx}][RNAP]RS(t) - k_{-Rx}RS(t) + 2\sqrt{k_{Rx}[P_{Rx}][RNAP]RS(t)}\Gamma_5(t), \quad (397)$$

where  $[P_{Rx}][RNAP] = P_{Rx}RNAP/\Omega_{Rx}^2$ ,  $\Gamma_5(t)$  is a white Gaussian noise [66], and the coefficient 2 takes into account the steady state noise generation from both the open complex formation and the open complex degradation reactions, in agreement with [105]. The upper-bound PSD  $\tilde{S}_{N_5}(\omega)$  of the noise source of the output transcription activation, as defined in (367), is equal to the following:

$$S_{N_5}(\omega) = 4k_{Rx}[P_{Rx}][RNAP]PoPS_{aux}H_D(\omega)PoPS_{in}[H_A(\omega)H_B(\omega)H_C(\omega)H_E(\omega)]_{\omega=0}, \quad (398)$$

where  $k_{Rx}$  is the open complex formation rate,  $[P_{Rx}]$  and  $[RNAP]$  are the constant concentrations of promoter sequences and RNA polymerase at the receiver cell, respectively. We computed the value of  $RS(t)$  by multiplying the auxiliary input signal  $PoPS_{aux}$ , assumed constant in time, by  $H_D(\omega)$  and by multiplying the constant input signal  $PoPS_{in}$  by the transfer function  $H_A(\omega)$ ,  $H_B(\omega)$ ,  $H_C(\omega)$ , and  $H_E(\omega)$  computed at frequency  $\omega = 0$ . Inspired by the theoretical framework in [105], we compute the PSD  $S_{N_F}(\omega)$  of the noise at the output of the output transcription activation functional block as follows:

$$S_{N_F}(\omega) = \frac{\tilde{S}_{N_5}(\omega)}{\omega^2 + k_{-Rx}^2} = 4\frac{PoPS_{in}PoPS_{aux}[P_{Rx}][RNAP]}{k_{Rx}} \cdot [H_A(\omega)H_B(\omega)H_C(\omega)H_E(\omega)]_{\omega=0} H_D(\omega) |H_F(\omega)|^2, \quad (399)$$

where we applied the expression of  $H_F(\omega)$  defined in (346).

Finally, the upper bound to the PSD  $S_{N_{Rx}}(\omega)$  of the noise generated at the receiver cell is equal to the sum of the PSD at the output of the receptor activator expression, ligand-receptor binding and output transcription activation functional blocks, respectively, multiplied by the squared absolute value  $||\cdot||^2$  of the transfer functions of the functional blocks to

reach the output of the receiver cell. This results in the following:

$$\mathcal{S}_{N_{Rx}}(\omega) = \mathcal{S}_{N_D}(\omega) |H_E(\omega)|^2 |H_F(\omega)|^2 + \mathcal{S}_{N_E}(\omega) |H_F(\omega)|^2 + \mathcal{S}_{N_F}(\omega) , \quad (400)$$

where  $\mathcal{S}_{N_D}(\omega)$ ,  $H_E(\omega)$ ,  $H_F(\omega)$ ,  $\mathcal{S}_{N_E}(\omega)$ , and  $\mathcal{S}_{N_F}(\omega)$  are computed through (393), (351), (346), (396), and (399), respectively.

#### 7.4.2 Noise in Diffusion

According to our results included in Section 4.3, the noise source that stems from the diffusion process, termed particle counting noise, can be stochastically modeled by considering the number  $S_{Rx}(t)$  of signaling molecules in the receiver cell at time  $t$  as the realization within the receiver volume  $\Omega_{Rx}$  of a spatial inhomogeneous Poisson counting process [66] whose rate is the concentration  $[S]_{rx}(t)$  of molecules in the receiver cell at time  $t$ , computed as if no noise were affecting the diffusion process. This stochastic model is expressed as follows:

$$S_{Rx}(t) \sim \text{Poiss}([S]_{rx}(t), \Omega_{Rx}) . \quad (401)$$

According to the Central Limit Theorem [66], under the assumption to have a sufficiently high value for the receiver volume  $\Omega_{Rx}$ , the Poisson distribution in (401) can be approximated as a Gaussian distribution with average value and variance equal to  $[S]_{rx}(t)\Omega_{Rx}$ . This approximation is expressed as follows:

$$S_{Rx}(t) \simeq [S]_{rx}(t)\Omega_{Rx} + \sqrt{[S]_{rx}(t)\Omega_{Rx}}\Gamma , \quad (402)$$

where  $\Gamma$  is a white Gaussian noise [66].

PSD  $\mathcal{S}_{N_{Diff}}(\omega)$  of the noise source at the diffusion process affecting the biological circuit is computed through the following formula [66]:

$$\mathcal{S}_{N_{Diff}}(\omega) = \text{E} \left[ \left| \mathcal{F} \left\{ \sqrt{[S]_{rx}(t)\Omega_{Rx}}\Gamma \right\} \right|^2 \right] , \quad (403)$$

where  $\text{E}[\cdot]$  and  $\mathcal{F}\{\cdot\}$  denote the average value and the Fourier transform [51] operators, respectively.

Under the assumption to have a constant value for the input signal  $PoPS_{in}$ , by considering that the previous noise sources in the biological circuit have a negligible effect on the concentration  $[S]_{rx}(t)$  of molecules in the receiver cell, as function of the time  $t$ , and by taking into account the expressions in (364) and in (365), we can derive an upper bound  $\tilde{\mathcal{S}}_{N_{Diff}}(\omega)$  to the PSD of the noise source at the diffusion process. This upper bound is expressed as

$$\tilde{\mathcal{S}}_{N_{Diff}}(\omega) = PoPS_{in}H_A(0)H_B(0)H_C(0)\Omega_{Rx} , \quad (404)$$

where we computed the value of the concentration  $[S]_{rx}(t)$  of molecules in the receiver cell by multiplying the constant input signal  $PoPS_{in}$  by the values of the transfer function  $H_A(\omega)$  of the signaling enzyme expression (333), the transfer function  $H_B(\omega)$  of the signaling molecule production (336), and the transfer function  $H_C(\omega)$  of the diffusion process (339) functional blocks at the frequency value  $\omega = 0$ .

### 7.4.3 Output Noise PSD Expression

The PSD  $\mathcal{S}_{N_{out}}(\omega)$  of the noise generated at the output of the biological circuit is computed through the sum of the following terms: i) the upper bounds to the PSD  $\mathcal{S}_{N_{Tx}}(\omega)$  of the noise generated at the transmitter cell multiplied by the cascade of the squared absolute values of the transfer functions of the functional blocks between the transmitter and the output of the biological circuit, namely,  $H_C(\omega)H_D(\omega)H_E(\omega)H_F(\omega)$ ; ii) the upper bound  $\tilde{\mathcal{S}}_{N_{Diff}}(\omega)$  to the PSD of the noise source at the diffusion process multiplied by the cascade of the squared absolute values of the transfer functions of the functional blocks in the transmitter cell, namely,  $H_D(\omega)H_E(\omega)H_F(\omega)$ ; iii) the upper bound to the PSD  $\mathcal{S}_{N_{Rx}}(\omega)$  of the noise generated at the receiver cell. This is expressed as follows:

$$\begin{aligned} \mathcal{S}_{N_{out}}(\omega) = & \mathcal{S}_{N_{Tx}}(\omega) |H_C(\omega)H_D(\omega)H_E(\omega)H_F(\omega)|^2 + \\ & \tilde{\mathcal{S}}_{N_{Diff}}(\omega) |H_D(\omega)H_E(\omega)H_F(\omega)|^2 + \mathcal{S}_{N_{Rx}}(\omega) , \end{aligned} \quad (405)$$

where  $\mathcal{S}_{N_{Tx}}(\omega)$ ,  $H_C(\omega)$ ,  $H_D(\omega)$ ,  $H_E(\omega)$ ,  $H_F(\omega)$ ,  $\tilde{\mathcal{S}}_{N_{Diff}}(\omega)$ , and  $\mathcal{S}_{N_{Rx}}(\omega)$  are expressed in (382), (339), (341), (351), (346), (404), and (400), respectively.

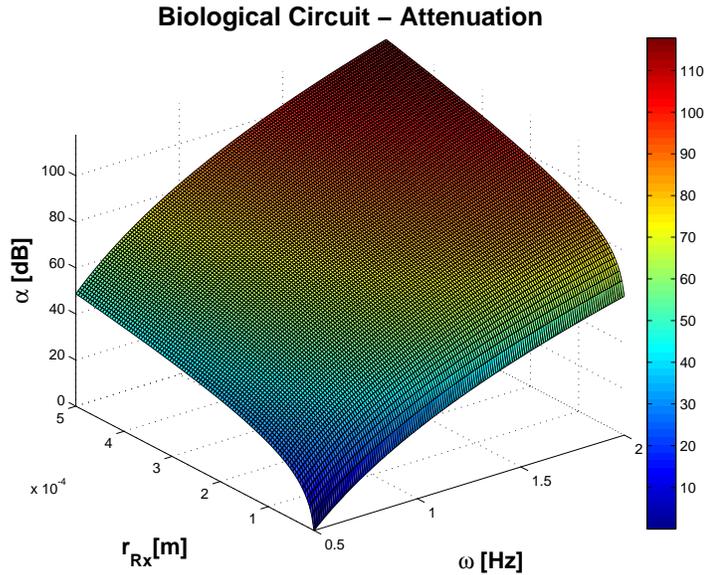
## 7.5 Numerical Results

In this section, we present some preliminary numerical results obtained through the evaluation of the expressions of the deterministic model and the stochastic model of the biological circuit for diffusion-based MC analyzed in this chapter of the Ph.D. thesis.

In the synthetic biology and biological circuit engineering literature, there are currently very few works that focus on the joint experimental determination of the biochemical parameters of a complete biological circuit implementation, such as the parameters we introduced in Section 7.2.2. Most of the results presented in the literature focus on the study of one specific element or biochemical reaction rather than a complete architecture. As a consequence, the values of the biochemical parameters used for our numerical results are taken from a diverse pool of papers and, although they satisfy the goal of having a realistic order of magnitude, they do not necessarily capture the values that they would have in a real implementation of the biological circuit we analyze here.

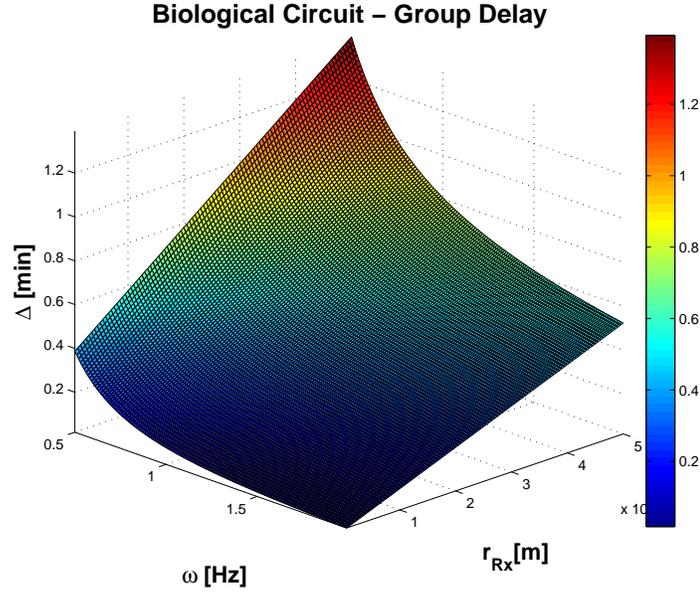
For the numerical results of both the deterministic model and the stochastic model, presented in the following, we applied the following parameter values from the LuxR-LuxI quorum sensing system [108] in *E. coli* bacteria, which has been already used for the engineering of a biological circuit for diffusion-based molecular communication, as in [21]. The rates of signaling enzyme and receptor activator translation equal to the rate of Lux protein translation from [109], namely,  $k_E = k_R = 9.6 \times 10^{-1} \text{ min}^{-1}$ , the rate of signaling enzyme degradation equal to the degradation rate of LuxI protein in [109], namely,  $k_{dE} = 1.67 \times 10^{-2} \text{ min}^{-1}$ , the rate of receptor degradation equal to the degradation rate of LuxR protein in [109], namely,  $k_{dR} = 2.31 \times 10^{-2} \text{ min}^{-1}$ , the Michaelis-Menten constant, defined in (379), equal to the constant computed for the hexanol homoserine lactone synthase activity in [110], namely,  $K_M = 130 \mu\text{M}$ , the complex formation rate from the binding of the signaling enzymes and the substrates equal to the forward LuxI-substrates reaction for autoinducer molecule production in [111], namely,  $k_{S_1} = 0.6 \text{ molecules}^{-1} \text{ min}^{-1}$ , the binding and unbinding rates between receptors and signaling molecules equal to the values

in [109], namely,  $k_{RS} = 6 \times 10^{-4} \text{ molecules}^{-1} \text{ min}^{-1}$  and  $k_{-RS} = 2 \times 10^{-2} \text{ min}^{-1}$ , respectively, and the rate of open complex formation upon output transcription activation and open complex dissociation at the receiver as in [109], namely,  $k_{Rx} = 10^{-2} \text{ molecules}^{-1} \text{ min}^{-1}$  and  $k_{-Rx} = 4 \times 10^{-2} \text{ min}^{-1}$ . The diffusion coefficient  $D \sim 60 \times 10^{-9} \text{ m}^2 \text{ min}^{-1}$  is set to the diffusion coefficient of molecules diffusing in a biological environment (cellular cytoplasm, [58]). We set the volume of the transmitter and receiver volume  $\Omega_{Tx} = \Omega_{Rx}$  to  $1 \text{ } \mu\text{m}^2$ , the number of substrates at the transmitter cell  $S_0 = 100$ , the auxiliary input signal  $PoPS_{aux} = 1$ , the number of promoter sequences and RNA polymerases at the receiver cell  $P_{Rx} = 1$  and  $RNAP = 100$ , respectively.



**Figure 73. Attenuation of the biological circuit for diffusion-based MC as function of the receiver distance from the transmitter  $r_{Rx}$  and the frequency  $\omega$ .**

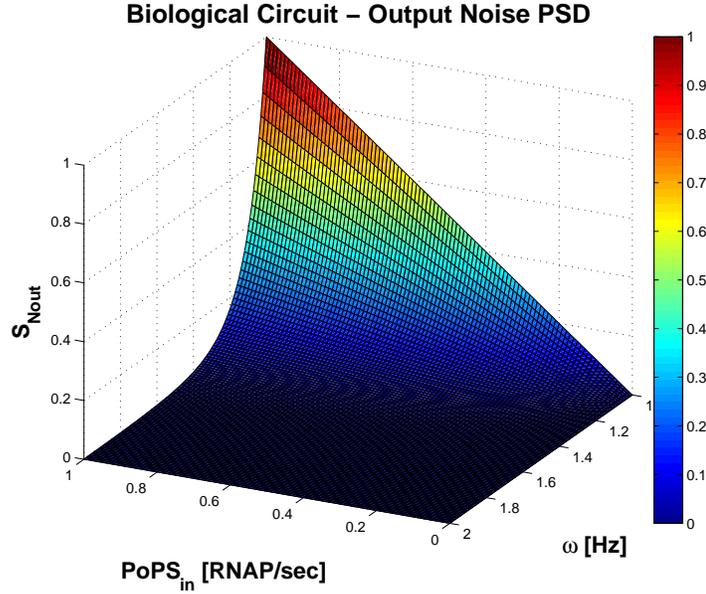
In Figure 73 and Figure 74 we show the numerical results for the attenuation  $\alpha$  and delay  $\Delta$  experienced by a signal through the biological circuit, computed by using the expressions in (354) and (356), respectively. The value of the receiver distance from the transmitter  $r_{Rx}$  ranges from 5 to  $500 \text{ } \mu\text{m}$ , while the range of observed frequency  $\omega$  values is between 0 and 1 Hz. The values for the attenuation  $\alpha$  of the biological circuit, shown in dB, range from a minimum of 0 at the minimum distance  $r_{Rx}$  and frequency  $\omega$  equal to 0, to a maximum



**Figure 74.** Delay of the biological circuit for diffusion-based MC as function of the receiver distance from the transmitter  $r_{Rx}$  and the frequency  $\omega$ .

of 117.8 dB for a distance  $r_{Rx}$  of 50  $\mu\text{m}$  and frequency  $\omega$  equal to 1 Hz with a monotonic increasing trend both when increasing the distance  $r_{Rx}$  and the frequency  $\omega$ . The values for the delay  $\Delta$  range from a minimum of 0 min at the minimum distance  $r_{Rx}$  and maximum frequency  $\omega$  equal to 1 Hz, to a maximum of 1.4 min at a distance  $r_{Rx}$  of 50  $\mu\text{m}$  and for a frequency  $\omega$  equal to 0. The curves of the delay  $\Delta$  always show a monotonically decreasing trend as function of the frequency  $\omega$ , more pronounced for higher values of the distance  $r_{Rx}$ .

In Figure 75, we show the values of the output noise PSD  $\mathcal{S}_{N_{out}}(\omega)$  computed through the expression in (405) for a range of the frequency  $\omega$  between 1 and 2 Hz, and for an input  $PoPS_{in}$ , assumed constant in time, ranging from a value of 0 to a maximum of 1 polymerase per second. The minimum PSD  $\mathcal{S}_{N_{out}}(\omega)$  close to 0 is achieved for a frequency  $\omega$  equal to 2 Hz and for a  $PoPS_{in}$  equal to 0 polymerase per second, and this result clearly shows the direct dependency between the input signal  $PoPS_{in}$  and the magnitude of the noise produced within the biological circuit. The maximum PSD  $\mathcal{S}_{N_{out}}(\omega)$ , close to 1 (polymerase per second)<sup>2</sup>, is achieved for a frequency  $\omega$  equal to 1 Hz and the maximum considered



**Figure 75.** PSD of the output noise from the biological circuit for diffusion-based MC as function of the frequency  $\omega$  and the input signal  $PoPS_{in}$ .

value for the input signal  $PoPS_{in}$  of 1 polymerase per second. This latter result shows that it is possible under some conditions to have an output noise from the biological circuit with a magnitude comparable to the input signal. The curves of the output noise PSD  $\mathcal{S}_{N_{out}}(\omega)$  show a quasi-constant trend as function of the input  $PoPS_{in}$  for high frequency  $\omega$ , while for low values of the frequency, the curves are monotonically increasing according to the value of the input  $PoPS_{in}$ .

## 7.6 Conclusion

In this chapter of the Ph.D. thesis, deterministic and stochastic communication engineering models are presented for a diffusion-based molecular communication system design based on biological circuits. Biological circuits are defined as genetic regulatory networks embedded in a biological cell, and they are envisioned to allow the future engineering of complete biological nanomachines. Some recent literature can be found on the analytical modeling of biological circuits, but with no specific mention to diffusion-based cell-to-cell communication through molecule exchange, for which only a biological description is

provided in some specific works.

In our work, first, a biological circuit for diffusion-based molecular communication is identified through a minimal subset of elements. Then, a mathematical model is detailed in terms of transfer functions, from which analytical expressions are derived for the attenuation and the delay experienced by an information signal through the biological circuits. Finally, the most significant noise sources within the biological circuit are identified, and statistical models for these sources are provided in terms of noise-generating random processes. For each statistical model, the impact of the generated noise on the biological circuit is quantified through the Power Spectral Density (PSD) parameter at the output of the biological circuit.

Numerical results for the attenuation and delay parameters show similar trends as in the physical end-to-end modeling presented in Chapter 3, while the general noise sources detailed in Chapter 4 find in this work a further confirmation of their validity through their stochastic modeling in a biological environment.

## CHAPTER 8

### CONCLUSION

Molecular communication (MC) is a promising bio-inspired paradigm for the exchange of information among autonomous intelligent nanotechnology-enabled devices, or nanomachines. MC realizes the exchange of information through the transmission, propagation, and reception of molecules, and it is proposed as a feasible solution for nanonetworks. This idea is motivated by the observation of nature, where MC is successfully adopted by cells for intracellular and intercellular communication. Thanks to the feasibility of MC in biological environments, MC-based nanonetworks have the potential to be the enabling technology for a wide range of applications, mostly in the biomedical, but also in the industrial and surveillance fields.

The focus of this Ph.D. thesis is on diffusion-based MC, where the propagation of information-bearing molecules between a transmitter and a receiver is realized through free diffusion in a fluid. This choice is motivated by a preliminary analysis, which identifies the diffusion-based as the most fundamental type of MC among different options suggested in the literature. Since there are profound differences between the diffusion-based MC paradigm and classical electromagnetic communication paradigms, the classical communication engineering models and techniques are not directly applicable for the study and the design of diffusion-based MC systems. As a consequence, there is a need of to build a complete understanding of the diffusion-based MC paradigm from the ground up.

The objectives of the research presented in this thesis are to analyze the diffusion-based MC paradigm from the point of view of communication engineering and information theory, and to provide solutions to the modeling and design of MC-based nanonetworks. First, a physical end-to-end model is realized to study each component in diffusion-based MC, as well as the overall system, in terms of gain and delay. Second, the noise sources affecting

the communication of information through diffusion-based MC are identified and statistically modeled. Third, upper and lower bounds to the capacity are derived to evaluate the information-theoretic performance of the diffusion-based MC paradigm. Fourth, a stochastic analysis of the interference when multiple transmitters access the diffusion-based MC channel simultaneously is provided. The main contributions included in each chapter of this Ph.D. thesis are summarized in the following.

Chapter 3 is devoted to the physical end-to-end model of diffusion-based MC. In particular, the main contributions are as follows:

- We provide a basic bio-inspired diffusion-based MC system design, which aims at an interpretation of the diffusion-based MC in terms three processes, namely, molecule emission, molecule propagation, and molecule reception.
- We study each process of the designed system by modeling the underlying physical phenomena with an equivalent electrical circuit model, for which it is possible to derive an input-output transfer function with a closed-form expression.
- We analyze the transfer function of each process, as well as their cascade that composes the overall system, in terms of gain and delay experienced by an information signal exchanged through the designed diffusion-based MC system, as function of the signal frequency components.

Through the results from this physical end-to-end model we learn that the gains of the molecule emission, molecule propagation, and molecule reception in a diffusion-based MC have non-linear curves as function of the frequency components of the information signal exchanged through the system, and as function of the distance range. Non-linear curves as function of the frequency components are also shown for the delays of the molecule reception and propagation processes. Moreover, the delay of the propagation process shows non-linear curves also as function of the distance range, but only for low-frequency components.

Chapter 4 contains the analysis of the noises affecting a diffusion-based MC system. In particular, the main contributions are as follows:

- We identify three noise sources affecting the diffusion-based MC, namely, the sampling noise, the counting noise, and the ligand-receptor-binding noise, which are related to the transmitter, the signal propagation in the channel, and the receiver, respectively.
- We provide for each noise source a physical model based on a mathematical analysis of the physical processes that generate the noise. Each physical model is characterized by an algorithm, summarized through a block scheme, which provides a means to simulate the generation of a particular noise in diffusion-based MC.
- We analytically derive for each noise source a stochastic model that aims at capturing the underlying physical processes through statistical parameters. Each stochastic model summarizes the generation of a particular diffusion-based MC noise using random processes and their associated parameters. We evaluated each stochastic model ability to capture the behavior of the physical processes that generate the noise through results of simulations based on the corresponding physical model.

Through the results of this noise analysis we learn that the diffusion-based MC noises arise from the discrete nature of the signaling molecules used for the transmission of the information signals, from the randomness of their Brownian motion propagation, and from the stochasticity of the chemical reactions in which they are involved. As a result of the peculiarities of these noise-generating processes, the statistical parameters of the noises in a diffusion-based MC system are functions of the amplitude of the information signal exchanged between the transmitter and the receiver.

Chapter 5 is focused on the capacity analysis of the diffusion-based MC paradigm. This analysis is performed by accounting for two main characteristics of the diffusion-based MC, which make impossible to find a closed-form analytical expression for the true

capacity. As a consequence, in the attempt to provide an analytical closed-form expression that relates the performance of a diffusion-based MC system to physical parameters, the contributions included in this chapter are based on some simplifying assumptions. In particular, these contributions are as follows:

- We apply solutions from statistical mechanics and equilibrium thermodynamics to find a relationship between thermodynamic entropy and information entropy in a diffusion-based MC system. On the basis of this relationship, we derive a closed-form expression of an upper bound to the true capacity in diffusion-based MC, as a function of the parameters from statistical mechanics. This preliminary expression is derived through the simplifying assumption of having a molecular system in equilibrium, and the dynamic effects of the diffusion-based channel are not taken into account.
- We provide lower-bound expression to the true information capacity of a diffusion-based MC system, by taking into account both the dynamic effects of the channel and the signal-dependent noise, termed molecular noise. Through a simplification of the molecule diffusion into the composition of two main processes, namely, the Fick's diffusion and the particle location displacement, we analytically derive a capacity lower-bound expression independent from any specific coding scheme, and expressed as a function of the average transmitted power.

Through the results of the capacity analysis we learn how the performance of a diffusion-based MC system depend on the system bandwidth, the distance range, and the average transmitted power. In particular, capacity values of a few [Kbit/sec] can be reached within a distance of tenth of  $\mu\text{m}$  between the transmitter and the receiver, and for an average transmitted power around 1 pW (Note that this power value should not be compared to the transmitted power values used for electrical devices, since the transmitted power in a MC system is a thermodynamic quantity).

Chapter 6 contains an analysis of the interference in diffusion-based MC systems. In particular, the main contributions are as follows:

- We jointly analyze the InterSymbol Interference (ISI) and the Co-Channel Interference (CCI) for a diffusion-based MC system having a limited number of transmitters in predetermined locations. We derive simple closed-form formulas for both the ISI and the CCI based on the attenuation and the dispersion of information signals encoded through Gaussian-pulses.
- We provide a statistical-physical modeling of the interference through an analytical expression of the Power Spectral Density (PSD) probability distribution of the received signal, independent from the transmitter number, specific transmitter locations or coding schemes. This PSD analytical expression is evaluated against the outcomes of a simulation environment.

Through the results of the interference analysis, we learn that the ISI and CCI of a Gaussian-pulse-encoded information signal have lower values when this signal is modulated by a carrier oscillation. Moreover, the higher is the frequency of the carrier oscillation, the lower are the values of the ISI and CCI. Through the statistical-physical modeling of the interference, we derive the PSD probability distribution of a received signal in a diffusion-based MC system as function of physical parameters, such as the diffusion coefficient, the transmitter density, and the average power of molecule emissions. As apparent from the PDF of the received PSD, the power of low frequency components of the received signal tends to a uniform distribution over the range of considered values, while the power of higher frequencies tends with more probability to lower values.

Chapter 7 is focused on the analysis of a diffusion-based MC system design built upon genetically-engineered biological circuits. In particular, the main contributions are as follows:

- We identify a minimal subset of elements necessary to realize diffusion-based molecular communication between biological cells. These elements are compliant with the current biological circuit standardization efforts where the units to measure the input and the output are defined as Polymerases Per Second (PoPS). We provide a description of the biological mechanisms underlying each element, which is then characterized in terms of chemical reactions.
- We mathematically analyze the elements of the biological circuit and provide deterministic models in terms of transfer functions. The cascade of these transfer functions allow to derive analytical closed-form expressions for the attenuation and delay experienced by an information signal through the biological circuit.
- We provide a stochastic model for each noise source within the biological circuit by stemming from the noise analysis in biochemical reactions. For each noise source, we quantify the impact of the generated noise in terms of Power Spectral Density (PSD) at the output of the biological circuit.

Through this analysis of a diffusion-based MC design, we learn that the engineering of biological circuits is a key technology for the implementation of this communication paradigm in future biological nanomachines. Moreover, in this chapter we provide a proof-of-concept of the modeling techniques and results presented for the general diffusion-based MC case in Chapter 3 and Chapter 4.

In the future, we plan to extend our work on molecular communication in several directions, including 1) an analytical framework to assess the energy consumptions for transmitting, receiving, and processing information in diffusion-based MC systems, with reference to the energy modeling in the biochemical reactions involved in the cellular bio-signaling; 2) the study of addressing mechanisms in diffusion-based MC systems; 3) the research on the notion of message “packet” within the diffusion-based MC paradigm; 3) the study of other molecular communication architectures, e.g., advection-based and walkway-based,

and the definition of end-to-end physical models, the analysis of the related noise sources, information capacity, interference, energy consumption, and addressing techniques; 4) the study of more specific designs for molecular communication, stemming from the engineering of synthetic biological circuits for the realization of biological nanomachines; 5) the design of cooperative systems based on the molecular-communication-mediated interaction within swarms of biological nanomachines; 6) the characterization of intrabody molecular transport networks, such as the gastrointestinal system, the endocrine system, or the cardiovascular systems, for the realization of molecular communication networks for future drug-delivery and diagnostic applications.

## CHAPTER 9

### APPENDICES

#### 9.1 Particle Binding and Release Rates

The rate  $k_+$  of particle binding is defined in chemical kinetics [112] as follows:

$$k_+ = ZF_C(E_a) , \quad (406)$$

where  $Z$  is the average collision frequency,  $E_a$  is the activation energy of the ligand-receptor binding and  $F_C(E_a)$  is the fraction of collisions having a higher energy than the activation energy. The average collision frequency  $Z$  quantifies how frequently a collision occurs between a particle and an unbound receptor and it is expressed as follows:

$$Z = \langle N_p \rangle \langle N_R - n_b \rangle \pi(r_p + r_R)^2 \langle \bar{v}_p \rangle , \quad (407)$$

where  $c_R(t)$  is the concentration of particles at the receiver,  $\langle N_p \rangle$  is the average number of particles inside the receptor space,  $\langle N_R - n_b \rangle$  is the average number of unbound chemical receptors,  $r_p$  and  $r_R$  are the radius of a particle and a receptor in the system, respectively, and  $\langle \bar{v}_p \rangle$  is the average velocity of the particles. The fraction  $F_C(E_a)$  of collisions having a higher energy than the activation energy is expressed through the Boltzmann distribution [113] as

$$F_C(E_a) = e^{-\frac{E_a}{k_B T}} , \quad (408)$$

where  $E_a$  is the activation energy,  $T$  is the absolute temperature of the system and  $k_B$  is the Boltzmann constant. The average collision frequency  $Z$  depends on the average velocity  $\langle \bar{v}_p \rangle$  of the particles subject to the Brownian motion and it is known from the kinetic theory [113] to have the following expression:

$$\langle \bar{v}_p \rangle = \sqrt{\frac{8k_B T}{\pi m_p}} , \quad (409)$$

where  $k_B$  is the Boltzmann constant,  $T$  is the absolute temperature of the system and  $m_p$  is the mass of a particle.

The rate  $k_-$  of particle release is defined in transition state theory [114] as follows:

$$k_- = \frac{k_B T}{h} \left( 1 - e^{-\frac{h\nu}{k_B T}} \right) e^{-\frac{E^\theta}{RT}}, \quad (410)$$

where  $k_B$  is the Boltzmann constant,  $T$  is the absolute temperature of the system,  $h$  is the Planck constant,  $\nu$  is the vibrational frequency of the bond,  $E^\theta$  is the unbinding energy at absolute zero and  $R$  is the universal gas constant.

## PUBLICATIONS

### Journals

1. M. Pierobon and I. F. Akyildiz, "Capacity of a diffusion-based molecular communication system with channel memory and molecular noise," *IEEE Transactions on Information Theory*, vol. 59, no. 2, pp. 942-954, Feb. 2013.
2. M. Pierobon and I. F. Akyildiz, "Noise analysis in ligand-binding reception for molecular communication in nanonetworks," *IEEE Transactions on Signal Processing*, vol. 59, no. 9, pp. 4168-4182, Sept. 2011.
3. M. Pierobon and I. F. Akyildiz, "Diffusion-based noise analysis for molecular communication in nanonetworks," *IEEE Transactions on Signal Processing*, vol. 59, no. 6, pp. 2532-2547, June 2011.
4. M. Pierobon and I. F. Akyildiz, "A physical end-to-end model for molecular communication in nanonetworks," *IEEE Journal on Selected Areas in Communications (JSAC)*, vol. 28, no. 4, pp. 602-611, May 2010.
5. I. F. Akyildiz, J. M. Jornet, and M. Pierobon, "Nanonetworks: a new frontier in communications," *Communications of the ACM*, vol. 54, no. 11, pp. 84-89, Nov. 2011.
6. I. Llatser, A. Cabellos-Aparicio, M. Pierobon, and E. Alarcon, "Detection techniques for diffusion-based molecular communication," *IEEE Journal on Selected Areas in Communications (JSAC)*, to appear, 2013.
7. Y. Chahibi, M. Pierobon, S. O. Song, and I. F. Akyildiz, "A molecular communication system model for particulate drug delivery systems," *IEEE Transactions on Biomedical Engineering*, to appear, 2013.
8. M. Pierobon and I. F. Akyildiz, "A statistical-physical model of interference in diffusion-based molecular nanonetworks," *submitted for journal publication*, 2013.

### Conferences

1. M. Pierobon and I. F. Akyildiz, "Information capacity of diffusion-based molecular communication in nanonetworks," in *Proc. of IEEE Intl. Conf. on Computer Communication, INFOCOM, Miniconference*, Shanghai, China, Apr. 2011.
2. M. Pierobon and I. F. Akyildiz, "Intersymbol and co-channel interference in diffusion-based molecular communication," in *Proc. of the 2nd IEEE Intl. Workshop on Molecular and Nano Scale Communication (MoNaCom), ICC*, Ottawa, Canada, June 2012.
3. I. Llatser, E. Alarcon, and M. Pierobon, "Diffusion-based channel characterization in molecular nanonetworks," in *Proc. of 1st IEEE Intl. Workshop on Molecular and Nano Scale Communication (MoNaCom), INFOCOM*, Shanghai, China, Apr. 2011.

4. N. Garralda, I. Llatser, A. Cabellos-Aparicio, and M. Pierobon, "Simulation-based evaluation of the diffusion-based physical channel in molecular nanonetworks," in *Proc. of 1st IEEE Intl. Workshop on Molecular and Nano Scale Communication (MoNaCom), INFOCOM*, Shanghai, China, Apr. 2011.
5. I. Llatser, I. Pascual, N. Garralda, A. Cabellos-Aparicio, E. Alarcon, and M. Pierobon, and J. Solé-Pareta, "Exploring the physical channel of diffusion-based molecular communication by simulation," in *Proc. of Global Telecommunications Conf. (GLOBECOM)*, Houston, USA, Dec. 2011.
6. I. F. Akyildiz, J. M. Jornet and M. Pierobon, "Propagation models for nanocommunication networks," in *Proc. of EUCAP 2010, Fourth European Conf, on Antennas and Propagation*, Barcelona, Apr. 2010.

## REFERENCES

- [1] I. F. Akyildiz, F. Brunetti, and C. Blazquez, “Nanonetworks: a new communication paradigm at molecular level,” *Computer Networks (Elsevier) Journal*, vol. 52, pp. 2260–2279, Aug. 2008.
- [2] D. L. Nelson and M. M. Cox, *Lehninger Principles of Biochemistry*, ch. 12.2, pp. 425–429. W. H. Freeman, 2005.
- [3] I. F. Akyildiz, J. M. Jornet, and M. Pierobon, “Nanonetworks: a new frontier in communications,” *Communications of the ACMs*, vol. 54, pp. 84–89, Nov. 2011.
- [4] E. Drexler, “Molecular engineering: Assemblers and future space hardware.” American Astronautical Society: AAS-86-415, 1986.
- [5] E. Drexler, *Nanosystems: Molecular Machinery, Manufacturing, and Computation*. John Wiley and Sons, Inc, 1992.
- [6] I. F. Akyildiz and J. M. Jornet, “Electromagnetic wireless nanosensor networks,” *Nano Communication Networks (Elsevier) Journal*, vol. 1, pp. 3–19, June 2010.
- [7] I. F. Akyildiz and J. M. Jornet, “The internet of nano-things,” *IEEE Wireless Communication Magazine*, vol. 17, pp. 58–63, Dec. 2010.
- [8] Y.-C. Hung, W.-T. Hsu, T.-Y. Lin, and L. Fruk, “Photoinduced write-once read-many-times memory device based on dna biopolymer nanocomposite,” *Appl. Phys. Lett.*, vol. 99, p. 253301, Oct. 2011.
- [9] B. Behkam and M. Sitti, “Bacterial flagella-based propulsion and on/off motion control of microscale objects,” *Appl. Phys. Lett.*, vol. 90, p. 023902, Jan. 2007.
- [10] D. Endy, “Foundations for engineering biology,” *Nature*, vol. 438, pp. 449–453, Nov. 2005.
- [11] D. Baker, G. Church, J. Collins, D. Endy, J. Jacobson, J. Keasling, P. Modrich, C. Smolke, and R. Weiss, “Engineering life: building a fab for biology,” *Scientific American*, vol. 294, pp. 44–51, June 2006.
- [12] H. Salis, A. Tamsir, and C. Voigt, “Engineering bacterial signals and sensors,” *Contrib. Microbiol.*, vol. 16, pp. 194–225, June 2009.
- [13] P. Tallury, A. Malhotra, L. M. Byrne, and S. Santra, “Nanobioimaging and sensing of infectious diseases,” *Advanced Drug Delivery Reviews*, vol. 62, pp. 424–437, Mar. 2010.

- [14] R. A. Freitas, *Nanomedicine, Volume I: Basic Capabilities*. Landes Bioscience, Georgetown, TX, 1999.
- [15] J. M. Dubach, D. I. Harjes, and H. A. Clark, “Fluorescent ion-selective nanosensors for intracellular analysis with improved lifetime and size,” *Nano Letters*, vol. 7, no. 6, pp. 1827–2831, 2007.
- [16] J. Li, T. Peng, and Y. Peng, “A cholesterol biosensor based on entrapment of cholesterol oxidase in a silicic sol-gel matrix at a prussian blue modified electrode,” *Electroanalysis*, vol. 15, no. 12, pp. 1031–1037, 2003.
- [17] J. W. Costerton, P. S. Stewart, and E. P. Greenberg, “Bacterial biofilms: a common cause of persistent infections,” *Science*, vol. 67, pp. 1318–1322, May 1999.
- [18] C. Kerr, K. Osborn, G. Robson, and P. Handley, “The relationship between pipe material and biofilm formation in a laboratory model system,” *Journal of Applied Microbiology*, vol. 85, pp. 295–385, 1999.
- [19] J. Riu, A. Maroto, and F. X. Rius, “Nanosensors in environmental analysis,” *Talanta*, vol. 69, no. 2, pp. 288–301, 2006.
- [20] C. R. Yonzon, D. A. Stuart, X. Zhang, A. D. McFarland, C. L. Haynes, and R. P. V. Duyne, “Towards advanced chemical and biological nanosensors - an overview,” *Talanta*, vol. 67, no. 3, pp. 438–448, 2005.
- [21] I. F. Akyildiz, F. Fekri, R. Sivakumar, C. R. Forest, and B. K. Hammer, “Monaco: fundamentals of molecular nano-communication networks,” *IEEE Wireless Communications Magazine*, vol. 19, pp. 12–18, Oct. 2012.
- [22] E. M. Comelli, S. Lariani, M. C. Zwahlen, G. Fotopoulos, J. A. Holzwarth, C. Cherbut, G. Dorta, I. Corthysy-Theulaz, and M. Grigorov, “Biomarkers of human gastrointestinal tract regions,” *Mamm Genome*, vol. 20, pp. 516–527, Aug. 2009.
- [23] E. Livak-Dahl, I. Sinn, and M. Burns, “Microfluidic chemical analysis systems,” *Annual Review of Chemical and Biomolecular Engineering*, vol. 2, pp. 325–353, Mar. 2011.
- [24] F. J. Holler, D. A. Skoog, and D. M. West, *Fundamentals of analytical chemistry*. Brooks/Cole, Cengage Learning, 9 ed., 2013.
- [25] S. Girotti, E. N. Ferri, M. Fumo, and E. Maiolini, “Monitoring of environmental pollutants by bioluminescent bacteria,” *Anal Chim Acta*, vol. 608, pp. 2–29, Feb. 2008.
- [26] J. Philibert, “One and a half century of diffusion: Fick, Einstein, before and beyond,” *Diffusion Fundamentals*, vol. 4, pp. 6.1–6.19, 2006.
- [27] E. L. Cussler, *Diffusion. Mass Transfer in Fluid Systems*. 2nd edition, Cambridge University Press, 1997.

- [28] M. Moore, A. Enomoto, T. Nakano, R. Egashira, T. Suda, A. Kayasuga, H. Kojima, H. Sakakibara, and K. Oiwa, “A design of a molecular communication system for nanomachines using molecular motors,” in *Proc. of Fourth Annual IEEE Intl. Conf. on Pervasive Computing and Communications Workshops*, pp. 6–12, Mar. 2006.
- [29] M.-J. Moore, T. Suda, and K. Oiwa, “Molecular communication: modeling noise effects on information rate,” *IEEE Transactions on NanoBioscience*, vol. 85, pp. 295–385, June 2009.
- [30] C. Bustamante, D. Keller, and G. Oster, “The physics of molecular motors,” *Accounts of Chemical Research*, vol. 34, no. 6, pp. 412–420, 2001.
- [31] T. Nakano, T. Suda, T. Koujin, T. Haraguchi, and Y. Hiraoka, “Molecular communication through gap junction channels: system design, experiments and modeling,” in *Proc. of 2nd Intl. Conf. on Bio-Inspired Models of Network, Information and Computing Systems (Bionetics)*, pp. 139–146, Dec. 2007.
- [32] M. Gregori and I. F. Akyildiz, “A new nanonetwork architecture using flagellated bacteria and catalytic nanomotors,” *IEEE Journal on Selected Areas in Communications (JSAC)*, vol. 28, pp. 602–611, May 2010.
- [33] M. Gregori, I. Llatser, A. Cabellos-Aparicio, and E. Alarcón, “Physical channel characterization for medium-range nanonetworks using flagellated bacteria,” *Computer Networks*, vol. 55, no. 3, pp. 779–791, 2011.
- [34] M. J. Berridge, “The AM and FM of calcium signalling,” *Nature*, vol. 386, pp. 759–780, Apr. 1997.
- [35] G. Alfano and D. Miorandi, “On information transmission among nanomachines,” in *Proc. of First Intl. Conf. on Nano-Networks and Workshops*, pp. 1–5, Sept. 2006.
- [36] D. J. Spencer, S. K. Hampton, P. Park, J. P. Zurkus, and P. J. Thomas, “The diffusion-limited biochemical signal-relay channel,” *Advances in Neural Information Processing Systems*, vol. 16, 2004.
- [37] B. Atakan and O. B. Akan, “An information theoretical approach for molecular communication,” in *Proc. of Second Intl. Conf. on Bio-Inspired Models of Network, Information and Computing Systems (Bionetics)*, pp. 33–40, Dec. 2007.
- [38] L. Parcerisa and I. F. Akyildiz, “Molecular communication options for long range nanonetworks,” *Computer Networks (Elsevier) Journal*, vol. 53, pp. 2753–2766, Aug. 2009.
- [39] T. Nakano, T. Suda, M. Moore, R. Egashira, A. Enomoto, and K. Arima, “Molecular communication for nanomachines using intercellular calcium signaling,” in *Proc. of Fifth IEEE Intl. Conf. on Nanotechnology*, vol. 2, pp. 478–481, July 2005.
- [40] W. H. Bossert and E. . Wilson, “The analysis of olfactory communication among animals,” *Journal of Theoretical Biology*, vol. 5, pp. 443–469, Nov. 1963.

- [41] S. Kadloor and R. Adve, "A framework to study the molecular communication system," in *Proc. of 18th Intl. Conf. on Computer Communications and Networks*, pp. 1–6, Aug. 2009.
- [42] B. Atakan and O. B. Akan, "On molecular multiple-access, broadcast, and relay channels in nanonetworks," in *Proc. of Third Intl. Conf. on Bio-Inspired Models of Network, Information and Computing Systems (Bionetics)*, pp. 16:1–16:8, Nov. 2008.
- [43] B. Atakan and O. B. Akan, "On channel capacity and error compensation in molecular communication," *Springer Transaction on Computational System Biology*, vol. 10, pp. 59–80, Dec. 2009.
- [44] A. Einolghozati, M. Sardari, A. Beirami, and F. Fekri, "Capacity of discrete molecular diffusion channels," in *Proc. of 2011 IEEE International Symposium on Information Theory (ISIT)*, pp. 603–607, July 2011.
- [45] M. Mahfuz, D. Makrakis, and H. Mouftah, "Characterization of intersymbol interference in concentration-encoded unicast molecular communication," in *Proc. of 24th Canadian Conference on Electrical and Computer Engineering (CCECE)*, pp. 139–146, May 2011.
- [46] M. S. Kuran, H. B. Yilmaz, T. Tugcu, and I. F. Akyildiz, "Interference effects on modulation techniques in diffusion based nanonetworks," *Nano Communication Networks (Elsevier) Journal*, vol. 3, pp. 65–73, Mar. 2012.
- [47] B. Atakan, S. Galmes, and O. B. Akan, "Nanoscale communication with molecular arrays in nanonetworks," *IEEE Transactions on NanoBioscience*, vol. 11, pp. 164–168, June 2012.
- [48] D. Arifler, "Capacity analysis of a diffusion-based short-range molecular nanocommunication channel," *Computer Networks Journal (Elsevier)*, vol. 55, pp. 1426–1434, Apr. 2011.
- [49] J.-P. Rospars, V. Křivan, and P. Lánský, "Perireceptor and receptor events in olfaction. Comparison of concentration and flux detectors: a modeling study," *Chemical Senses*, vol. 25, pp. 293–311, June 2000.
- [50] M. Reed and R. Rohrer, *Applied introductory circuit analysis for electrical and computer engineers*. Prentice-Hall, Inc. Upper Saddle River, NJ, USA, 1999.
- [51] B. Davies, *Integral transforms and their applications*. Springer, New York, 2002.
- [52] L. Clancy, *Aerodynamics*. Pitman Publishing Limited, 1975.
- [53] E. Eckert and R. Drake, *Analysis of Heat and Mass Transfer*. McGraw-Hill, 1972.
- [54] Y. Ali and L. Zhang, "Relativistic heat conduction," *International Journal of Heat and Mass Transfer*, vol. 48, pp. 2397–2406, Mar. 2005.

- [55] G. Arfken, “Green’s functions—two and three dimensions,” *Mathematical Methods for Physicists*, pp. 480–491, 1985.
- [56] R. Aleixo and E. C. de Oliveira, “Greens function for the lossy wave equation,” *Revista Brasileira de Ensino de Fisica*, vol. 30, no. 1, p. 1302, 2008.
- [57] A. Sezginer and W. C. Chew, “Closed form expression of the greens function for the time-domain wave equation for a lossy two-dimensional medium,” *IEEE Transactions on Antennas and Propagation*, vol. 32, no. 5, pp. 527–528, 1984.
- [58] B. S. Donahue and R. F. Abercrombie, “Free diffusion coefficient of ionic calcium in cytoplasm,” *Cell Calcium*, vol. 8, pp. 437–48, Dec. 1987.
- [59] M. A. Model and G. M. Omann, “Ligand-receptor interaction rates in the presence of convective mass transport,” *Biophysical Journal*, vol. 69, no. 5, pp. 1712–1720, 1995.
- [60] G. Vasilescu, *Electronic Noise and Interfering Signals: Principles and Applications*. Springer Series in Signals and Communication Technology, Springer, 2005.
- [61] H. Berg and E. Purcell, “Physics of chemoreception,” *Biophysical Journal*, vol. 20, pp. 193–219, Nov. 1977.
- [62] W. Bialek and S. Setayeshgar, “Physical limits to biochemical signaling,” *Proceedings of the National Academy of Sciences (PNAS) of the USA*, vol. 102, pp. 10040–10045, July 2005.
- [63] K. Prank, F. Gabbiani, and G. Brabant, “Coding efficiency and information rates in transmembrane signaling,” *Biosystems*, vol. 55, pp. 15–22, Feb. 2000.
- [64] J.-Q. Liu and T. Nakano, “An information theoretic model of molecular communication based on cellular signaling,” in *Proc. of Second Intl. Conf. on Bio-Inspired Models of Network, Information and Computing Systems (Bionetics)*, pp. 316–321, Dec. 2007.
- [65] H. Nyquist, “Certain topics in telegraph transmission theory,” *Reprint as classic paper in: Proc. IEEE*, vol. 90, pp. 280–305, Feb. 2002.
- [66] A. Papoulis and S. U. Pillai, *Probability, Random Variables and Stochastic Processes, 4th ed.* McGraw-Hill, 2002.
- [67] D. T. Gillespie, “Stochastic simulation of chemical kinetics,” *Annual Review of Physical Chemistry*, vol. 58, pp. 35–55, May 2007.
- [68] D. A. McQuarrie, “Kinetics of small systems. I,” *The Journal of Chemical Physics*, vol. 38, pp. 433–436, Jan. 1963.
- [69] C. E. Shannon, “A mathematical theory of communication,” *The Bell System Technical Journal*, vol. 27, pp. 379–423, July 1948.

- [70] B. Atakan and O. B. Akan, “Deterministic capacity of information flow in molecular nanonetworks,” *Nano Communication Networks Journal (Elsevier)*, vol. 1, pp. 31–42, Mar. 2010.
- [71] A. Einolghozati, M. Sardari, and F. Fekri, “Capacity of diffusion-based molecular communication with ligand receptors,” in *Proc. of 2011 IEEE Information Theory Workshop (ITW)*, pp. 85–89, Oct. 2011.
- [72] A. Einolghozati, M. Sardari, and F. Fekri, “Collective sensing-capacity of bacteria populations,” in *Proc. of 2011 IEEE International Symposium on Information Theory (ISIT)*, pp. 2959–2963, July 2012.
- [73] A. Einolghozati, M. Sardari, and F. Fekri, “Molecular communication between two populations of bacteria,” in *Proc. of 2012 IEEE Information Theory Workshop (ITW)*, pp. 437–441, Sept. 2012.
- [74] E. Schrödinger, *Statistical thermodynamics*. Dover Publications New York, 1989.
- [75] E. T. Jaynes, “Information theory and statistical mechanics,” *The Physical Review*, vol. 106, pp. 620–630, May 1957.
- [76] A. Ben-Naim, *A Farewell to Entropy: Statistical Thermodynamics Based on Information*. World Scientific Publishing Company, 2008.
- [77] T. M. Cover and J. A. Thomas, *Elements of Information Theory, 2nd Edition*. Wiley, 2006.
- [78] R. A. Freitas, “Pharmacytes: An ideal vehicle for targeted drug delivery,” *Journal of Nanoscience and Nanotechnology*, vol. 6, pp. 2769–2775, 2006.
- [79] A. Einstein, “On the electrodynamics of moving bodies,” *Annalen der Physik*, vol. 17, pp. 891–921, 1905.
- [80] E. Fermi, *Thermodynamics*. Prentice-Hall Company, 1936.
- [81] J. W. Gibbs, H. A. Bumstead, and R. G. V. Name, *Scientific Papers of J. Willard Gibbs, Vol. 1: Thermodynamics*. Longmans, Green and Company, 1961.
- [82] M. L. McGlashan, *Physico-Chemical Quantities and Units*. London: Royal Institute of Chemistry, 1968.
- [83] R. P. Feynman, “There’s plenty of room at the bottom,” *Journal of Microelectromechanical Systems*, vol. 1, pp. 60–66, 1992.
- [84] M. P. Langevin, “Paul langevins 1908 paper on the theory of brownian motion,” *American Journal of Physics*, vol. 65, pp. 1079–1081, Nov. 1997.
- [85] M. S. Bartlett, *An introduction to stochastic processes, with special reference to methods and applications*. Press Syndicate of the University of Cambridge, 1978.

- [86] A. Mandelis, *Diffusion-wave fields: mathematical methods and Green functions*. Springer-Verlag, 2001.
- [87] A. Gelman, J. B. Carlin, H. S. Stern, and D. B. Rubin, *Bayesian Data Analysis, Second Edition*. Chapman and Hall/CRC, 2003.
- [88] M. Merkle, “Convexity in the theory of the gamma function,” *International Journal of Applied Mathematics & Statistics*, vol. 11, pp. 103–117, Nov. 2007.
- [89] J. M. Bernardo, “Psi (digamma) function,” *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, vol. 25, no. 3, pp. 315–317, 1976.
- [90] S. S. Zumdahl, *Thermochemistry*. Chemistry. Cengage Learning., 2008.
- [91] E. Fermi, *Thermodynamics*. New York: Dover Publications, 1956.
- [92] M. S. Kuran, H. B. Yilmaz, T. Tugcu, and I. F. Akyildiz, “Interference effects on modulation techniques in diffusion based nanonetworks,” *Nano Communication Networks (Elsevier) Journal*, vol. 3, pp. 65–73, Mar. 2012.
- [93] J. Ilow and D. Hatzinakos, “Analytic alpha-stable noise modeling in a poisson field of interferers or scatterers,” *IEEE Transactions On Signal Processing*, vol. 46, pp. 1601–1611, June 1998.
- [94] D. T. Gillespie, “The chemical langevin equation,” *Journal of Chemical Physics*, vol. 113, pp. 297–306, July 2000.
- [95] K. Gulati, B. L. Evans, J. G. Andrews, and K. R. Tinsley, “Statistics of co-channel interference in a field of poisson and poisson-poisson clustered interferers,” *IEEE Transactions on Signal Processing*, vol. 58, pp. 6207–6222, Dec. 2010.
- [96] A. Mandelis, “Diffusion waves and their uses,” *Physics Today*, vol. 53, pp. 29–34, Aug. 2000.
- [97] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. New York: Dover Publications, 1972.
- [98] J. S. Simonoff, *Smoothing Methods in Statistics*. Springer Series in Statistics, Springer-Verlag New York, 1998.
- [99] C. J. Myers, *Engineering genetic circuits*. Chapman & Hall/CRC, Mathematical and Computational Biology Series, 2009.
- [100] S. Mukherji and A. van Oudenaarden, “Synthetic biology: understanding biological design from synthetic circuits,” *Nature Reviews Genetics*, vol. 10, pp. 859–871, Dec. 2009.
- [101] B. Canton, A. Labno, and D. Endy, “Refinement and standardization of synthetic biological parts and devices,” *Nature Biotechnology*, vol. 26, pp. 787–793, July 2008.

- [102] U. Alon, *An Introduction to Systems Biology - Design Principles of Biological Circuits*. Chapman & Hall/CRC, 2006.
- [103] S. S. Andrews, T. Dinh, and A. P. Arkin, *Stochastic models of biological processes*, vol. 9. Meyers, Robert (Ed.), Springer, NY, 2009.
- [104] C. D. Cox, J. M. McCollum, D. W. Austin, M. S. Allen, R. D. Dar, and M. L. Simpson, “Frequency domain analysis of noise in simple gene circuits,” *Chaos.*, vol. 16, p. 026102, June 2006.
- [105] M. L. Simpson, C. D. Cox, M. S. Allen, J. M. McCollum, R. D. Dar, D. K. Karig, and J. F. Cooke, “Noise in biological circuits,” *Wiley Interdiscip Rev Nanomed Nanobiotechnol.*, vol. 1, pp. 214–225, Mar. 2009.
- [106] J. E. Ladbury and S. T. Arold, “Noise in cellular signaling pathways: causes and effects,” *Trends Biochem Sci.*, vol. 37, pp. 173–178, May 2012.
- [107] C. Bustamante, Y. Chelma, N. Forde, and D. Izhaky, “Mechanical processes in biochemistry,” *Annual Review of Biochemistry*, vol. 73, pp. 705–748, 2004.
- [108] W. C. Fuqua, S. C. Winans, and E. P. Greenberg, “Quorum sensing in bacteria: the luxR-luxI family of cell density-responsive transcriptional regulators,” *J Bacteriol.*, vol. 176, pp. 269–275, Jan. 1994.
- [109] A. B. Goryachev, D. J. Toh, and T. Lee, “Systems analysis of a quorum sensing network: Design constraints imposed by the functional requirements, network topology and kinetic constants,” *Biosystems*, vol. 83, pp. 178–187, Feb. 2006.
- [110] A. L. Schaefer, B. L. H. D. L. Val, J. J. E. Cronan, and E. P. Greenberg, “Generation of cell-to-cell signals in quorum sensing: Acyl homoserine lactone synthase activity of a purified *vibrio fischeri* luxI protein,” *Proc Natl Acad Sci U S A*, vol. 93, pp. 269–275, Sept. 1996.
- [111] J. Garcia-Ojalvo, M. B. Elowitz, and S. H. Strogatz, “Modeling a synthetic multicellular clock: repressilators coupled by quorum sensing,” *Proc Natl Acad Sci U S A*, vol. 101, pp. 10955–60, July 2004.
- [112] M. J. Pilling and P. W. Seakins, *Reaction Kinetics*. Oxford University Press, 1995.
- [113] F. W. Sears, G. L. Salinger, and J. E. Lee, *Thermodynamics, Kinetic Theory, and Statistical Thermodynamics*. Addison-Wesley, 1975.
- [114] K. J. Laidler and M. C. King, “Development of transition-state theory,” *The Journal of Physical Chemistry*, vol. 87, pp. 2657–2664, July 1983.

## VITA

Massimiliano Pierobon was born on June, 21, 1979, in Bollate (Milan), Italy. He received the Master of Science (B.S.+M.S.) degree in Telecommunication Engineering from the Politecnico di Milano, Milan, Italy, in 2005. During 2006, Massimiliano Pierobon worked as a researcher in the R&D department of Siemens Carrier Networks, Milan, where he coauthored two filed patents on jitter buffer management. From January 2007 to July 2009 he was a graduate research assistant at the Politecnico di Milano in the fields of signal processing and pattern recognition. In November 2008 Massimiliano Pierobon joined the BWN lab, first as a visiting researcher and, from August 2009, as a Georgia Tech Ph.D. student. Massimiliano Pierobon received the BWN Lab Researcher of the Year Award at the Georgia Institute of Technology for his outstanding research achievements in 2011. He is a member of IEEE, ACM, and ACS. His current research interests are in molecular communication theory for nanonetworks, communication engineering applied to intelligent drug delivery systems and biological circuit network engineering.