RELIABLE AND EFFICIENT COMMUNICATION IN WIRELESS UNDERGROUND SENSOR NETWORKS

A Thesis
Presented to
The Academic Faculty

by

Zhi Sun

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

Georgia Institute of Technology
August 2011

RELIABLE AND EFFICIENT COMMUNICATION IN WIRELESS UNDERGROUND SENSOR NETWORKS

Approved by:

Professor Ian F. Akyildiz, Advisor
School of Electrical and Computer Engineering
Georgia Institute of Technology

Professor Ying Zhang
School of Electrical and Computer Engineering
Georgia Institute of Technology

Professor Geoffrey Ye Li
School of Electrical and Computer Engineering
Georgia Institute of Technology

Professor Mostafa H. Ammar
College of Computing
Georgia Institute of Technology

Professor Mary Ann Ingram
School of Electrical and Computer Engineering
Georgia Institute of Technology

Date Approved: June 3rd 2011

To my family,

for their endless love and support.

ACKNOWLEDGEMENTS

I would like to express my sincere thanks to my advisor Dr. Ian F. Akyildiz for giving me the opportunity to work as his student, and for his guidance, support, and encouragement during the entire Ph.D program. I am grateful to him for his trust, patience, and constructive criticisms as well as rewarding praises, which helped me improve the quality of my research. Dr. Akyildiz not only trained me to be a good researcher, he has also taught me in many other ways that could lead me to success in my future career.

I wish to express my gratitude to all the academic members of the Electrical and Computer Engineering Department at the Georgia Institute of Technology for their excellent advice, constructive criticism, helpful and critical reviews throughout the Ph.D. program. A special thank goes to Dr. Ye (Geoffrey) Li, Dr. Mary Ann Ingram, Dr. Ying Zhang, and Dr. Mostafa Ammar, who kindly agreed to serve in my Ph.D. Defense Committee. Their invaluable comments and enlightening suggestions have helped me to achieve a solid research path towards this thesis.

I would also like to thank all former and current members of the Broadband Wireless Networking Laboratory (BWN Lab) for their support and friendship. The excellent family-like atmosphere they created gave me a warm memory during my Ph.D student life.

Finally, I would like to express my deep gratitude to my parents and my wife, for their patience, continuous support and encouragement throughout this thesis.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

Wireless Underground Sensor Networks (WUSNs) are the networks of wireless sensors that operate below the ground surface. These sensors are either buried completely in soil medium, or placed within a bounded open underground space, such as underground mines and tunnels. WUSNs enable a wide variety of novel applications, including intelligent irrigation, underground structure monitoring, and border patrol and intruder detection.

This thesis is concerned with establishing reliable and efficient communications in the network of wireless sensor nodes that are deployed in either soil medium or underground mines and tunnels. The unique characteristics of the WUSNs in different underground channels are first analyzed. Then and the communication and networking solutions are developed based on the understanding of the underground channels.

In particular, to realize WUSNs in soil medium, two types of signal propagation techniques including Electromagnetic (EM) waves and Magnetic Induction (MI) are explored. For EM wave-based WUSNs, the heterogeneous network architecture and dynamic connectivity are investigated based on a comprehensive channel model in soil medium. Then a spatio-temporal correlation-based data collection schemes is developed to reduce the sensor density while keeping high monitoring accuracy. For MI-based WUSNs, the MI channel is first analytically characterized. Then based on the MI channel model, the MI waveguide technique is developed in order to enlarge the underground transmission range. Finally, the optimal deployment algorithms for MI waveguides in WUSNs are analyzed to construct the WUSNs with high reliability and low costs.

To realize WUSNs in underground mines and tunnels, a mode-based analytical channel model is first proposed to accurately characterize the signal propagation in both empty and obstructed mines and tunnels. Then the Multiple-Input and Multiple-Output (MIMO) system and cooperative communication system are optimized to establish reliable and efficient communications in underground mines and tunnels.

CHAPTER I

INTRODUCTION

1.1 Backgroud

Wireless Underground Sensor Networks (WUSNs) [2] are the networks of wireless sensor nodes operating below the ground surface. As a natural extension to the well-established wireless sensor networks (WSNs) [3] paradigm, WUSNs are envisioned to provide real-time monitoring capabilities in two types of underground environments: soil medium and underground mines and tunnels. Based on the monitored environments, the WUSNs can be further divided into two categories: the WUSNs in soil medium and the WUSNs in underground mines and tunnels. In the former case, networks of wireless nodes are buried underground and communicate through soil. In the latter case, although the network is located underground, the communications take place through the air, i.e., through the voids that exist underground. Compared with existing underground monitoring strategies, WUSNs have the advantages in timeliness of data, ease of deployment and data collection, concealment, reliability, and coverage density [2].

A wide variety of novel and essential applications are enabled by WUSNs [2, 91, 92], including:

- **Intelligent Irrigation:** With the real time monitoring of the soil moisture, temperature, among other soil properties, the WUSNs can accurately determine when and where to irrigate the crops. Considering that the irrigations constitute more than 70% fresh water consumption all over the world [1], the WUSNs can greatly enhance the water sustainability.

- **Mine Disaster Prevention and Rescue:** No existing techniques support communications and localization after mine disasters, especially when RF wireless channel

is blocked due to tunnel collapses and wired communication is cut off due to cable damages. Since the WUSN is able to work in harsh underground environment, it can greatly enhance current mine safety and productivity.

- **Earthquake and Landslide Monitoring:** Up to now, earthquake and landslide are still difficulty to be accurately predicted. WUSNs provide us a novel way to monitor the signs of earthquake and landslides in real time with small deployment and maintenance cost. As a result, the personal injury and property loss caused by those natural disasters can be minimized.

- **Underground Pipeline and Power Grid Monitoring:** Underground pipelines constitute one of the most important ways to transport large amounts of fluid (e.g. oil and water) through long distances. However, existing leakage detection techniques do not work well due to the harsh underground environmental conditions. Moreover, in current underground power grid, various faults, such as underground power gird fire caused by overloading and cable break caused by conductor theft or careless digging, are difficult to be avoided, detected, localized, and fixed due to the inaccessible environments. The WUSNs can provide real time monitoring to help the administrators prevent the potential faults and fix existing faults in those underground structures.

- **Border Patrol and Intruder Detection:** Border Patrol is important for national security. The conventional border patrol systems suffer from intensive human involvement. WUSNs deployed along the border provide a low cost, reliable, and concealed way to detect the intruders crossing the border.

Despite the potential advantages of WUSNs, the underground environment is a hostile place for wireless communication and requires existing networking solutions and communication protocols for terrestrial WSNs be reexamined. Specifically, the key difference between the WUSNs and the terrestrial WSNs is the communication medium. For the WUSNs deployed in soil, the propagation medium is no longer air but soil, rock and water.

Although the well established terrestrial signal propagation techniques based on electromagnetic (EM) waves may still work in soil medium, the unique channel characteristics of EM waves in this environments needs to be modeled. Besides EM waves, alternate signal propagation techniques, such as magnetic induction (MI) can also be used for wireless communications in soil and need to be investigated. For the WUSNs deployed in underground mines and road/subway tunnels, the EM waves are suitable for wireless signal propagation, since the radio signal propagates through the air in this case. However, the propagation characteristics of EM waves are significantly different from those of terrestrial wireless channels due to the restrictions caused by the walls and ceilings in mines and tunnels. Moreover, since different physical layer techniques have to be developed to solve the challenges brought by the harsh underground environments, the corresponding higher layers of the protocol stack also need to be redesigned.

1.2 Research Objectives and Solutions

The objective of this thesis is to analyze the unique characteristics of the WUSNs in different underground environments and to find out the solutions to realize the reliable and efficient communication in WUSNs. For WUSNs in soil medium, we develop two types of WUSNs based on either EM wave techniques or MI technique to overcome the unique challenges brought by the soil transmission medium. For WUSNs in underground mines and tunnels, we utilize the Multiple-Input and Multiple-Output (MIMO) system and cooperative communication system to establish reliable and efficient communications.

1.2.1 EM Wave-based WUSNs in Soil Medium

In soil medium, the well established wireless communication techniques using EM waves do not work well [4]. First, EM waves experience high levels of attenuation due to the absorption by soil, rock, and water in the soil medium. Since the underground sensor devices have limited radio power due to the energy constraint, the transmission range between two sensor nodes is extremely small (no more than 4 meters). Second, the path loss of the

EM waves in soil medium is highly dependent on numerous soil properties such as water content, soil makeup (sand, silt, or clay), and density. Those soil properties can change dramatically with time (e.g., soil water content increases after a rainfall) and location (e.g., soil properties change dramatically over short distances). Consequently, the transmission range of the underground sensors also varies dramatically in different times and positions.

Besides the communication channel between underground sensors, the channels between underground (UG) sensor nodes and aboveground (AG) data sinks also needs to be analyzed. Hence, three types of channels exist in WUSNs in soil medium, including: underground-to-underground (UG-UG) channel, underground-to-aboveground (UG-AG) channel, and aboveground-to-underground (AG-UG) channel. For the UG-AG channel, the transmission range is much longer than the UG-UG channel [16, 98, 79, 78]. This is because a large portion of the radiation energy can penetrate the air-ground interface from the soil to the air, and the path loss in the air is much smaller than that in the soil. For the AG-UG channel, the transmission range is much smaller than the UG-AG channel since most of the radiation energy is reflected back when penetrating the air-ground interface from the air to the soil. Similar to the UG-UG channel, the transmission ranges of the UG-AG and AG-UG channel are also dramatically influenced by many environmental conditions and system configurations, including soil water content, soil composition, UG sensor burial depth, AG sink antenna height, and signal operating frequency [2, 4, 53, 79, 78].

The complex characteristics of the UG-UG, UG-AG, and AG-UG channel create unique challenges in the design of WUSNs in soil medium.

- First, in the envisioned applications of WUSNs in soil medium, the underground sensor nodes are expected to transmit sensing data to one or multiple aboveground data sinks via single or multi-hop paths. Hence, the connectivity in WUSNs is essential for the system functionalities. Because of the complex channel characteristics, the connectivity analysis in the WUSNs is much more complicated than in the terrestrial wireless sensor networks and ad hoc networks.

- Moreover, the number of underground sensors is expected to be as small as possible due to the high deployment/maintenance cost. However, an extremely high density of underground sensors is required to maintain the full connectivity of WUSNs due to the harsh underground channel conditions. This conflict constitutes one of the greatest challenges to deploy the WUSNs.

In this research, we first quantitatively model the channel characteristics of the three types of channels of WUSNs in soil medium. Based on the channel model, we propose a heterogeneous network architecture and analyze the dynamic connectivity of such network that captures the influence of multiple system and environmental parameters. Moreover, we introduce aboveground mobile sinks to WUSNs and developed a spatio-temporal correlation-based data collection scheme, which significantly reduces the sensor density while keeping high monitoring accuracy. Finally, we propose a theoretical method to determine the optimal sensor density under the proposed scheme, which provides principles and guidelines for the design and deployment of WUSNs.

1.2.2 MI-based WUSNs in Soil Medium

As discussed previously, the EM wave-based techniques encounter two major problems in soil medium: the high path loss and the dynamic channel condition. If the sensors of WUSNs are buried in the shallow depth, sensor can communicate with the aboveground data sinks directly using EM waves since the UG-AG channel has relatively large communication range. However, many WUSN applications, such as underground structure monitoring, require the sensors buried deep underground, where only UG-UG channel is available.

MI is a promising alternative physical layer technique for WUSNs in deep burial depth. Using MI technique could have several benefits. One of these is that the underground medium such as soil and water cause little variation in the attenuation rate of magnetic fields from that of air, since the magnetic permeabilities of each of these materials are

similar [2]. This fact guarantees that the MI channel conditions remain constant for a certain path in different times. However, MI is generally unfavorable for terrestrial wireless communication, since the magnetic field strength falls off much faster than the EM waves in terrestrial environments. In soil medium, although it is known that the soil absorption causes high signal attenuation in the EM waves systems but does not affect the MI systems, it needs to be analyzed whether the total path loss of the MI system is lower than that of the EM waves system or not.

In this research, we conduct detailed analysis on the path loss and the bandwidth of the MI system in underground soil medium. Based on the channel analysis, we develop the MI waveguide technique in order to reduce the high path loss of the traditional EM wave system and the ordinary MI system. By utilizing the passive relay coils, the MI waveguide system dramatically increases the the transmission range of underground sensors in soil medium. Moreover, we analyze the deployment strategies of MI waveguides in WUSNs. We develop optimal deployment algorithms to use the MI relay coils to connect the underground sensors. The proposed algorithm provides guidelines to deploy MI-based WUSNs with high reliability and low costs.

1.2.3 WUSNs in Underground Mines and Tunnels

The WUSNs in underground mines and tunnels are necessary to improve the safety and productivity in mines, to realize intelligent transportation system in road/subway tunnels, and to avoid attacks by monitoring these vulnerable areas.

In underground mines and tunnels, wireless networking using EM waves propagation is a more flexible and efficient solution than the wire-based or leaky coaxial cable guided systems [29] because it is low-cost, easy to implement, and scalable. However, radio waves do not propagate well in underground mines and road tunnels [33]. Due to the reflections of the EM waves on the tunnel walls, the multipath fading in these environments is much more significant than in the terrestrial wireless channels. Moreover, the tunnels in operation are

filled with mobile vehicles with random size and positions. The reflections and the diffractions on the vehicles make the wireless channel in the tunnel even more complicated. To setup reliable and efficient WUSNs in underground mines and tunnels, the analytical channel model that explicitly contains the dependence on the tunnel geometry, vehicular traffic information, and other communication parameters is needed. After the channel model in underground mines and tunnels is derived, suitable communication protocols can be developed to solve the impact of the multi-path fading in these environments.

In this research, we first developed a mode-based analytical channel model that can accurately characterize the signal propagation in empty mines and tunnels. Then we analyze the influence of the vehicular traffic flow on the signal propagation in mines and tunnels by utilizing the uniform theory of diffraction (UTD) [48] and the traffic flow theory [37]. Based on the signal propagation model in mines and tunnels, we analyzed the capacity distribution and outage behavior of MIMO and cooperative systems in such environments. Finally, we developed an optimal antenna geometry design strategy for MIMO system and an optimal relay assignment protocol for cooperative system. With these optimizations, significantly higher spectral efficiency and link reliability are achieved in underground mines and tunnels.

1.3 Thesis Outline

This thesis is organized as follows. In Chapter 2, the EM wave-based WUSNs in soil medium are developed. In particular, the models of the three types of channels, i.e., UG-UG channel, UG-AG channel, and AG-UG channel, are first developed. Then based on the channel model, the network architecture and the dynamic connectivity in EM wave-based WUSNs in soil are investigated. At the end of this chapter, a spatio-temporal correlation-based data collection scheme is developed for WUSNs in soil medium. In Chapter 3, the MI-based WUSNs in soil medium are introduced. Specifically, the MI channel model for

WUSNs in soil medium is first provided. Then, the MI waveguides are developed to significantly enlarge the UG-UG communication range. At the end of this chapter, the optimal deployment algorithms for MI waveguide are presented. In Chapter 4, the WUSNs in underground mines and tunnels are explored. Particularly, the channel model of the WUSNs in empty and obstructed mines and tunnels are first derived. Based on the channel model, the MIMO and cooperative communication solutions for WUSNs in underground mines and tunnels are proposed to mitigate the severe multipath fading problem. Finally, Chapter 5 summarizes the research contributions and identifies several future research directions.

CHAPTER II

EM WAVE-BASED WUSNS IN SOIL MEDIUM

2.1 Motivation and Related Work

The EM wave-based wireless signal propagation technique is widely adopted in existing wireless communications and networks. The underground soil medium brings unique characteristics of the wireless channel using EM waves. Three types of channels with dramatically different transmission ranges, including UG-UG channel, UG-AG channel, and AG-UG channel are introduced.

According to the EM wave channel characteristics and the envisioned applications of WUSNs in soil medium, a practical WUSN network consists of UG sensors deployed in the sensing field, fixed AG data sinks set around the sensing field, and a small number of mobile data sinks carried by people or machineries inside the sensing field. Specifically, if there is only one single AG data sink, a prohibitively high density of UG sensors is required to guarantee the full connectivity, due to the small and dynamic transmission range of the UG-UG channel. If the cost of deployment and maintenance is considered, the extremely high density of UG sensors is unacceptable. To solve this problem, multiple AG data sinks have to be introduced [2, 16]. Since the transmission range of the UG-AG channel is much larger than the UG-UG channel, the WUSNs can be connected with much lower UG sensor density if multiple AG data sinks are employed. The AG data sinks can be either fixed or mobile. Fixed AG data sinks are deployed at random positions inside the monitored field, while mobile AG data sinks can be handsets that are carried by people or machineries working inside the monitored field. The mobile sink moves randomly in the monitored field, and collects data from the UG sensors when moving into their transmission range. Therefore, if the WUSN applications can tolerate a certain level of latency, the isolated UG

sensors can expect a mobile sink coming and collecting their data.

In WUSN applications, real-time underground environmental conditions in different locations are monitored, collected, and processed by the WUSNs to achieve the application goals. The underground sensor nodes are envisioned to send the measurements to the aboveground data sinks to guarantee a certain level of monitoring accuracy. To achieve this requirement, the network connectivity and corresponding data collection scheme should be investigated for the EM wave-based WUSNs in soil medium.

According to the above discussion, the connectivity analysis in WUSNs is a complicated problem since the network consists of three types of wireless nodes (UG sensors, AG fixed sinks and AG mobile sinks) in two different mediums (soil and air) with three different transmission ranges (UG-UG, UG-AG and AG-UG). In addition, the connectivity in WUSNs is highly dynamic due to the dynamic underground channel characteristics and the random movement of the mobile sinks. First, the transmission ranges of the three types of channels are all highly dynamic due to the changes of the environmental conditions as well as the sensor burial depth and sink antenna height. Consequently, the network connectivity varies in different time and locations. Second, although the mobile AG sinks can improve the network connectivity, the random movement also bring fluctuations of the network connectivity. The tradeoff between the good connectivity and the low latency needs to be analyzed. Moreover, since the channels between AG and UG devices are asymmetrical, the network connectivity is also asymmetrical.

Besides network connectivity, the data collection scheme is also need to be designed to address the unique challenges in WUSNs. Specifically, one of the greatest barrier in designing WUSNs is the conflict between the high deployment cost of underground sensors and the high underground sensor density required to achieve fully connected network. On the one hand, since each sensor needs to be buried underground, the deployment and maintenance costs are extremely high compared with terrestrial sensor networks. Hence the sensor density should be minimized. On the other hand, due to the material absorption

in soil medium, the communication range between underground sensors is very limited (≤ 4 m) [4]. Consequently, a prohibitively high density of underground sensors (nearly 1 sensor per m^2) is required to guarantee the network connectivity [87]. Moreover, the highly dynamic soil water content significantly affects sensor's communication range. As a result, the network connectivity of WUSNs is not guaranteed even with high underground sensor density. To solve the above conflict, the data collection scheme of the WUSNs needs to be reconsidered. Specifically, it may be not necessary to collect the measurements of every sensor at every time stamp since the measured data over an area is usually spatio-temporally correlated [39]. Then the requirements of the network connectivity in the WUSNs can be lowered so that the number of the underground sensors can be reduced. The bottom line of the WUSNs is to achieve satisfying monitoring/estimation accuracy of every position at every time stamp.

Although a few recent papers are specifically concerned with the communication problems of WUSNs in soil medium, the literature on the subject is extremely limited. In [2], application scenarios and research challenges of the WUSNs in soil medium are discussed, and open research issues are described. In [4, 53, 100], the channel characteristics of EM waves in soil medium are investigated. The analysis shows that the path loss is much higher than the terrestrial case due to the material absorption. In addition, the communication success significantly depends on the composition of the soil and the operating frequency. The feasible transmission range of the underground sensors in soil medium is no more than 4 meters. The theoretical analysis of [4, 53, 100] is validated by field experiments in [79, 78].

To date, no existing work has analyzed the connectivity problems in WUSNs. However, the connectivity in the homogenous ad hoc networks has been well analyzed. In [40], the necessary and sufficient scaling of the transmission range is analyzed to achieve the full connectivity. In [13], the upper bound of the connectivity probability is proposed as a function of the node density. Comprehensive simulation results for the connectivity in mobile ad

hoc network are provided in [74]. The above connectivity analysis are based on the deterministic disk shaped model. In [14], the impacts of large scale lognormal shadowing on the network connectivity are analyzed. In[72], the network connectivity is investigated in the presence of both large scale fading and small scale fading as well as the unreliable nodes. In [47], the dynamic connectivity caused by unreliable links is analyzed. All the above works are based on the homogenous network architecture with only one types of nodes, which is much simpler than the case in WUSNs where three types of wireless devices are deployed in two types of mediums. Moreover the simple terrestrial channel models cannot characterize the complex channels among devices in both underground and aboveground. The connectivity of ad hoc networks with a heterogeneous network architecture is analyzed in [32]. It is proved that the connectivity of ad hoc networks can be improved by deploying base stations under certain conditions. This result is also suitable in wireless sensor networks by replacing the base stations by the data sinks. In [31], the connectivity in a sensor network with node sleeping scheme is analyzed. However, the authors assume that only one data sink exists. Therefore the connectivity criteria is the same as in the ad hoc networks. In [35], multiple sinks are considered in the connectivity analysis in wireless sensor networks. However, the authors assume that the sensors can be connected to the sinks only in a single-hop fashion, which is not true in most multi-hop wireless sensor networks. The above works are based on the determined terrestrial channel model and do not consider the possible connectivity improvement introduced by mobile data sinks.

The spatio-temporal correlations have been widely used in the environmental monitoring. In [99], the spatio-temporal correlations in wireless sensor networks are exploited to improve the performance of communication protocols. In [43], a simplified spatio-temporal soil moisture model is proposed. This model is utilized to design the WUSN in [30]. These works assume that the sensor networks are fully connected so that all measurements are available at the monitoring center. However, due to the harsh underground channel condition and the high deployment cost of underground sensors, the fully connected network is

difficult to achieve in WUSNs. In [66], the spatio-temporal planning is summarized as a problem of the cooperative control of multiple mobile robots. As an example of the spatio-temporal planning in sensor networks, an event collection scheme using a single mobile sink is developed in [106]. The authors assume that all the sensors are isolated and can only communicate with the mobile sink. However, in WUSNs, although the network is not necessarily to be fully connected, there may still exist connections between adjacent underground sensors. The multiple aboveground mobile sinks can either communicate with a single sensor or a cluster of sensors. Moreover, the connectivity is subject to change due to the dynamic underground channel conditions.

In this chapter, we first extend the underground channel model provided in [4, 53, 100] and quantitatively analyze the characteristics of all the three types of channels in WUSNs, including the UG-UG, UG-AG, and AG-UG channel. Then based on the channel model, we investigate the dynamic network connectivity of the WUSNs in soil medium. A mathematical framework to determine the lower and upper bounds of the connectivity probability in WUSNs is developed, which analytically captures the effects of the density and distribution of both the UG sensors and the AG fixed sinks, the number and mobility of the AG mobile sinks, the soil properties especially the dynamic soil moisture, the UG sensor burial depth, the AG sink antenna height, the tolerable latency of the envisioned application, the radio power, and the system operating frequency. Finally, we develop a spatio-temporal correlation-based data collection scheme to reduce the WUSN deployment cost while maintaining satisfying monitoring accuracy. The optimal sensor density in WUSNs is also derived by jointly analyzing the underground channel characteristics, the spatio-temporal correlation, the dynamic network connectivity, and the random or controlled mobility of multiple mobile sinks.

2.2 Channel Modeling of EM Waves in Soil Medium

As discussed in Chapter 1, the complex channel characteristics of the UG-UG channel, the UG-AG channel, and the AG-UG channel constitute one of the major challenges in the connectivity analysis in WUSNs. We have developed the channel model for UG-UG channel in our previous works [4, 53]. In this section, we extend this channel model to characterize all the three types of channels and provide the formulas to calculate the transmission ranges of those channels. Since the WUSNs are mainly deployed in spacious fields (e.g. crop field or sports field), the multi-path fading effects can be ignored.

2.2.1 UG-UG Channel

The channel model for UG-UG channel proposed in [4, 53] is first overviewed. Assuming that $L_{UG}(d)$ is the signal loss of an underground soil path with length d (meters), then

$$L_{UG}(d) = 6.4 + 20\log d + 20\log\beta + 8.69\alpha d \ , \tag{1}$$

where α is the attenuation constant with the unit of $1/m$, and β is the phase shifting constant with the unit of $radian/m$. The values of α and β depend on the dielectric properties of soil:

$$\begin{aligned} \alpha &= 2\pi f \sqrt{\frac{\mu\epsilon'}{2}\left[\sqrt{1+(\frac{\epsilon''}{\epsilon'})^2}-1\right]}, \\ \beta &= 2\pi f \sqrt{\frac{\mu\epsilon'}{2}\left[\sqrt{1+(\frac{\epsilon''}{\epsilon'})^2}+1\right]}, \end{aligned} \tag{2}$$

where f is the operating frequency, μ is the magnetic permeability, ϵ' and ϵ'' are the real and imaginary parts of the relative dielectric constant of soil medium:

$$\begin{aligned} \epsilon' &= 1.15[1+\frac{\rho_b}{\rho_s}(\epsilon_s^{\alpha'}) + m_v^{\beta'}\epsilon_{f_w}^{\prime\alpha'} - m_v]^{1/\alpha'} - 0.68, \\ \epsilon'' &= [m_v^{\beta''}\epsilon_{f_w}^{\prime\prime\alpha'}]^{1/\alpha'} \ , \end{aligned} \tag{3}$$

where m_v is the volumetric water content (VWC) of the soil medium, ρ_b is the bulk density, $\rho_s = 2.66\ g/cm^3$ is the specific density of the solid soil particles, $\alpha' = 0.65$ is an empirically determined constant, ϵ'_{f_w} and ϵ''_{f_w} are the real and imaginary parts of the relative

dielectric constant of water, β' and β'' are empirically determined constants, dependent on soil composition in terms of sand and clay.

Since the UG sensors are buried near the air-ground interface (the burial depth is less than 2 m), the reflection from the air-ground interface needs to be considered. If the burial depth of UG sensors is h_u, the total path loss of the UG-UG channel L_{UG-UG} is deduced as [4, 53]:

$$L_{UG-UG} = L_{UG}(d) - 10 \log V(d, h_u) \,, \tag{4}$$

where $V(d, h)$ is the attenuation factor due to the second path:

$$\begin{aligned} V^2(d) =& 1 + \left(\Gamma \cdot \exp\left(-\alpha \Delta r\right)\right)^2 \\ & - 2\Gamma \exp\left(-\alpha \Delta r\right) \cos\left(\pi - \left(\phi - \frac{2\pi f}{c\sqrt{\epsilon'}} \Delta r\right)\right) \,, \end{aligned} \tag{5}$$

where Γ and ϕ are the amplitude and phase angle of the reflection coefficient at the reflection point, c is the velocity of light in vacuum, and $\Delta r = \sqrt{d^2/4 + h_u^2} - d$, is the difference of the two paths.

Assuming that the transmit power of the UG sensor is P_t^u, the antenna gains of the receiver and transmitter are g_r and g_t. Then the received power, P_r^{U-U}, at a receiver sensor node d meters away is $P_r^{U-U} = P_t^u + g_r + g_t - L_{UG-UG}$. Consequently, the transmission range of the UG-UG channel is:

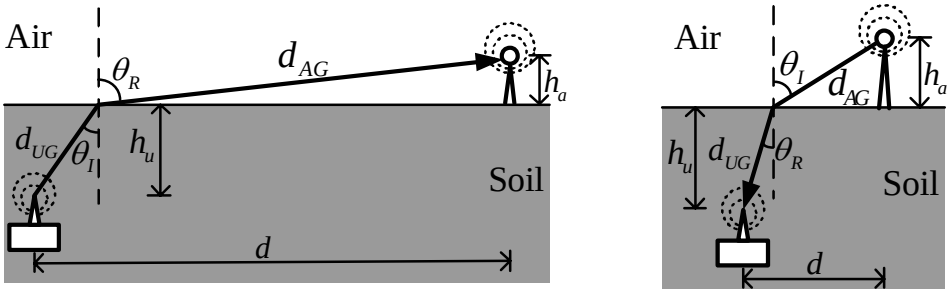
$$R_{UG-UG} = \max\{d : P_r^{U-U}/P_n > SNR_{th}\} \,, \tag{6}$$

where P_n is the noise power; and SNR_{th} is the minimum signal-to-noise ratio required by the receiver.

2.2.2 UG-AG Channel

The path loss of the UG-AG channel L_{UG-AG} consists of three parts: the UG path loss L_{UG}, the AG path loss L_{AG} and the refraction loss from soil to air L_{UG-AG}^R:

$$L_{UG-AG} = L_{UG}(d_{UG}) + L_{AG}(d_{AG}) + L_{UG-AG}^R \,, \tag{7}$$



(a) (b)

Figure 1: Illustration of (a) UG-AG channel and (b) AG-UG channel

where d_{UG} is the length of the UG path, and the d_{AG} is the length of the AG path, as shown in Fig. 1(a). The UG path loss L_{UG} can be derived from (1). The AG path loss L_{AG} is:

$$L_{AG}(d) = -147.6 + 20\log d + 20\log f \ , \tag{8}$$

Since the dielectric constant of soil is much larger than the air, the signals with an incident angle θ_I that is larger than the critical angle θ_c will be completely reflected. Moreover, because the length of the AG path d_{AG} is much larger than the height of the AG sink antenna h_a, the incident angle θ_I is approximately equal to θ_c; and the refracted angle θ_R is approximately equal to 90°, as shown in Fig. 1(a). Then the horizontal distance d between the UG sensor and AG sink is approximately equal to d_{AG}. And

$$d_{UG} \simeq \frac{h_u}{\cos\theta_c}; \quad \theta_c \simeq \arcsin\frac{1}{\sqrt{\epsilon'}}. \tag{9}$$

The refraction loss L^R_{UG-AG} can be calculated as:

$$L^R_{UG-AG} \simeq 10\log\frac{(\sqrt{\epsilon'}+1)^2}{4\sqrt{\epsilon'}}. \tag{10}$$

Then the received power is $P_r^{U-A} = P_t^u + g_r + g_t - L_{UG-AG}$ at the AG sink. Consequently the transmission range of the UG-AG channel is calculated as:

$$R_{UG-AG} \simeq \max\{d_{AG} : P_r^{U-A}/P_n > SNR_{th}\} \ . \tag{11}$$

2.2.3 AG-UG Channel

Similar to the UG-AG channel, the path loss of the AG-UG channel is:

$$L_{AG-UG} = L_{UG}(d_{UG}) + L_{AG}(d_{AG}) + L^R_{AG-UG} \ , \tag{12}$$

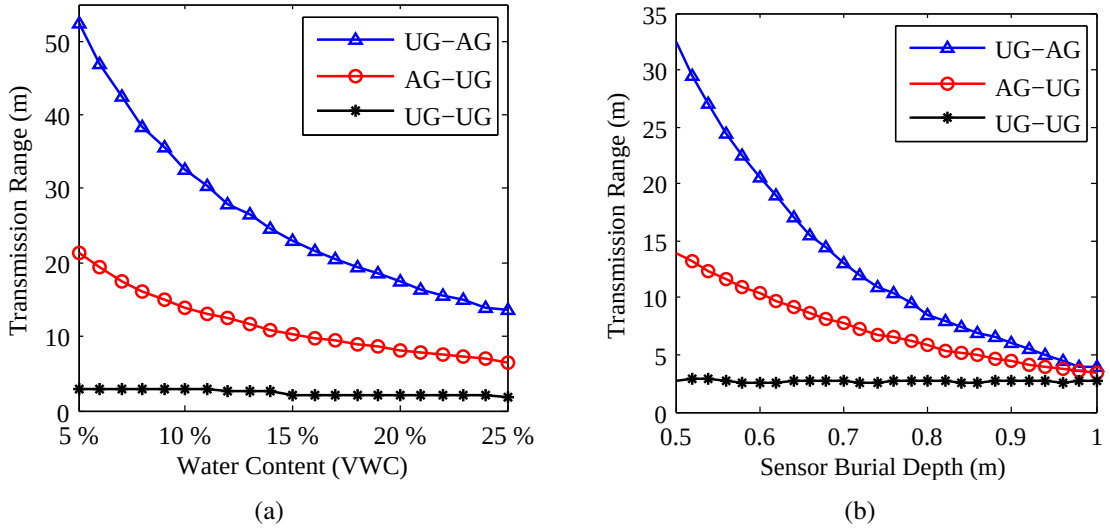


Figure 2: Transmission ranges of the three types of channels in WUSNs as functions of (a) volumetric water content and (b) sensor burial depth.

where L^R_{AG-UG} is the refraction loss from air to soil. As shown in Fig. 1(b), because the dielectric constant of soil is much larger than the air, most radiation energy from the AG sink will be reflected back if the incident angle θ_I is large. Therefore, we only consider the signal with small incident angle. Consequently, the refracted angle θ_R in the soil is even smaller hence it can be viewed approximately as zero. Then the UG path length $d_{UG} \simeq h_u$, and the horizontal distance d between the UG sensor and AG sink is:

$$d \simeq \sqrt{d_{AG}^2 - h_a^2}\,, \quad \cos\theta_I = \frac{h_a}{d_{AG}}\,. \tag{13}$$

The refraction loss L^R_{AG-UG} can be calculated as:

$$L^R_{AG-UG} \simeq 10\log\frac{(\cos\theta_I + \sqrt{\epsilon' - \sin^2\theta_I})^2}{4\cos\theta_I\sqrt{\epsilon' - \sin^2\theta_I}}\,. \tag{14}$$

If the transmit power of the AG sink is P_t^a, then the received power is $P_r^{A-U} = P_t^a + g_r + g_t - L_{AG-UG}$ at the UG sensor. Therefore the transmission range of the UG-AG channel is calculated as:

$$R_{AG-UG} \simeq \max\{d : P_r^{A-U}/P_n > SNR_{th}\}\,. \tag{15}$$

2.2.4 Numerical Results

The numerical results of the transmission ranges of the channels in WUSNs are given in Fig. 2. It shows that R_{UG-UG} is the smallest ($\leq$ 5 m) among the three channels. R_{UG-AG} and R_{AG-UG} are in the range of 10 m to 50 m, depending on the soil water content and the sensor burial depth. R_{UG-AG} is larger than R_{AG-UG} due to the reflection and refraction on the air-ground interface. Moreover, Fig. 2(a) shows that the soil water content has significant influences on all the three types of channels in WUSNs. Fig. 2(b) shows that the sensor burial depth only affects the UG-AG channel and AG-UG channel while does not dramatically influence the UG-UG channel. It should be noted that the range changes of the UG-UG channel is not showed clearly in Fig. 2 due to the much smaller value of the UG-UG channel range compared with the ranges of the other two channels. When soil water content increases from 5% to 25%, the UG-UG range R_{UG-UG} decreases dramatically from 3.42 m to 2.36 m. However, when sensor burial increases from 0.5 m to 1 m, the UG-UG range R_{UG-UG} does not change a lot but fluctuates between 2.7 m to 2.9 m. Beside the soil water content and the sensor burial depth, the antenna height of the AG sinks also has obvious effect on the AG-UG channel, the numerical result of which is not given due to the page limit.

2.3 Dynamic Connectivity in WUSNs

2.3.1 Problem Formulation

After the channel models of the three types of channels in WUSNs are provided, we formulate the problem of the connectivity analysis in WUSNs in this section. We consider a WUSN deployed in a bounded region $\mathbb{R}^2$, as shown in Fig. 3. The UG sensors $\{N_i, i = 1, 2, \cdots\}$ are distributed inside the region $\mathbb{R}^2$ according to a homogeneous Poisson point process of constant spatial intensity λ_u. The AG fixed sinks $\{S_j, j = 1, 2, \cdots\}$ are distributed inside $\mathbb{R}^2$ according to another homogeneous Poisson point process with spatial

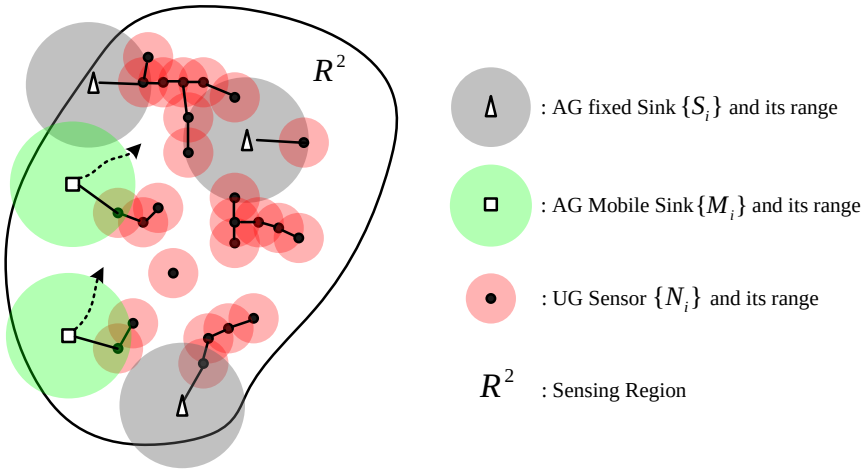


Figure 3: The network model of the WUSNs. The gray disk is the range in which other nodes can connect to the node in the center of the disk.

intensity λ_a. In addition, there are m AG mobile sinks $\{M_k, k = 1, 2, \cdots m\}$ carried by people or machineries inside the region $\mathbb{R}^2$. The monitored region $\mathbb{R}^2$ is much larger than the transmission range of the UG-UG channel. Hence the scale of the network is large and the border effects can be ignored.

In this section, we analyze the probability of the full connectivity of such WUSNs. A WUSN is defined to be fully connected if every UG sensor is connected to at least one AG data sink in a multi-hop fashion within the tolerable latency. Specifically, we introduce the following definition of the full connectivity in WUSNs:

Definition 1: In a WUSN, a UG sensor is connected if either of the following statements is true.

- The UG sensor is connected to at least one fixed AG sink directly or in a multi-hop fashion;
- The UG sensor is connected to at least one mobile AG sink directly or in a multi-hop fashion within the duration t_{max}, where t_{max} is the maximum tolerable latency.

Definition 2: A WUSN is fully connected if all its UG sensors are connected.

The functionalities of the WUSNs include two phases: the sensing phase and the control phase. In the sensing phase, the UG sensors report sensing data to the AG sinks, while in the control phase, the AG sinks send control messages to the UG sensors. Since the UG-AG channel and the AG-UG channel are asymmetrical, we analyze the connectivity in the two phases separately. In the sensing phase, the UG-UG and the UG-AG channels are used, while in the control phase UG-UG and the AG-UG channel are utilized. The maximum tolerable latencies in the sensing phase and control phase are t_s and t_c, respectively. $t_s \ll t_c$ in most envisioned applications. Since the only differences between the connectivity analysis in the two phases are the transmission ranges and the tolerable latencies, we calculate the connectivity probability in the sensing phase in the following sections. The connectivity probability in the control phase can be derived from the developed formulas by changing the values of the transmission range and the tolerable latency.

The connectivity in WUSNs is highly dynamic due to the dynamic underground channel characteristics and the random movement of the mobile sinks. Hence, we mathematically formulate these two randomness in the rest part of this section.

2.3.1.1 Randomness Caused by Dynamic Channel Characteristics

As discussed in Section 2.2, the transmission ranges of the three types of channels in WUSNs are functions of soil water content, the sensor burial depth, and sink antenna height. In most applications of WUSNs, those three environmental and system parameters are either temporally or spatially random, which cause the randomness of connectivity in WUSNs.

Soil Water Content:

According to [43] among many other previous works, the daily soil water content data can be well-fitted by a gamma distribution. The gamma distribution can be completely characterized by its mean and variance, which are given by [43]:

$$\mu_{m_v} = \frac{b2\pi\zeta}{a{r_R}^2\eta\beta}, \qquad \sigma^2_{m_v} = \frac{4\pi\zeta}{\eta\beta^2{r_R}^2}\frac{b^2}{a(\eta+a)}, \tag{16}$$

where ζ is the intensity of the Poisson rain process; a is the normalized soil water loss; b is the rain/irrigation coefficient; $1/r_R$, $1/\eta$, and $1/\beta$ are the mean cell radius, duration, and intensity of each rain, respectively. Then, the probability density function (PDF) of the soil water content can be derived:

$$f(m_v) = m_v^{-1+\mu_{m_v}^2/\sigma_{m_v}^2} \cdot \frac{e^{-m_v \cdot \mu_{m_v}/\sigma_{m_v}^2}}{\left(\frac{\sigma_{m_v}^2}{\mu_{m_v}}\right)^{\mu_{m_v}^2/\sigma_{m_v}^2} \cdot \Gamma(\frac{\mu_{m_v}^2}{\sigma_{m_v}^2})}, \tag{17}$$

where $\Gamma(x)$ is the Gamma function [61]; μ_{m_v} and $\sigma_{m_v}^2$ are given in (16). It should be noted that the randomness brought by the dynamic soil water content is only in temporal scale. In a give time stamp, the soil water content throughout the monitored field can be considered to be the same.

Sensor Burial Depth and Sink Antenna Height:

In WUSN applications, the burial depths of all UG sensors are not necessarily the same. The data at different soil levels may be required and the depth deviations may be incurred during the deployment processes. Hence, the sensor burial depth throughout the whole WUSN is a random variable. Similarly, the antenna height of each mobile AG sink is different since different people or machineries may carry the sink handsets at different positions. Moreover, the antenna heights of different fixed AG sinks are different due to the deviations in the deployment process. Therefore, the antenna heights of all the AG sinks throughout the monitored field are also random variables.

In this section, we model the random sensor burial depths and sink antenna heights as uniformly distributed variables. Specifically, the UG sensor burial depths are uniformly distributed in $[h_u^{min}, h_u^{max}]$; and the antenna heights of the AG fixed and mobile sinks are uniformly distributed in $[h_a^{min}, h_a^{max}]$. It should be noted that the randomness brought by the different sensor burial depth and the sin antenna height is only in the spatial scale. After the deployment, the burial depth and the antenna height are assumed to be remain the same during the WUSN operation.

2.3.1.2 *Randomness Caused by AG Sink Mobility:*

The employment of the AG mobile sinks can improve the connectivity in WUSNs if a certain level of delay is allowed. Meanwhile, the random movement of the AG mobile sinks also brings randomness. Since the mobile AG sinks are carried by people or machineries, the movement of the AG mobile sinks can be modeled by the widely used Random Waypoint (RWP) Model [20]. In RWP model, the random movement of a mobile sink is modeled as a sequence of steps. A step includes a flight and a following pause. In a flight, the sink first select a destination that is uniformly distributed in the whole region $\mathbb{R}^2$. Then the sink starts to move towards the destination with a constant speed v m/s. After it arrives the destination, the sink pauses for τ second and then starts the next step. The speed v and the pause τ are chosen uniformly from $[v_{min}, v_{max}]$ and $[0, \tau_{max}]$, respectively.

2.3.2 Lower Bound of Connectivity Probability in WUSNs

According to the channel models derived in Section 2.2 and the network, environment, and mobility models derived in Section 2.3.1, the connectivity in WUSNs depends on various environmental and system parameters. Here and in the next section, the lower and upper bounds for the connectivity probability in WUSNs are derived analytically. These theoretical bounds enable the quantitative analysis of the effects of multiple system and environmental parameters on the connectivity in WUSNs.

From Definition 2, the full connectivity probability P_c of WUSNs can be expressed as:

$$\begin{aligned} P_c &= P(\text{Every UG sensor is connected}) \\ &= \sum_{n=0}^{\infty} P(\text{All } n \text{ UG sensors are connected}) \times P(\text{There are } n \text{ UG sensors in } \mathbb{R}^2) \ . \end{aligned} \tag{18}$$

According to the FKG inequality [62],

$$P(\text{All } n \text{ UG sensors are connected}) \begin{cases} \geq \prod_{i=1}^{n} P(N_i \text{ is connected}) \ , & \text{if } n \geq 1 \\ = 0 \ , & \text{if } n = 0 \ . \end{cases} \tag{19}$$

Since each UG sensor node is assumed to be identically distributed, then

$$\prod_{i=1}^{n} P(N_i \text{ is connected}) = P^n(N_i \text{ is connected})$$

Additionally, since the UG sensors are distributed according to a Poisson point process,

$$P(\text{There are } n \text{ UG sensors in } \mathbb{R}^2) = \frac{(\lambda_u S_{\mathbb{R}^2})^n}{n!} e^{-\lambda_u S_{\mathbb{R}^2}} , \tag{20}$$

where $S_{\mathbb{R}^2}$ is the area of the region $\mathbb{R}^2$. Then

$$\begin{aligned} P_c &\geq \sum_{n=0}^{\infty} P^n(N_i \text{ is connected}) \cdot \frac{(\lambda_u S_{\mathbb{R}^2})^n}{n!} e^{-\lambda_u S_{\mathbb{R}^2}} \\ &= \exp\Big\{ -\lambda_u S_{\mathbb{R}^2} \cdot [1 - P(N_i \text{ is connected})] \Big\} \\ &= \exp\Big\{ -\lambda_u S_{\mathbb{R}^2} \cdot P(N_i \text{ is not connected}) \Big\} . \end{aligned} \tag{21}$$

Next, we evaluate the upper bound of $P(N_i \text{ is not connected})$ in (21), the probability that a single UG sensor node N_i is not connected. According to Definition 1, we have

$$P(N_i \text{ is not connected}) = P(N_i \nleftrightarrow \text{fixed sink} \cap N_i \nleftrightarrow \text{mobile sink within } t_s) \tag{22}$$

where $A \nleftrightarrow B$ indicates that A is not connected to B; t_s is the maximum tolerable latency in the sensing phase given in Section IV. Since the event $\{N_i \nleftrightarrow \text{fixed sink}\}$ and event $\{N_i \nleftrightarrow \text{mobile sink within } t_s\}$ can be viewed as independent, then

$$P(N_i \text{ is not connected}) = P(N_i \nleftrightarrow \text{fixed sink}) \cdot P(N_i \nleftrightarrow \text{mobile sink within } t_s) . \tag{23}$$

According to (18) to (23), to derive the lower bound of the connectivity probability P_c in WUSNs, the upper bounds of two probabilities need to be found out. The two probabilities are: the probability that the UG sensor N_i is not connected to all fixed AG sinks, $P(N_i \nleftarrow \nrightarrow \text{fixed sink})$, and the probability that the UG sensor N_i is not connected to all mobile AG sink within time t_s, $P(N_i \nleftrightarrow \text{mobile sink within } t_s)$. In the rest part of this section, the upper bounds of the two probabilities are developed.

2.3.2.1 Upper Bound of $P(N_i \nleftrightarrow$ fixed sink$)$

The probability that the UG sensor N_i is not connected to any fixed AG sinks $P(N_i \nleftrightarrow$ fixed sink) can be further developed as

$$P(N_i \nleftrightarrow \text{fixed sinks}) = \sum_{n=0}^{\infty} P(N_i \nleftrightarrow S_1 \cap N_i \nleftrightarrow S_2 \cap ... \cap N_i \nleftrightarrow S_n) \times P(\text{There are } n \text{ fixed AG sinks in } \mathbb{R}^2) , \quad (24)$$

where S_j is j^{th} fixed AG sink. Since AG fixed sinks are distributed according to a Poisson point process with density λ_a, the probability P(There are n fixed AG sinks in $\mathbb{R}^2$) can be calculated using (20) by just replacing λ_u with λ_a. $\{N_i \nleftrightarrow S_{j_1}\}$ and $\{N_i \nleftrightarrow S_{j_2}\}$ $(j_1 \neq j_2)$ can be viewed as independent events. Then

$$\begin{aligned} P(N_i \nleftrightarrow \text{fixed sinks}) &= \sum_{n=0}^{\infty} P^n(N_i \nleftrightarrow S_j) \cdot \frac{(\lambda_a S_{\mathbb{R}^2})^n}{n!} e^{-\lambda_a S_{\mathbb{R}^2}} \\ &= \exp\Big\{ -\lambda_a S_{\mathbb{R}^2} \cdot [1 - P(N_i \nleftrightarrow S_j)] \Big\} \\ &= \exp\Big\{ -\lambda_a S_{\mathbb{R}^2} \cdot P(N_i \leftrightarrow S_j) \Big\} , \end{aligned} \quad (25)$$

where $A \leftrightarrow B$ indicates that A is connected to B.

Since the position of the UG sensor N_i and the position of the fixed sink S_j are distributed according to two different homogeneous Poisson point processes, then

$$P(N_i \nleftrightarrow \text{fixed sinks}) = \exp\Big\{ -\lambda_a S_{\mathbb{R}^2} \cdot \int_{\mathbb{R}^2}\int_{\mathbb{R}^2} \Big(\frac{1}{S_{\mathbb{R}^2}}\Big)^2 \cdot P(\boldsymbol{x}_i \leftrightarrow \boldsymbol{z}_j)\, d\boldsymbol{x}_i d\boldsymbol{z}_i \Big\} , \quad (26)$$

where $\boldsymbol{x}_i$ is the vector position of the UG sensor node N_i; $\boldsymbol{z}_j$ is the vector position of the fixed sink S_j; $P(\boldsymbol{x}_i \leftrightarrow \boldsymbol{z}_j)$ is the probability that the UG sensor at $\boldsymbol{x}_i$ is connected to the AG fixed sink at $\boldsymbol{z}_j$.

Next, we investigate the lower bound of the probability $P(\boldsymbol{x}_i \leftrightarrow \boldsymbol{z}_j)$. To derive the lower bound, we first map the WUSN on a discrete lattice, as shown in Fig. 4. The square lattice L over the region $\mathbb{R}^2$ is constructed as follows. The location of the UG sensor $\boldsymbol{x}_i$ is on one vertex of the lattice, which is set as the origin of the lattice. The straight line ef connecting

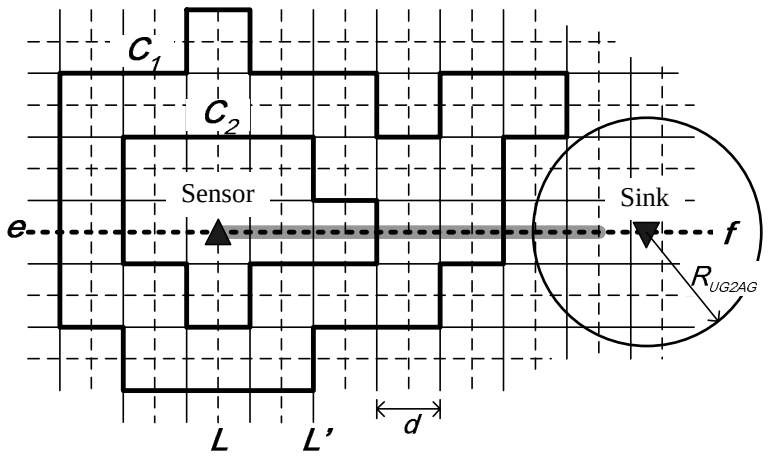


Figure 4: Mapping the WUSN on a lattice **L** (dashed) and its dual **L'** (plain).

$\boldsymbol{x}_i$ and $\boldsymbol{z}_j$ forms a sequence of horizontal edges of the lattice L. The length of each edge is d. Let L' be the dual lattice of L. The vertexes of L' are placed in the center of every square of L. The edges of L' crosse every edge of L. According to the above structure, there exists a one-to-one relation between the edges of L and the edges of L'. L and L' have the same edge length $d = \frac{1}{\sqrt{5}} R_{UG-UG}(m_v)$. The value is chosen so that two UG sensors deployed in two adjacent squares of the dual lattice L' are guaranteed to be able to connect to each other.

Note that the soil water content m_v is a random variable as discussed in Section IV, and $R_{UG-UG}(m_v)$ is a function of the soil water content m_v. Hence the edge length of the lattice d is also a random variable. According to [43], at one time stamp, the soil water contents can be viewed as the same throughout the whole monitored field since the water contents are highly spatio-correlated. Therefore, all the edges have the same length d in the lattice L and L' at one time stamp. The edge length $d(m_v)$ is random in different time stamps.

Some definitions are first given before the next step.

Definition 3: An edge l' of the L' is said to be open if both squares adjacent to l contains at least one UG sensor.

Definition 4: An edge l of the L is said to be open if and only if the corresponding edge of L' is open.

Definition 5: A path of the L or L's is said to be open (closed) if all edges forming the

path are open (closed).

If a open path of the L is given, all the UG sensors in the squares in L' along the open path are connected to each other.

Now consider the connection between the sensor at $\boldsymbol{x}_i$ and the sink at $\boldsymbol{z}_j$. The region $\mathbb{R}^2$ is divided into two parts, the region inside the circle C_{z_j} and the region outside the circle. C_{z_j} is defined as the circle with radius $R_{UG-AG}(m_v, h_u)$ and center located at $\boldsymbol{z}_j$, as shown in Fig. 4, where the UG-AG channel range $R_{UG-AG}(m_v, h_u)$ is a function determined by two random variables: the soil water content m_v and the burial depth of the last hop UG sensor that is directly connected to the AG fixed sink h_u. All UG sensors located inside C_{z_j} are connected to the sink directly. Note that the UG-AG channel range $R_{UG-AG}(m_v, h_u)$ is used since we aim to calculate the connectivity probability in sensing phase of the WUSNs. For the control phase of the WUSNs, the UG-AG channel range $R_{AG-UG}(m_v, h_u, h_a)$ is used and the connectivity probability of the control phase can be derived in the similar way.

If there is an open path of L connecting $\boldsymbol{x}_i$ and a vertex V of L inside C_{z_j}, then UG sensor and the AG sink are guaranteed to be connected by each other. Note that the square in L' containing vertex V should be completely inside the circle C_{z_j}. The set of these open paths is denoted as $\boldsymbol{P}_o = \{P_o^1, P_o^2, ...\}$, where $P_o^1, P_o^2, ...$ denote all possible open paths. Then,

$$P(\boldsymbol{x}_i \leftrightarrow \boldsymbol{z}_j | m_v, h_u) = P\left(\cup_i^{\infty} P_o^i\right) \geq \begin{cases} \max_i\{P(P_o^i)\}, \\ P(|\boldsymbol{P}_o| > 0) = 1 - P(|\boldsymbol{P}_o| = 0), \end{cases} \tag{27}$$

where $P(\boldsymbol{x}_i \leftrightarrow \boldsymbol{z}_j | m_v, h_u)$ is the conditional probability assuming that m_v and h_u are given; $|\boldsymbol{P}_o|$ is the number of the existing open paths. In (77), two bounds of $P(\boldsymbol{x}_i \leftrightarrow \boldsymbol{z}_j | m_v, h_u)$ are given, which are the maximum probability that a certain open path exists, i.e. $\max_i\{P(P_o^i)\}$, and the probability that there is at least one open path, i.e. $P(|\boldsymbol{P}_o| > 0)$. The larger one of the two bounds is utilized as the lower bound of $P(\boldsymbol{x}_i \leftrightarrow \boldsymbol{z}_j | m_v, h_u)$, which is determined by the UG sensor density λ_u. Hence,

$$P(\boldsymbol{x}_i \leftrightarrow \boldsymbol{z}_j | m_v, h_u) \geq \max\left\{ \max_i\{P(P_o^i)\},\ 1 - P(|\boldsymbol{P}_o| = 0) \right\}, \tag{28}$$

We first calculate $\max_i\{P(P_o^i)\}$ in (28). Since the UG sensors are distributed according to a homogeneous Poisson point process, the shortest open path connecting $\boldsymbol{x}_i$ and $\boldsymbol{z}_j$ can yield the maximum existing probability. Specifically, the shortest path is the line segment on ef between $\boldsymbol{x}_i$ and the first vertex of L inside C_{z_j}. This line segment is illustrated by the thick gray segment in Fig. 4. The length of the line segment is Wd, where

$$W=\begin{cases}\left\lceil \frac{\|\boldsymbol{x}_i-\boldsymbol{z}_j\|-R_{UG-AG}(m_v,h_u)}{d}\right\rceil+1, & \text{if } \|\boldsymbol{x}_i-\boldsymbol{z}_j\|\geq R_{UG-AG}(m_v,h_u)\\ 0, & \text{if } \|\boldsymbol{x}_i-\boldsymbol{z}_j\|<R_{UG-AG}(m_v,h_u)\end{cases} \tag{29}$$

where $\lceil a\rceil$ means rounding a to the nearest integer $\geq a$. Hence, the maximum probability that a certain open path exists is

$$\begin{aligned}\max_i\{P(P_o^i)\} &= P(\text{There exists an open path with length } W)\\ &= P^{W+1}(\text{There exists at least one sensor in a square } d^2)\\ &= (1-q)^{W+1}\ ,\end{aligned} \tag{30}$$

where

$$q = P(\text{There is no sensor in a square } d^2) = e^{-\lambda_u d^2}. \tag{31}$$

We then calculate $P(|\boldsymbol{P}_o|=0)$ in (28). $\left|\{P_o\}\right| = 0$, if and only if the sensor $\boldsymbol{x}_i$ lies in the interior of some closed circuits of the dual lattice L', which do not contain a whole common square that is also inside the circle C_{z_j}, such as C_1 and C_2 (thick black circuits) in Fig. 4. Hence, $P(|\boldsymbol{P}_o|=0)$ can be evaluated by counting the number of such closed circuits in L'. Let $\rho(n)$ be the number of circuits in L' which have length nd and contain $\boldsymbol{x}_i$ in their interiors. To contain $\boldsymbol{x}_i$ in their interiors, those circuits pass through some point on the line ef, as shown in Fig. 4. The position of the corresponding pass vertex in L' has the form of $(kd-\frac{1}{2}d, \frac{1}{2}d)$. k cannot be larger than $\lfloor\frac{n}{2}\rfloor-1$. Otherwise the circuits would have a length larger than n. Thus, such a circuit contains a self-avoiding walk of length $n-1$ starting from a vertex at $(kd-\frac{1}{2}d, \frac{1}{2}d)$ and $k>\lfloor\frac{n}{2}\rfloor-1$. Moreover, to contain $\boldsymbol{x}_i$ inside, the length of

the circuits $n \geq 4$. The number of self avoiding walks of L' having length n and beginning at a vertex is denoted as $\sigma(n)$. It has been proven in [62] that $\sigma(n) \leq 4 \cdot 3^{n-1}$ in a 2-D plane.

Since those closed circuits do not contain a whole common square with the circle C_{z_j}, such as C_1 or C_2 in Fig. 4, they must pass through at least one point on the shortest path connecting $\boldsymbol{x}_i$ and $\boldsymbol{z}_j$ (illustrated by the thick gray segment in Fig. 4). Hence, those closed circuits contain a self-avoiding walk of length $n-1$ ($n \geq 4$) starting from a vertex at $(kd - \frac{1}{2}d, \frac{1}{2}d)$ and $k \leq \min\{\lfloor \frac{n}{2} \rfloor - 1, W\}$. The total number of such closed circuits is denoted as CN. Based on the above discussions, the upper bound of CN can be calculated as follows.

$$CN \leq \sum_{n=4}^{\infty} \sigma(n-1) + \sum_{n=6}^{\infty} \sigma(n-1) + \cdots \sum_{n=2W}^{\infty} \sigma(n-1) \ . \tag{32}$$

Then the upper bound of $P(|\boldsymbol{P}_o| = 0)$, the probability that there is no open path connecting the UG sensor at $\boldsymbol{x}_i$ and the AG fixed sink at $\boldsymbol{z}_j$, is:

$$\begin{aligned} P(|\boldsymbol{P}_o| = 0) &\leq \sum_{n=4}^{\infty} \sigma(n-1) \cdot q^n + \sum_{n=6}^{\infty} \sigma(n-1) \cdot q^n + \ldots + \sum_{n=2W}^{\infty} \sigma(n-1) \cdot q^n \\ &= \begin{cases} 36 \cdot q^4 \cdot \frac{1-(3q)^{2W-2}}{(1+3q)(1-3q)^2} \ , & \text{if } q < \frac{1}{3} \\ 1 \ , & \text{if } q \geq \frac{1}{3} \end{cases} \end{aligned} \tag{33}$$

Substituting (80) and (33) into (28), we derive

$$P(\boldsymbol{x}_i \leftrightarrow \boldsymbol{z}_j \big| m_v, h_u) \geq \begin{cases} \max\left\{(1-q)^{W+1}, \ 1 - \frac{36q^4[1-(3q)^{2W-2}]}{(1+3q)(1-3q)^2}\right\}, & \text{if } q < \frac{1}{3} \\ (1-q)^{W+1}, & \text{if } q \geq \frac{1}{3} \end{cases} \tag{34}$$

$$\stackrel{\text{def}}{=} \gamma_1(\boldsymbol{x}_i, \boldsymbol{z}_j, \lambda_u, m_v, h_u) \tag{35}$$

Based on the discussion in Section IV.A, the probability without conditions $P(\boldsymbol{x}_i \leftrightarrow \boldsymbol{z}_j)$ can be calculated by

$$P(\boldsymbol{x}_i \leftrightarrow \boldsymbol{z}_j) = \frac{1}{h_u^{max} - h_u^{min}} \iint P(\boldsymbol{x}_i \leftrightarrow \boldsymbol{z}_j \big| m_v, h_u) \cdot f(m_v) \, dm_v dh_u, \tag{36}$$

where $f(m_v)$ is the PDF of the soil water content given in (17); h_u^{max} and h_u^{min} are the maximum and the minimum UG sensor burial depths defined in Section IV.A. Substituting (34)

and (36) into (26) yields the upper bound of $P(N_i \nleftrightarrow$ fixed sink$)$:

$$P(N_i \nleftrightarrow \text{fixed sink}) \leq \exp\Big\{ - \frac{\lambda_a}{h_u^{max} - h_u^{min}} \cdot \frac{1}{S_{\mathbb{R}^2}} \cdot \int_{\mathbb{R}^2}\int_{\mathbb{R}^2}\int_{m_v}\int_{h_u} f(m_v) \cdot \gamma_1(\boldsymbol{x}_i, \boldsymbol{z}_j, \lambda_u, m_v, h_u)\, d\boldsymbol{x}_i\, d\boldsymbol{z}_i\, dm_v\, dh_u \Big\}. \tag{37}$$

2.3.2.2 *Upper Bound of $P(N_i \nleftrightarrow$ mobile sink within $t_s)$*

Beside fixed sinks, mobile sinks also contribute to the network connectivity in WUSNs. In this subsection, we calculate the upper bound of the probability that a UG sensor N_i is not connected to any mobile sink within time t_s, i.e. $P(N_i \nleftrightarrow$ mobile sink within $t_s)$ given in (23). Due to the mobility of the mobile sinks, the contributions of the multi-hop connection is much smaller than those of the direct connection. Therefore, only the direct connection is considered while deriving the upper bound of the probability, i.e.

$$P(N_i \nleftrightarrow \text{mobile sink within } t_s) \leq P(N_i \overset{\text{direct}}{\nleftrightarrow} \text{mobile sink within } t_s)\,. \tag{38}$$

As discussed in Section IV, m mobile sinks randomly move in region $\mathbb{R}^2$ according to the RWP model. The stationary node distribution of RWP model is provided in [42], while the intermeeting time between the mobile nodes in RWP model is proved to be exponentially distributed in [19]. We utilize their results to derive the upper bound of the probability $P(N_i \overset{\text{direct}}{\nleftrightarrow}$ mobile sink within $t_s)$.

The UG sensor N_i is regarded as directly connected by the mobile sinks if at least one of the m mobile sinks visits the UG-AG communication range around N_i at least once during the time slot $[0, t_s]$. Let $H_k(t)$ be the event that the k^{th} mobile sink does not directly cover the sensor N_i at time stamp t, then

$$P(N_i \overset{\text{direct}}{\nleftrightarrow} \text{mobile sink within } t_s) = P\Big(\cap_{t\in[0,t_s]} \cap_{k=1,\ldots,m} H_k(t)\Big), \tag{39}$$

Note that the event $H_k(t)$ is determined by the position of the sensor N_i and the k^{th} mobile sink, the soil water content, and the sensor burial depth. Let $\boldsymbol{y}_k(t)$ denotes the position of the k^{th} sink at time stamp t. If sensor node N_i's position $\boldsymbol{x}_i$, the soil water

content m_v, and the UG sensor burial depth h_u are given, the event $H_k(t)$ can be further expressed as

$$\{H_k(t)\big|\boldsymbol{x}_i, m_v, h_u\} = \{\|\boldsymbol{y}_k(t) - \boldsymbol{x}_i\| > R_{UG-AG}(m_v, h_u)\}, \tag{40}$$

where $R_{UG-AG}(m_v, h_u)$ is the communication range of the UG-AG channel as a function of m_v and h_u. Then the probability that event $\{\cap_{k=1,...,m} H_k(t)\}$ in (39) happens is

$$P\Big(\cap_{k=1,...,m} H_k(t)\Big) = \frac{1}{h_u^{max} - h_u^{min}} \frac{1}{S_{\mathbb{R}^2}} \iiint P\Big(\cap_{k=1...m} H_k(t)\big|\boldsymbol{x}_i, m_v, h_u\Big) \cdot f(m_v)\, d\boldsymbol{x}_i dm_v dh_u, \tag{41}$$

where the conditional probability $P\Big(\cap_{k=1,...,m} H_k(t)\big|\boldsymbol{x}_i, m_v, h_u\Big)$ can be calculated by

$$\begin{aligned} P\Big(\cap_{k=1,...,m} H_k(t)\big|\boldsymbol{x}_i, m_v, h_u\Big) &= \left(\int_{\boldsymbol{x} \in \mathbb{R}^2 - \mathbb{C}^2[\boldsymbol{x}_i, R_{UG-AG}(m_v, h_u)]} \xi(\boldsymbol{x})\, d\boldsymbol{x}\right)^m \\ &\stackrel{\text{def}}{=} \gamma_2^m(\boldsymbol{x}_i, m_v, h_u)\ , \end{aligned} \tag{42}$$

where $\mathbb{C}^2[\boldsymbol{x}_i, R_{UG-AG}(m_v, h_u)]$ is the disk region centered at $\boldsymbol{x}_i$ with radius $R_{UG-AG}(m_v, h_u)$; $\xi(\boldsymbol{x})$ is the PDF that a sink visit the position $\boldsymbol{x}$ at arbitrary time stamp (stationary node distribution), which is defined by the RWP model; the detailed expression of $\xi(\boldsymbol{x})$ is given in [42].

Given the convex region $\mathbb{R}^2$, the maximum flight length is denoted by D, which is the maximum length of a line segment in $\mathbb{R}^2$. Then the maximum time duration t_D for a sink to finish two sequential flights is:

$$t_D = 2(\tau_{max} + D/v_{min})\ . \tag{43}$$

where τ_{max} and v_{min} are the maximum pause time and the minimum velocity of each flight, respectively, which are defined in Section IV.B.

The current positions of all the sinks are independent with their positions t_D ago since all the sinks have already finished at least two flights. We choose an index set of time stamps in $[0,\ t_s]$:

$$\boldsymbol{T}_D = \left\{0,\ t_D,\ 2t_D, ..., \lfloor \frac{t_s}{t_D} \rfloor \cdot t_D \right\}\ .$$

Then the events $\{\cap_{k=1,\ldots,m} H_k(t_j), t_j \in \boldsymbol{T}_D\}$ are all independent. Hence,

$$P\Big(\cap_{t\in[0,t_s]}\cap_{k=1,\ldots,m}H_k(t)\Big) \le P\Big(\cap_{t_j\in\boldsymbol{T}_D}\cap_{k=1,\ldots,m}H_k(t_j)\Big) = P^{\lfloor t_s/t_D\rfloor}\Big(\cap_{k=1,\ldots,m} H_k(t)\Big), \quad (44)$$

By substituting (39), (41), (42), and (44) into (38), the upper bound of the probability $P(N_i \leftrightarrows \text{mobile sink within } t_s)$ is derived:

$$P(N_i \leftrightarrows \text{mobile sink within } t_s) \le \left(\frac{1}{h_u^{max}-h_u^{min}}\cdot\frac{1}{S_{\mathbb{R}^2}}\cdot\int_{\mathbb{R}^2}\int_{m_v}\int_{h_u}\gamma_2^m(\boldsymbol{x}_i, m_v, h_u)\cdot f(m_v)\, d\boldsymbol{x}_i dm_v dh_u\right)^{\lfloor\frac{t_s}{t_D}\rfloor}. \quad (45)$$

2.3.2.3 *Lower Bound of the Connectivity Probability in WUSNs*

According to the above analysis, the lower bound of the connectivity probability in WUSNs can be derived by substituting (23) into (21):

$$P_c \ge \exp\Big[-\lambda_u \cdot S_{\mathbb{R}^2} \cdot P(N_i \leftrightarrows \text{fixed sink}) \cdot P(N_i \leftrightarrows \text{mobile sink within } t_s)\Big]. \quad (46)$$

where $P(N_i \leftrightarrows \text{mobile sink within } t_s)$ is given by (45); and $P(N_i \leftrightarrows \text{fixed sink})$ is given by (37).

2.3.3 Upper Bound of Connectivity Probability in WUSNs

The absence of isolated UG sensor is a necessary but not sufficient condition for the full connectivity in WUSNs. Hence the probability that there are no isolated UG sensors, denoted by P(no isolated UG sensor), is an upper bound for the connectivity probability in WUSNs. Therefore we have:

$$P_c \le P(\text{no isolated UG sensor}) = \sum_{n=0}^{\infty} P(\text{All } n \text{ UG sensors are not isolated}) \times P(\text{There are } n \text{ UG sensors in } \mathbb{R}^2). \quad (47)$$

The isolation events of each node can be viewed as independent according to [13, 74]. Hence

$$\begin{aligned} &P(\text{no isolated UG sensor}) \\ &= \sum_{n=0}^{\infty} P^n(N_i \text{ is not isolated}) \cdot P(\text{There are } n \text{ UG sensors in } \mathbb{R}^2) \,. \end{aligned} \tag{48}$$

Then using the same strategy in (20) and (21), we derive:

$$P_c \leq \exp\Big\{ -\lambda_u \, S_{\mathbb{R}^2} \cdot P(N_i \text{ is isolated}) \Big\} \,. \tag{49}$$

To derive the upper bound of P_c in (49), we analyze the lower bound of the probability $P(N_i \text{ is isolated})$. Due to the randomness of the soil water content m_v and the sensor burial depth h_u, the $P(N_i \text{ is isolated})$ is calculated by utilizing the conditional probability $P(N_i \text{ is isolated} \big| m_v, h_u)$, i.e.

$$P(N_i \text{ is isolated}) = \frac{1}{h_u^{max} - h_u^{min}} \iint P(N_i \text{ is isolated} \big| m_v, h_u) \cdot f(m_v) \, dm_v dh_u, \,. \tag{50}$$

Moreover, a UG sensor is isolated, if and only if no other UG sensors, AG fixed sinks and AG mobile sinks exist inside its transmission range. Note that the three events are independent. Then

$$\begin{aligned} P(N_i \text{ is isolated} \big| m_v, h_u) &= P(\text{no sensor, fixed sink, mobile sink in } N_i\text{'s range} \big| m_v, h_u) \\ &= P(\text{no other UG sensor in } N_i\text{'s range} \big| m_v) \\ &\quad \times P(\text{no fixed sink in } N_i\text{'s range} \big| m_v, h_u) \\ &\quad \times P(\text{no mobi. sink moves in } N_i\text{'s range within } t_s \big| m_v, h_u) \,. \end{aligned} \tag{51}$$

Since the UG sensors are distributed according to a homogeneous Poisson point process with density λ_u, we have

$$\begin{aligned} &P(\text{no fixed sink in } N_i\text{'s range} \big| m_v, h_u) \\ &= P(N_i \text{ has no sensor neighbor} \big| m_v) = e^{-\lambda_u \pi R^2_{UG-UG}(m_v)} \,. \end{aligned} \tag{52}$$

Similarly, the AG fixed sinks are distributed according to a Poisson point process with density λ_a. Therefore,

$$
\begin{aligned}
&P(\text{no fixed sink in } N_i\text{'s range} \big| m_v, h_u) \\
&= \sum_{n=0}^{\infty} P(\text{There are } n \text{ fixed sinks in } \mathbb{R}^2) \cdot P(\text{All the } n \text{ fixed sinks are not in } N_i\text{'s range} \big| m_v, h_u) \\
&= \sum_{n=0}^{\infty} \frac{(\lambda_a S_{\mathbb{R}^2})^n}{n!} e^{-\lambda_a S_{\mathbb{R}^2}} \cdot \left(\frac{S_{\mathbb{R}^2} - \pi \cdot R_{UG-AG}^2(m_v, h_u)}{S_{\mathbb{R}^2}} \right)^n \\
&= e^{-\lambda_a \pi R_{UG-AG}^2(m_v, h_u)} \, . \qquad (53)
\end{aligned}
$$

The probability that a UG sensor is connected to a mobile sink is affected by the position of the UG sensor. Hence,

$$
\begin{aligned}
&P(\text{no mobi. sink moves in } N_i\text{'s range within } t_s \big| m_v, h_u) \qquad (54) \\
&= \frac{1}{S_{\mathbb{R}^2}} \int_{S_{\mathbb{R}^2}} P(\text{no mobi. sink in } \boldsymbol{x}_i\text{'s range within } t_s \big| m_v, h_u) \, d\boldsymbol{x}_i \, ,
\end{aligned}
$$

where

$$
\begin{aligned}
&P(\text{no mobi. sink in } \boldsymbol{x}_i\text{'s range within } t_s \big| m_v, h_u) \qquad (55) \\
&= P^m(k^{th} \text{ mobi. sink is not in } \boldsymbol{x}_i\text{'s range within } t_s \big| m_v, h_u) \\
&= \Big[1 - P(k^{th} \text{ mobi. sink is in } \boldsymbol{x}_i\text{'s range within } t_s \big| m_v, h_u) \Big]^m .
\end{aligned}
$$

Since the mobile sinks have limited moving velocity, i.e. $v < v_{max}$, the upper bound of the probability $P(k^{th}$ mobi. sink is in $\boldsymbol{x}_i$'s range within $t_s \big| m_v, h_u)$ can be derived by assuming that the mobile sink moves towards $\boldsymbol{x}_i$'s range with its maximum velocity at the time stamp 0. Therefore,

$$
\begin{aligned}
&P(k^{th} \text{ mobi. sink is in } \boldsymbol{x}_i\text{'s range within } t_s \big| m_v, h_u) \qquad (56) \\
&\leq \int_{\boldsymbol{x} \in \mathbb{C}^2[\boldsymbol{x}_i, R_{UG-AG}(m_v, h_u) + v_{max} \cdot t_s]} \xi(\boldsymbol{x}) \, d\boldsymbol{x} \\
&\stackrel{\text{def}}{=} \gamma_3(\boldsymbol{x}_i, m_v, h_u) \, ,
\end{aligned}
$$

where $\mathbb{C}^2[\boldsymbol{x}_i, R_{UG-AG}(m_v, h_u) + v_{max} \cdot t_s]$ is the circular region centered at $\boldsymbol{x}_i$ with radius $R_{UG-AG}(m_v, h_u) + v_{max} \cdot t_s$; $\xi(\boldsymbol{x})$ is the PDF of the stationary node distribution in the RWP model, which is given in [42].

By substituting (50)-(56) into (49), the upper bound of the connectivity probability in WUSNs is obtained.

$$P_c \leq \exp\Bigg\{-\frac{\lambda_u}{h_u^{max} - h_u^{min}} \int_{m_v}\int_{h_u}\int_{\boldsymbol{x}_i} e^{-\pi\left[\lambda_u \cdot R^2_{UG-UG}(m_v) + \lambda_a \cdot R^2_{UG-AG}(m_v, h_u)\right]} \cdot\Big[1 - \gamma_3(\boldsymbol{x}_i, m_v, h_u)\Big]^m \cdot f(m_v)\, dm_v\, dh_u\, d\boldsymbol{x}_i\Bigg\}. \tag{57}$$

2.3.4 Numerical Evaluation

According to the analytical results shown in (46) and (57), the lower and upper bounds of the connectivity probability in WUSNs are functions of multiple system and environmental parameters, including the UG sensor node density λ_u, the AG fixed sink density λ_a, the number of AG mobile sinks m, the mobility model of the mobile sinks, the tolerable latency (t_s in the sensing phase and t_c in the control phase), the transmission ranges (R_{UG-UG}, R_{UG-AG} in sensing phase and R_{UG-UG}, R_{AG-UG} in control phase), the operating frequency, the distribution of the random soil water content, the sensor burial depth, and the sink antenna height. In this section, we numerically analyze the effects of the above system and environmental parameters on the connectivity in WUSNs. The theoretical probability bounds are validated by the simulations in the meantime. Note that the analysis is based on the sensing phase unless otherwise specified.

Except studying the effects of certain parameters, the default values are set as follows: The monitored region is a 500 $m \times 500$ m square. The UG sensors are deployed according to a homogeneous Poisson point process of spatial intensity λ_u with random burial depths. The density of the UG sensor node λ_u is in the range from 0.05 m^{-2} to 1.6 m^{-2}. The mean number of the UG sensor node is calculated by multiplying the region area by the UG sensor node density λ_u. The burial depths of all the UG sensors are uniformly distributed in the

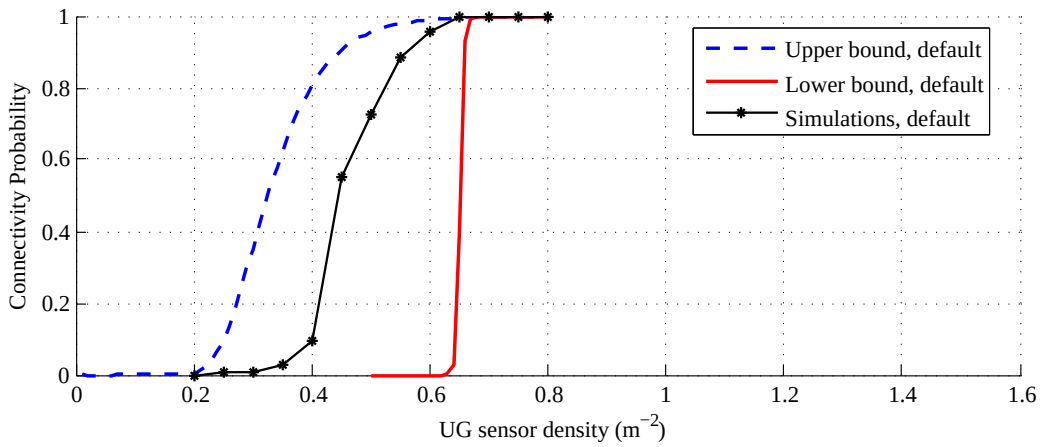


Figure 5: Connectivity probability in WUSNs as a function of UG sensor node density with default system and environmental parameters.

interval $[0.4, 0.6]$ m (i.e. the mean burial depth is 0.5 m). The density of the fixed AG sinks λ_a is $0.001m^{-2}$. There are 10 mobile AG sinks moving inside the region according to RWP model. The velocity of each flight is uniformly chosen from $[1, 2]$ m/s. The pause duration is uniformly chosen from $[0, 30]$ sec. The tolerable latencies are $t_s = t_c = 30$ sec in both the sensing phase and the control phase. All the transceivers in sensors and sinks are assumed to be the same. The transmitting power is 10 mW at 900 MHz. The minimum received power for correct demodulation is -90 dBm. The antenna gains $g_t = g_r = 5$ dB. The antenna heights of all AG fixed and mobile sinks are uniformly distributed in the interval $[0.8, 1.2]$ m (i.e. the mean antenna height is 1 m). In the soil medium, the sand particle percent is 50%. The clay percent is 15%. The bulk density is 1.5 grams/cm^3, and the solid soil particle density is 2.66 grams/cm^3. The volumetric water content (VWC) in the soil is randomly distributed according to a gamma distribution defined in (17), where the mean is $\mu_{m_v} = 8\%$ and the variance $\sigma^2_{m_v} = 10^{-4}$.

In Fig. 5 to Fig. 11, the theoretical upper and lower bounds are compared with the simulation results with various system and environmental parameters. Each simulated connectivity probability is calculated based on 500 simulation iterations. The lower and upper bounds are calculated by (46) and (57) respectively. As shown in Fig. 5 to Fig. 11, the

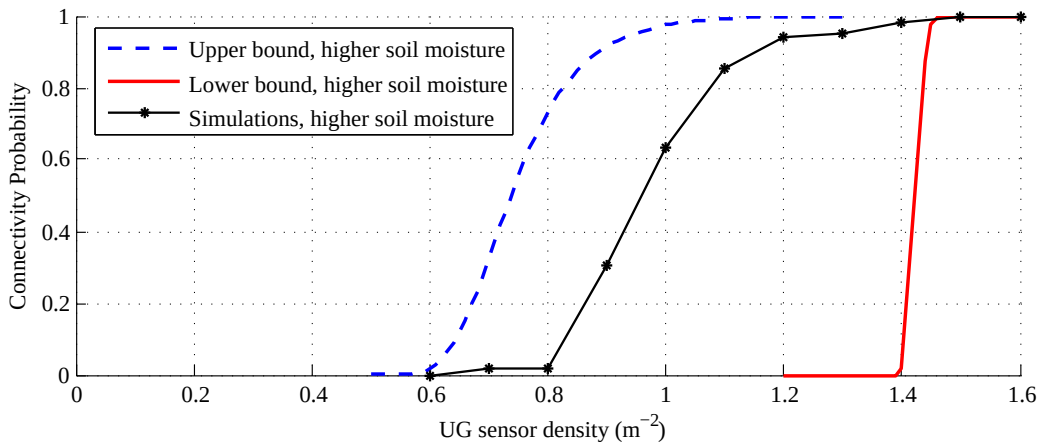


Figure 6: Connectivity probability in WUSNs as a function of UG sensor node density in soil medium with higher soil moisture (VWC=22%).

theoretical upper and lower bounds are valid in all the simulation scenarios. It should be noted that the upper bound is tighter than the lower bound, since the sufficient condition of the connectivity (lower bound) is more difficult to achieve than the necessary condition (upper bound).

Fig. 5 shows the upper bound, lower bound, and the simulation results of the connectivity in a WUSN with the default parameters. The connectivity probability increases as the UG sensor density increases. There exists a turning point in x-axis, where the WUSN has a high probability to be fully connected if the UG sensor density is larger than the turning point. This result is consistent with the connectivity analysis of terrestrial wireless networks [40]. In the following part of this section, the unique effects of various parameters of the WUSN system and the underground environments on the WUSN connectivity are discussed.

2.3.4.1 Soil Moisture

The effects of the higher soil moisture on the WUSNs' connectivity are illustrated in Fig. 6, where the connectivity probabilities are given as a function of UG sensor node density in soil medium with much higher soil moisture. Instead of the 8% mean VWC in default settings, the mean VWC in Fig. 6 is 22%. The variance $\sigma^2_{m_v}$ remains the same. It indicates

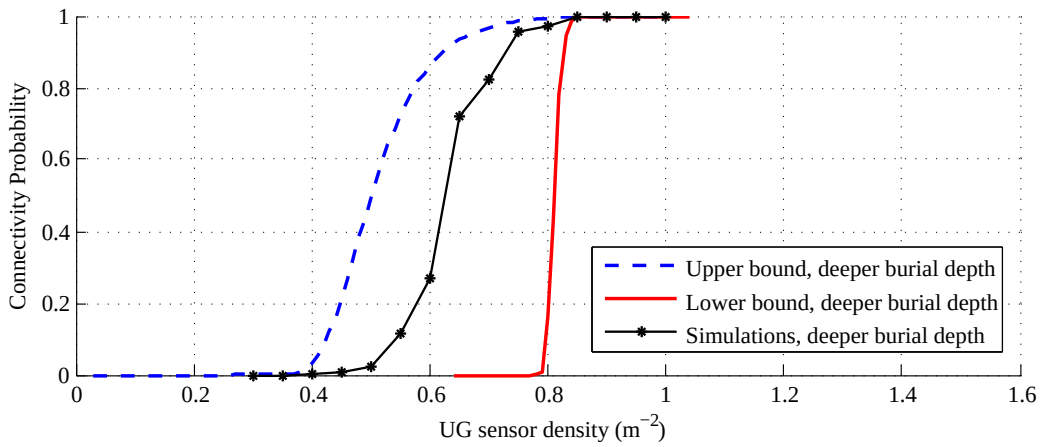


Figure 7: Connectivity probability in WUSNs as a function of UG sensor node density with deeper sensor burial depth (mean depth is 1 m).

that the connectivity in WUSNs highly depends on the soil moisture. To achieve equal connectivity probability, the UG sensor node density of the WUSN in wet soil ($\mu_{m_v} = 22\%$) is more than twice of the density required in dry soil ($\mu_{m_v} = 8\%$). This is because the transmission ranges of both the UG-UG and the UG-AG channel are significantly reduced when the water content in the soil increases, as discussed in Section III.

2.3.4.2 Sensor Burial Depth

In Fig. 7, the effects of the deeper sensor burial depth on the WUSNs' connectivity are captured, where the mean sensor burial depth is doubled, i.e. the burial depth is uniformly distributed in the interval $[0.8, 1.2]$ m. Similar to the influence of the soil moisture, the connectivity probability in WUSNs dramatically decreases if the sensor burial depth increases, since the transmission range of the UG-AG channel significantly decreases as sensor burial depth increases. Note that the impacts of the sensor burial depth are smaller that the impacts of the soil moisture, since the burial depth does not dramatically affect the UG-UG channel while the soil moisture influence both the UG-UG channel and the UG-AG channel.

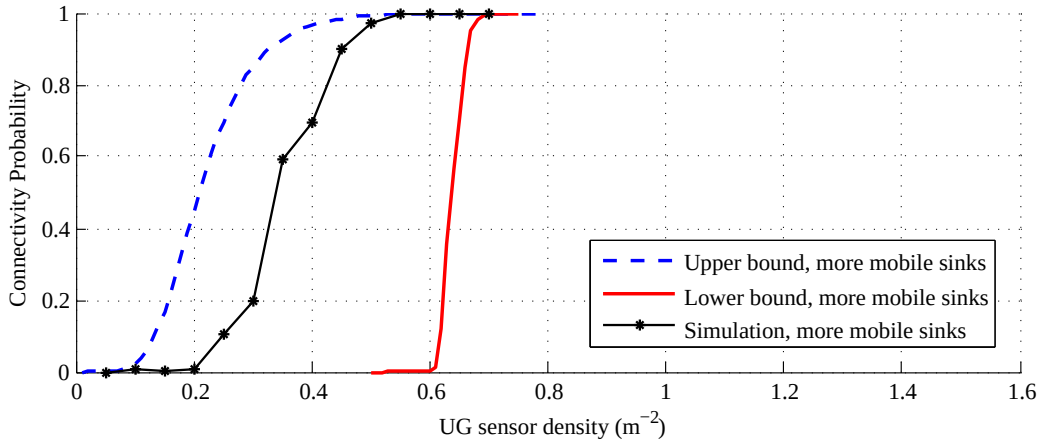


Figure 8: Connectivity probability in WUSNs as a function of UG sensor node density with four times more AG mobile sinks ($m = 50$).

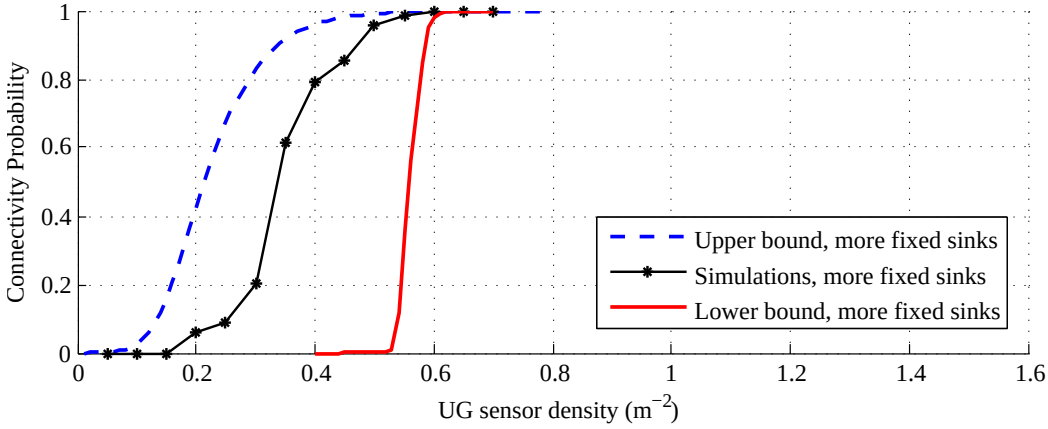


Figure 9: Connectivity probability in WUSNs as a function of UG sensor node density with two times AG fixed sink density ($\lambda_a = 0.002\ m^{-2}$).

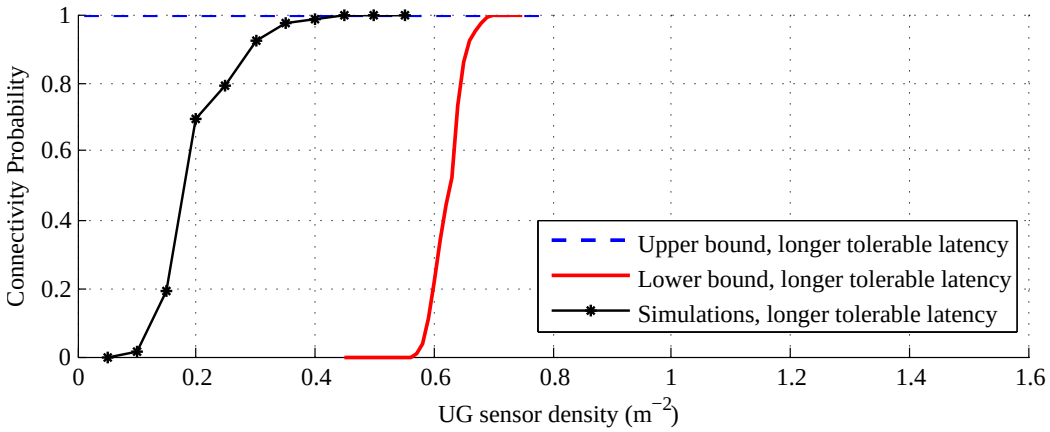


Figure 10: Connectivity probability in WUSNs as a function of UG sensor node density with longer tolerable latency (t_s = 300 sec).

2.3.4.3 Number of Mobile Sinks and Fixed Sink Density

In Fig. 8 and Fig. 9, the effects of mobile sink number and fixed sink density on the connectivity in WUSNs are investigated. Specifically, in Fig. 8, four times more AG mobile sinks are added in the monitored field (m = 50), while in Fig. 9, the density of the AG fixed sinks is doubled (λ_a = 0.002 m^{-2}) compared with the default parameters. It is shown that the connectivity probabilities increase if the number of mobile sinks or the fixed sink density increases, which can be explained by the definition of WUSN connectivity. With larger fixed sink density, both upper and lower bound of the connectivity probability dramatically increase. However, the lower bound of connectivity probability does not significantly increase with more mobile sinks because of the following reason. Due to the highly random mobility of the mobile sinks, the sufficient conditions (lower bound) are not becoming significantly easier to achieve with more mobile sinks.

2.3.4.4 Tolerable Latency and Sink Mobility

Fig. 10 shows the effect of the longer tolerable latency on the network connectivity, where the tolerable latency is prolonged from 30 sec to 300 sec. As expected, the connectivity

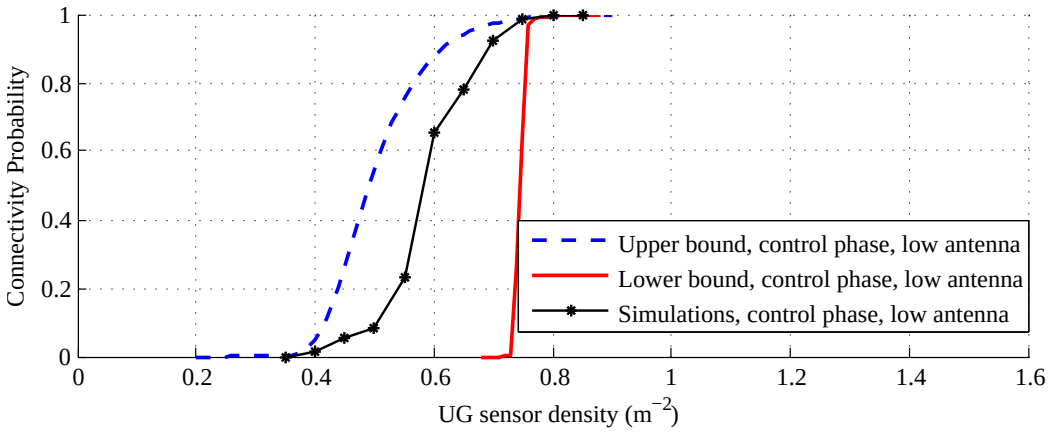


Figure 11: Connectivity probability in WUSNs as a function of UG sensor node density in control phase with lower sink antenna height (mean height is 0.2 m).

probability increases with longer tolerable latency. Therefore, there exists a tradeoff between the lower latency and higher connectivity probability. In Fig. 10, with the 300 sec tolerable latency, the upper bound of the WUSN connectivity probability become constant 100% since the mobile sink can move to any position in the monitored region within the prolonged tolerable latency in the best case. However, similar to the effects of the mobile sink number, the tolerable latency does not have obvious effects on the lower bound of the WUSN connectivity due to the highly random mobility model. It should be noted that the effects of the mobility model parameters (moving velocity and pause time), are similar to the tolerable latency, since the tolerable latency and mobility model parameters have equal effects in determining whether the mobile sink can move into the range of a UG sensor or not.

2.3.4.5 Connectivity in Control Phase

Due to the asymmetrical channel between the UG sensors and AG sinks, the connectivity performances of the sensing phase and the control phase are different. In Fig. 12, the connectivity probability of a WUSN with the default parameters in the control phase is shown as a function of the UG sensor density. Compared with the sensing phase, the

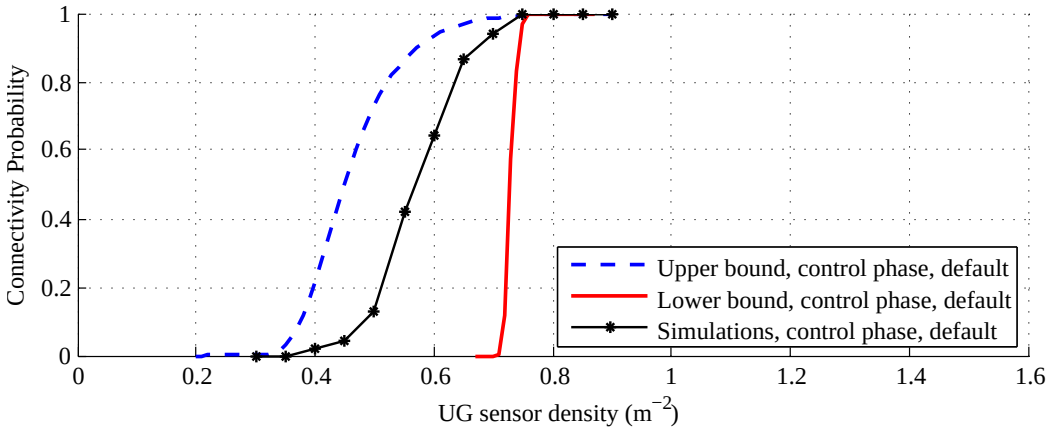


Figure 12: Connectivity probability in WUSNs as a function of UG sensor node density in control phase with default parameters.

connectivity probability in the control phase is obviously lower due to the following reason. In the control phase, the AG-UG channel is utilized. Since the transmission range of the AG-UG channel is much smaller than the UG-AG channel as discussed in Section III, the coverages of either the fixed sinks or the mobile sinks in the control phase are much smaller. Consequently, the connectivity probability decreases in the control phase.

The effects of all the system and environmental parameters on the WUSN connectivity in sensing phase are similar in control phase. Besides, the antenna height of the AG fixed and mobile sinks may influence the connectivity in WUSNs since the AG-UG channel is affected by the AG sink antenna heights. In Fig. 11, the connectivity in a WUSN with lower sink antenna heights is investigated, where the antenna heights of the AG fixed and mobile sinks are uniformly distributed in the interval $[0.1, 0.3]$ m, i.e. the mean antenna height is 0.2 m. Fig. 11 shows that the connectivity probability in WUSNs with lower sink antenna heights slightly decreases. The influence of the antenna heights is not as significant as the influence of the sensor burial depth since the path loss in the soil is much larger than the path loss in the air.

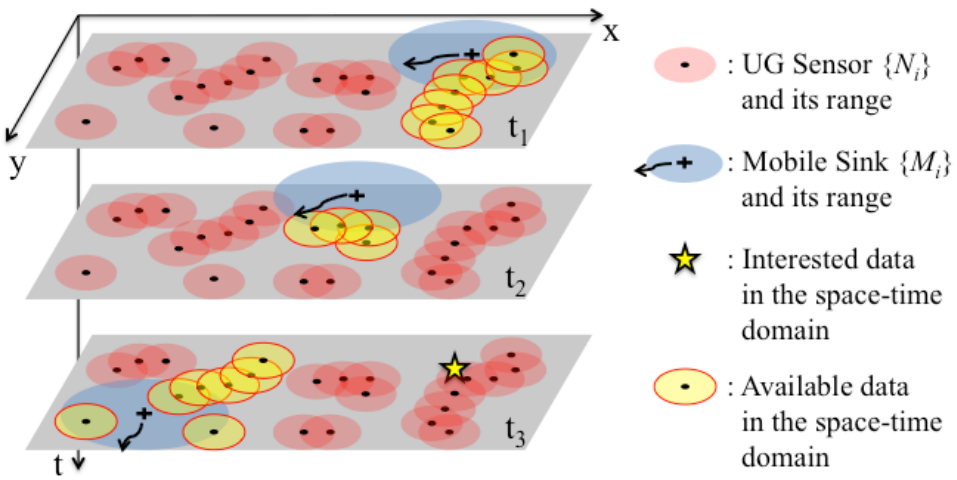


Figure 13: The snapshots of the dynamic network topology of the WUSN at three sequential time stamps in the space-time domain. (Only one of the multiple mobile sinks is plotted here for clear illustration.)

2.4 Spatio-Temporal Correlation-based Data Collection in WUSNs

In this section, we proposed a spatio-temporal correlation-based data collection scheme in WUSNs to release the unfeasible sensor density requirement of the full connectivity. By utilizing the spatio-temporal correlations and the AG mobile sinks, the WUSNs are not necessary to be fully connected so that the UG sensor density can be reduced. Meanwhile, the WUSN is divided into multiple unconnected clusters due to the reduced density of the UG sensors. The number of the clusters can range from one (fully connected) to the number of all UG sensors (totally isolated), depending on the UG sensor density. The UG sensors in an unconnected cluster in the network can only report their data when any one of the UG sensors in the same cluster connects to an AG mobile sink.

The scheme of the spatio-temporal correlation-based data collection in WUSNs is illustrated in Fig. 13, where the dynamic topologies of the WUSNs are plotted in the space-time domain. When an AG mobile sink moves into the communication range of one UG sensor, the communication is initiated by the data request from the AG mobile sink. This request is then broadcasted by this connected UG sensor to all other UG sensors in the same cluster. Finally, all the UG sensors in this cluster report their measurement data to the AG mobile

sink in a multi-hop fashion. Due to the reduced network connectivity and the usage of AG mobile sinks, not every UG sensor's data is available at the monitoring center at every time stamp. The time stamp when a certain data is available depends on the network connectivity and the sink mobility. Moreover, not every position in the field has a UG sensor due to the limited sensor density. Since all the monitored data are spatio-temporally correlated, the unavailable data at any interested locations and time stamps can be estimated by the least-squares linear regression (kriging) algorithms.

2.4.1 Sensor Density Optimization in WUSNs

After the spatio-temporal data collection scheme is derived, we develop the analytical solution for sensor density optimization in WUSNs under this data collection scheme. The network model and spatio-temporal correlation model are first described. Then the sensor density optimization is formalized and decomposed into a network connectivity analysis and a sink mobility analysis. Finally, the optimal sensor density is explicitly expressed as a function of multiple system and environmental parameters.

2.4.1.1 Network Model

The network model is similar to the model used in Section 2.3. The only difference is that the fixed AG sinks are not considered since the the underground sensors directly covered by a fixed AG sink do not need the spatio-temporal data collection scheme. The m mobile AG sinks are carried by people, machineries, or robots inside region $\mathbb{R}^2$. The movement of those mobile sinks can be either random (if they are carried by the people) or under control (if they are carried by machineries or robots).

2.4.1.2 Spatio-Temporal Correlation Model

By using the least-squares linear regression (kriging) algorithms [39], the unavailable data $z(\boldsymbol{x}, t)$ at an interested location $\boldsymbol{x}$ and time stamp t can be estimated. Assuming that the mean $E[z(\boldsymbol{x}, t)] = \mu$ is constant throughout the region $\mathbb{R}^2$ at arbitrary time stamp. If there

are n sensors, the unbiased estimated data $z^*(\boldsymbol{x}, t)$ can be expressed as a linear combination of the latest available measurement data of all the n UG sensors $\{z(\boldsymbol{x}_i, t_i), i = 1, 2, ... n\}$:

$$z_n^*(\boldsymbol{x}, t) = \sum_{i=1}^{n} \alpha_i \Big[z(\boldsymbol{x}_i, t_i) - \mu \Big] + \mu \ , \tag{58}$$

where the weights $\{\alpha_1, \alpha_2, ..., \alpha_n\}$ are determined to minimize the error variance $\sigma^2 = \text{Var}\{z^*(\boldsymbol{x}, t) - z(\boldsymbol{x}, t)\}$. The optimal weights $\{\alpha_1^{opt}, \alpha_2^{opt}, ..., \alpha_n^{opt}\}$ can be obtained by setting to zero each of the n partial first derivatives $\{\frac{\partial \sigma^2}{\partial \alpha_i}, i = 1, 2, ..., n\}$. Then the minimum error variance can be derived as

$$\sigma_n^2 = C(0) \Big[1 - \sum_{i=1}^{n} \alpha_i^{opt} \cdot \rho(\boldsymbol{x}_i, \boldsymbol{x}, \Delta t_i) \Big] \ , \tag{59}$$

where $C(0)$ is the variance of the data; $\rho(\boldsymbol{x}_i, \boldsymbol{x}, \Delta t_i)$ is the correlation function between the data at the space-time coordinate $(\boldsymbol{x}_i, t_i)$ and the data at the coordinate $(\boldsymbol{x}, t)$; $\Delta t_i = t - t_i$.

The mean, variance, and correlation function of the monitored data can be derived by the spatio-temporal models, which varies from case to case in different applications. In this section, we use the soil moisture as an example of the monitored data. Note that the monitored data can be easily changed to other physical quantity (such as temperature and vibration) by changing its statistical parameters. For soil moisture, the mean, variance, and correlation function are provided in [43, 30]:

$$\begin{aligned} \mu = \frac{b 2\pi\zeta}{a {r_R}^2 \eta\beta} \ , \quad C(0) = \frac{4\pi\zeta}{\eta\beta^2 {r_R}^2} \frac{b^2}{a(\eta + a)} \ , \\ \rho(\boldsymbol{x}_i, \boldsymbol{x}_j, \Delta t_{ij}) = \frac{\eta e^{-a\Delta t_{ij}} - a e^{-\eta \Delta t_{ij}}}{\eta - a} \left(1 + \frac{r_R d_{ij}}{4} \right) e^{-r_R \frac{d_{ij}}{2}} \ , \end{aligned} \tag{60}$$

where ζ is the intensity of the Poisson rain/irrigation process; a is the normalized soil water loss; b is the rain/irrigation coefficient; $1/r_R$, $1/\eta$, and $1/\beta$ are the mean cell radius, duration, and intensity of each rain/irrigation, respectively; d_{ij} is the distance between the two locations $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$; $\Delta t_{ij} = t_i - t_j$. By substituting (60) into (58) and (59), the monitored data at any coordinate in the space-time domain can be estimated and the corresponding estimation error can be calculated.

2.4.1.3 Optimization Problem Formalization

The optimal sensor density is actually the minimum sensor density that can guarantee a certain level of overall monitoring accuracy in the WUSNs. The monitoring accuracy of a certain position at a certain time stamp is measured by the estimation error given in (59). Hence, the overall monitoring accuracy in a WUSN is measured by the average error of every position and every time stamp through out the WUSN, which is denoted as $E[\sigma^2]$. Then the sensor density optimization problem can be formalized as

$$\begin{aligned} Given: &\quad \text{Underground channel conditions,} \\ &\quad \text{Spatio-temporal correlation model,} \\ &\quad \text{Number and mobility model of mobile sinks.} \\ Find: &\quad \min \lambda \\ s.t.: &\quad E[\sigma^2] < \sigma^2_{max}, \end{aligned} \tag{61}$$

where σ^2_{max} is the maximum tolerable mean error. To solve this optimization problem, we first calculate the objective function $E[\sigma^2]$. Due to the highly random network topology, it is impossible to find out the exact expression of $E[\sigma^2]$. Hence, we use the upper bound of $E[\sigma^2]$ as the new objective function, which can guarantee the required monitoring accuracy.

Before the calculation, notations are described first: t_{now} is the current time stamp; $\{\boldsymbol{a} \longleftrightarrow \boldsymbol{b}\}$ denotes that the sensor located at position $\boldsymbol{a}$ is connected to the sensor at position $\boldsymbol{b}$ by single or multiple hops; $\{\boldsymbol{a} \longleftrightarrow \text{sink at } \Delta t\}$ denotes the event that, at time stamp $t_{now} - \Delta t$, the sensor at $\boldsymbol{a}$ is connected to a mobile sink for the last time by single or multiple hops; and $\{\boldsymbol{a} \overset{\text{direct}}{\longleftrightarrow} \text{sink at } \Delta t\}$ denotes the event that, at time stamp $t_{now} - \Delta t$, the sensor at $\boldsymbol{a}$ is directly covered by a mobile sink for the last time. Then the average error is given by

$$E[\sigma^2] = E\left[E[\sigma^2 \big| n \text{ sensors}]\right] = \sum_{n=1}^{\infty} E[\sigma_n^2] \frac{(\lambda S_{\mathbb{R}^2})^n}{n!} e^{-\lambda S_{\mathbb{R}^2}}, \tag{62}$$

where the probability that there are n sensors is calculated according to the Poisson point process of the sensor distribution; $S_{\mathbb{R}^2}$ is the area of the region $\mathbb{R}^2$; $E[\sigma_n^2]$ is calculated by

(59). To avoid calculating the partial derivatives to get the optimal weights $\{\alpha_1^{opt}, \alpha_2^{opt}, ..., \alpha_n^{opt}\}$ in (59), a simple weight setting $\{\alpha_1^*, \alpha_2^*, ..., \alpha_n^*\}$ is used to calculate the upper bound of the error variance, where $\alpha_i^* = 1$ if $\boldsymbol{x}_i$ is the closest to $\boldsymbol{x}$ among the n sensors, otherwise $\alpha_i^* = 0$. This weight setting is equal to the commonly used strategy that the data at the closest sensor is utilized. Then

$$E[\sigma_n^2] \leq 2C(0)\left\{1 - E\Big[\rho(\boldsymbol{x}_i, \boldsymbol{x}, \Delta t_i)\Big|\boldsymbol{x}_i \text{ is closest to } \boldsymbol{x}, i \in \{1, ..., n\}\Big]\right\} . \tag{63}$$

We first define three events as follows:

A : One of the n sensors is located at $\boldsymbol{x}'$;

B : This sensor $\longleftrightarrow$ sink at $\Delta t \leq \Delta t'$;

C : This sensor is closest to $\boldsymbol{x}$ than any other $n-1$ sensors.

Then,

$$\begin{aligned} & E\Big[\rho(\boldsymbol{x}_i, \boldsymbol{x}, \Delta t_i)\Big|\boldsymbol{x}_i \text{ is closest to } \boldsymbol{x}, i \in \{1, ..., n\}\Big] \\ & \geq E\Big[\rho(\boldsymbol{x}', \boldsymbol{x}, \Delta t')\Big|A, B, C\Big] \\ & = \int_{\boldsymbol{x}\in\mathbb{R}^2} \frac{1}{S_{\mathbb{R}^2}} \int_0^{\Delta t_{max}} \frac{1}{\Delta t_{max}} \int_{\boldsymbol{x}'\in\mathbb{R}^2} \rho(\boldsymbol{x}', \boldsymbol{x}, \Delta t')\, f_{\boldsymbol{x}'}(A,B,C)\, d\boldsymbol{x}\, d\boldsymbol{x}'\, d\Delta t' , \end{aligned} \tag{64}$$

where Δt_{max} is the maximum usable time deviation, i.e. for $\Delta t > \Delta t_{max}$, the temporal correlation is very small.

The probability density function (pdf) of the conditions $\{A, B, C\}$ can be calculated as

$$f_{\boldsymbol{x}'}(A,B,C) = f_{\boldsymbol{x}'}(A) \cdot P(C|A) \cdot P(B\,|\,A,C) , \tag{65}$$

where

$$f_{\boldsymbol{x}'}(A) = \sum_{i=1}^{n} f(\text{sensor } N_i \text{ is located at } \boldsymbol{x}') = n\frac{1}{S_{\mathbb{R}^2}} , \tag{66}$$

since the UG sensors are distributed according to the homogeneous Poisson point process.

$$P(C|A) = P(C_1, C_2, ..., C_{n-1}) , \tag{67}$$

where C_i denotes the event that the i^{th} sensor of the $n-1$ sensors is outside the region $\mathbb{C}^2$; $\mathbb{C}^2$ is the circular region centered at $\boldsymbol{x}$ with radius $d_{xx'} = \|\boldsymbol{x}' - \boldsymbol{x}\|$. Since all the UG sensors are distributed according to a Poisson point process, $\{C_1, ..., C_{n-1}\}$ are independent events. Then

$$\begin{aligned} P(C_1, C_2, ..., C_{n-1}) &= P(C_1) \cdot P(C_2) \cdots P(C_{n-1}) \\ &= P^{n-1}(C_i) = \Big(1 - \frac{S_{\mathbb{C}^2 \cap \mathbb{R}^2}}{S_{\mathbb{R}^2}}\Big)^{n-1} , \end{aligned} \tag{68}$$

where $S_{\mathbb{C}^2 \cap \mathbb{R}^2}$ is the area of the joint region of $\mathbb{C}^2$ and $\mathbb{R}^2$.

So far, the first two terms $f_{\boldsymbol{x}'}(A)$ and $P(C|A)$ in (65) is calculated. In the last term in (65), the event $\{B \,|\, A,C\}$ is equivalent to the event B', where

B' : $\boldsymbol{x}' \longleftrightarrow$ sink at $\Delta t \leq \Delta t'$ without using relay sensors inside region $\mathbb{C}^2$;

If the positions of the other $n-1$ sensors is denoted as $\{\boldsymbol{x}_S^1, \boldsymbol{x}_S^2, ..., \boldsymbol{x}_S^{n-1}\}$, event B' can be further developed as the union of a set of sub-events $B'_0 \cup \{B'_i, i = 1, 2, ..., n-1\}$:

B'_0 : $\boldsymbol{x}' \overset{\text{direct}}{\longleftrightarrow}$ sink at $\Delta t \leq \Delta t'$;

B'_i : $\boldsymbol{x}' \longleftrightarrow \boldsymbol{x}_S^i$ not via relay sensors inside region $\mathbb{C}^2$, and $\boldsymbol{x}_S^i \overset{\text{direct}}{\longleftrightarrow}$ sink at $\Delta t \leq \Delta t'$.

Then

$$P(B \,|\, A,C) = P(B') = P(B'_0) + \big[1 - P(B'_0)\big] \cdot P(\cup_{i=1}^{n-1} B'_i) , \tag{69}$$

Due to the homogeneous sensor distribution, $P(B'_1) = P(B'_2) = ... = P(B'_{n-1}) = P(B'_i)$. Hence,

$$P(B \,|\, A,C) \simeq P(B'_0) + \big[1 - P(B'_0)\big] \cdot (n-1) \cdot P(B'_i) , \tag{70}$$

By substituting (63)-(70) into (62) and using the identical equation $\sum_{n=0}^{\infty} \frac{\lambda^n}{n!} e^{-\lambda} \equiv 1, \forall \lambda$, we derive:

$$E[\sigma^2] \leq 2C(0)\Big\{1 - E[\text{closest } \rho]\Big\} , \tag{71}$$

where

$$E[\text{closest } \rho] \geq \frac{\lambda}{\Delta t_{max} S_{\mathbb{R}^2}} \iiint_{\substack{\boldsymbol{x},\boldsymbol{x}' \in \mathbb{R}^2 \\ \Delta t' \in [0, \Delta t_{max}]}} \rho(\boldsymbol{x}', \boldsymbol{x}, \Delta t') \cdot e^{-\lambda S_{\mathbb{C}^2 \cap \mathbb{R}^2}} \cdot \left\{ P(B_0') + \lambda \cdot (S_{\mathbb{R}^2} - S_{\mathbb{C}^2 \cap \mathbb{R}^2}) \cdot [1 - P(B_0')] \cdot P(B_i') \right\} \cdot d\boldsymbol{x} d\boldsymbol{x}' d\Delta t' \ , \tag{72}$$

where

$$P(B_0') = P(\boldsymbol{x}' \overset{\text{direct}}{\longleftrightarrow} \text{sink at } \Delta t \leq \Delta t') \ , \tag{73}$$

$$P(B_i') = \frac{1}{S_{\mathbb{R}^2} - S_{\mathbb{C}^2 \cap \mathbb{R}^2}} \int_{\boldsymbol{x}_S \in \mathbb{R}^2 - \mathbb{C}^2} P(\boldsymbol{x}_S \overset{\text{direct}}{\longleftrightarrow} \text{sink at } \Delta t \leq \Delta t') \cdot P(\boldsymbol{x}' \leftrightarrow \boldsymbol{x}_S \text{ not via } \mathbb{C}^2) \, d\boldsymbol{x}_S \, , \tag{74}$$

where $P(\boldsymbol{x}' \leftrightarrow \boldsymbol{x}_S \text{ not via } \mathbb{C}^2)$ is the probability that the sensor at $\boldsymbol{x}'$ is connected to the other sensor at $\boldsymbol{x}_S$ without using relay sensors inside region $\mathbb{C}^2$.

According to (71) and (72), the upper bound of the average monitoring error $E[\sigma^2]$ is determined by the correlation function $\rho(\boldsymbol{x}_i, \boldsymbol{x}_j, \Delta t_{ij})$ in (60) and the probabilities $P(B_0')$ and $P(B_i')$ in (73) and (74). To calculate $P(B_0')$ and $P(B_i')$, two probabilities, $P(\boldsymbol{x}' \leftrightarrow \boldsymbol{x}_S \text{ not via } \mathbb{C}^2)$ and $P(\boldsymbol{y} \overset{\text{direct}}{\longleftrightarrow} \text{sink at } \Delta t \leq \Delta t')$ need to be analyzed, where $\boldsymbol{y} = \boldsymbol{x}'$ or $\boldsymbol{x}_S$. Hence, the sensor density optimization problem is decomposed into a network connectivity probability analysis on $P(\boldsymbol{x}' \leftrightarrow \boldsymbol{x}_S \text{ not via } \mathbb{C}^2)$ and a sink mobility analysis on $P(\boldsymbol{y} \overset{\text{direct}}{\longleftrightarrow} \text{sink at } \Delta t \leq \Delta t')$.

2.4.1.4 Network Connectivity Analysis

In this subsection, the lower bound of the probability $P(\boldsymbol{x}' \leftrightarrow \boldsymbol{x}_S \text{ not via } \mathbb{C}^2)$ is calculated. The notations are first described: $\lceil a \rceil_{rnd}$ means rounding a to the nearest integer; and $\lceil a \rceil$ means rounding a to the nearest integer $\geq a$.

Proposition 1. *The lower bound of the probability that a UG sensor located at* $\mathbf{x}'$ *is connected to another UG sensor located at* $\mathbf{x}_S$ *by single or multi-hops without using the relay nodes inside the circular region* $\mathbb{C}^2$ *is given by*

$$P(\mathbf{x}' \leftrightarrow \mathbf{x}_S \textit{ not via } \mathbb{C}^2) \geq \left(1 - e^{-\frac{2}{5}\lambda R_{UG-UG}^2}\right)^{\epsilon(\mathbf{x}', \mathbf{x}_S, \mathbf{x})} \ , \tag{75}$$

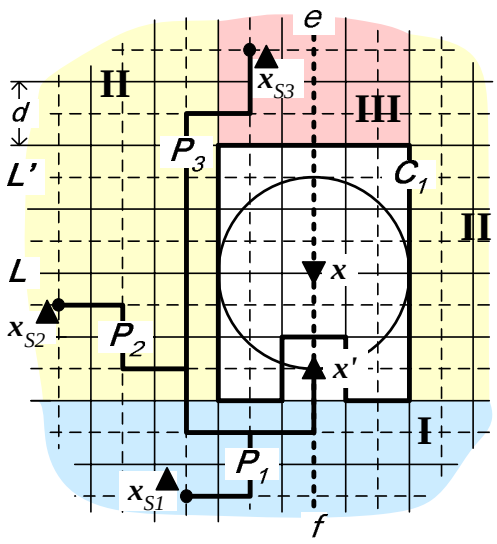


Figure 14: Mapping the WUSN on a lattice L (plain) and its dual L' (dashed).

where R_{UG-UG} is the communication range of the UG-UG channel that is derived in Section 2.2; the detailed expression of the function $\epsilon(\mathbf{x}', \mathbf{x}_S, \mathbf{x})$ is given in the proof.

Proof. We use the similar strategy in Section 2.3.2.1 to proof this proposition. First map the WUSN on a square lattice L (plain) and its dual L' (dashed), as shown in Fig. 14. The vertices of L' are placed in the center of every square of L. The edges of L' cross every edge of L. Hence, there exists a one-to-one relation between the edges of L and L'. L and L' have the same edge length $d = \frac{1}{\sqrt{5}} R_{UG-UG}$. The edge length is designed so that two UG sensors deployed in two adjacent squares of the lattice L are guaranteed to be connected to each other. One vertex of the dual lattice is located at $\boldsymbol{x}'$. The straight line ef connecting $\boldsymbol{x}'$ and $\boldsymbol{x}$ forms a sequence of vertical edges of the dual lattice L', as shown in Fig. 14.

Given an open path of the L', all the UG sensors in the squares in L along this open path belong to the same cluster and are connected to each other. The states of the edges (open or closed) are independent from each other. The probability that an edge is closed is denoted as q. According to the Poisson point process of the UG sensors:

$$q = P^2(\text{No sensor in a square}) = e^{-\frac{2}{5}\lambda R_{UG-UG}^2} \ . \tag{76}$$

If there is an open path of L' connecting the two squares in L where $\boldsymbol{x}$ and $\boldsymbol{x}_S$ are located, these two sensors are guaranteed to be connected. The set of all open paths connecting $\boldsymbol{x}$ and $\boldsymbol{x}_S$ without using the relay nodes inside region $\mathbb{C}^2$ is denoted as $\{P_i^o, i = 1, 2, ...\}$, then

$$P(\boldsymbol{x}' \leftrightarrow \boldsymbol{x}_S \text{ not via } \mathbb{C}^2) = P\Big(\cup_i^\infty P_i^o \Big) \geq P(P_i^o) \ , \tag{77}$$

where the probability of a certain open path $P(P_i^o)$ is used as the lower bound of $P(\boldsymbol{x}' \leftrightarrow \boldsymbol{x}_S$ not via $\mathbb{C}^2)$. To maximize the lower bound, the shortest open path is selected. Hence, the probability of the shortest open path in L' connecting $\boldsymbol{x}'$ and $\boldsymbol{x}_S$ is calculated as the lower bound of $P(\boldsymbol{x}' \leftrightarrow \boldsymbol{x}_S$ not via $\mathbb{C}^2)$:

$$P(\boldsymbol{x}' \leftrightarrow \boldsymbol{x}_S \text{ not via } \mathbb{C}^2) \geq (1-q)^{\epsilon(\boldsymbol{x}', \boldsymbol{x}_S, \boldsymbol{x})} \ , \tag{78}$$

where q is the close edge probability given in (76); $\epsilon(\boldsymbol{x}', \boldsymbol{x}_S, \boldsymbol{x})$ is the length of the shortest open path connecting $\boldsymbol{x}'$ and $\boldsymbol{x}_S$.

The shortest open path may not be a simple straight line since relay nodes can not be inside region $\mathbb{C}^2$. As shown in Fig. 14, a rectangular circuit C_1 is set up so that if the open path does not via the squares inside C_1, the relay nodes along the open path are guaranteed to be outside of circular region $\mathbb{C}^2$. The width and length of the rectangular circuit C_1 are $w_c d$ and $l_c d$, respectively, where

$$\begin{aligned} w_c &= \left\lceil 2\sqrt{5}\, d_{xx'}/R_{UG-UG} + 0.5 \right\rceil \ , \\ l_c &= 2\left\lceil \sqrt{5}\, d_{xx'}/R_{UG-UG} - 0.5 \right\rceil + 1 \ , \end{aligned} \tag{79}$$

where $d_{xx'} = \|\boldsymbol{x}' - \boldsymbol{x}\|$ is the distance between $\boldsymbol{x}'$ and $\boldsymbol{x}$, which is the radius of region $\mathbb{C}^2$. Construct a new Cartesian coordinate by setting $\boldsymbol{x}'$ as the origin, ef as the $y-$axis ($\boldsymbol{x}$ is on the positive of y-axis). The new coordinate of $\boldsymbol{x}_S$ is $(x_S^{\text{new}}, y_S^{\text{new}})$. As shown in Fig. 14, the possible positions of $\boldsymbol{x}_S$ are divided into three regions. In different regions, the shortest path connecting $\boldsymbol{x}'$ and $\boldsymbol{x}_S$ is different, e.g. path P_1, P_2, and P_3 in Fig. 14. The length of the

shortest path is $\epsilon(\boldsymbol{x}', \boldsymbol{x}_S, \boldsymbol{x})$, and

$$\epsilon(\boldsymbol{x}', \boldsymbol{x}_S, \boldsymbol{x}) = \begin{cases} \left\lceil \frac{\sqrt{5}\,|x_S^{\text{new}}|}{R_{UG-UG}} \right\rceil_{rnd} + \left\lceil \frac{\sqrt{5}\,|y_S^{\text{new}}|}{R_{UG-UG}} \right\rceil_{rnd}, & \text{if } \boldsymbol{x}_S \in \text{Region I} \\ 2 + \left\lceil \frac{\sqrt{5}\,|x_S^{\text{new}}|}{R_{UG-UG}} \right\rceil_{rnd} + \left\lceil \frac{\sqrt{5}\,|y_S^{\text{new}}|}{R_{UG-UG}} \right\rceil_{rnd}, & \text{if } \boldsymbol{x}_S \in \text{Region II} \\ l_c + 3 - \left\lceil \frac{\sqrt{5}\,|x_S^{\text{new}}|}{R_{UG-UG}} \right\rceil_{rnd} + \left\lceil \frac{\sqrt{5}\,|y_S^{\text{new}}|}{R_{UG-UG}} \right\rceil_{rnd}, & \text{if } \boldsymbol{x}_S \in \text{Region III} \end{cases}, \quad (80)$$

where the regions are defined as follows:

Region I : $y_S^{\text{new}} \leq -\frac{1}{2}d$;

Region II : $y_S^{\text{new}} > -\frac{1}{2}d$ and $|x_S^{\text{new}}| \geq \frac{1}{2}l_c d$;

Region III : $y_S^{\text{new}} > (w_c - \frac{1}{2})d$ and $|x_S^{\text{new}}| < \frac{1}{2}l_c d$.

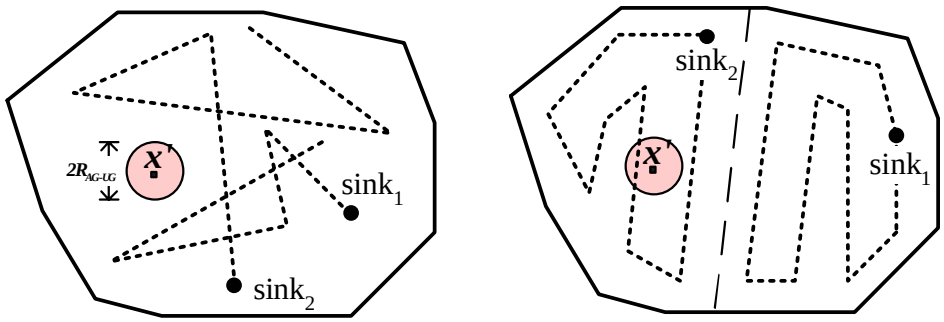
Since $\boldsymbol{x}_S$ cannot appear inside the circuit C_1, there is only one undiscussed region for $\boldsymbol{x}_S$ in the plain: the square in L that contains $\boldsymbol{x}'$, where $\epsilon(\boldsymbol{x}', \boldsymbol{x}_S, \boldsymbol{x}) \equiv 0$. Finally, substituting (80) into (78) completes the proof. □

2.4.1.5 Sink Mobility Analysis

In this subsection, we analyze the random and controlled mobility of the AG mobile sinks to derive the probability $P(\boldsymbol{y} \overset{\text{direct}}{\longleftrightarrow} \text{sink at } \Delta t \leq \Delta t')$ in (73) and (74). Before the analysis, it should be noted that due to the query-based data collection scheme, the effective communication range of the mobile sink is R_{AG-UG}, which is a function of the sensor burial depth as discussed in Section 2.2.

Random Sink Mobility:

In most WUSN applications, the mobile sinks are carried by the people and vehicles working inside the monitored field. As shown in Fig. 15(a), the mobility of those people and vehicles can be modeled by the widely used Random Waypoint (RWP) Model [42], which is described in Section 2.3.1.2. Using the same strategy in Section 2.3.2.2, we have the following proposition.



(a) (b)

Figure 15: (a) Random sink mobility and (b) controlled sink mobility.

Proposition 2. *Given m mobile sinks in region $\mathbb{R}^2$, at time stamp $t_{now} - \Delta t$, the sensor at coordinate* $\mathbf{y}$ *is directly covered by a mobile sink for the last time. Then, the probability that $\Delta t \leq \Delta t'$ is lower bounded by*

$$P(\mathbf{y} \overset{direct}{\longleftrightarrow} sink \text{ at } \Delta t \leq \Delta t') \geq 1 - \gamma^{\lfloor \frac{\Delta t'}{t_D} \rfloor} , \tag{81}$$

where

$$\gamma = \frac{1}{z_{max} - z_{min}} \int_{z_{min}}^{z_{max}} \left(\int_{\mathbf{x} \in \mathbb{R}^2 - \mathbb{C}_S^2(\mathbf{y})} \xi(\mathbf{x}) \, d\mathbf{x} \right)^m dz' ,$$

$$t_D = 2(\tau_{max} + D/v_{min}) , \tag{82}$$

where z' is the burial depth of the sensor; $\xi(\mathbf{x})$ is the pdf that a sink visit position $\mathbf{x}$ *at arbitrary time stamp (stationary node distribution) given in [42]; $\mathbb{C}_S^2(\mathbf{y})$ is the circular region centered at* $\mathbf{y}$ *with radius R_{AG-UG}; D is the maximum flight length in the convex region $\mathbb{R}^2$; v_{min} and τ_{max} are the minimum moving speed and the maximum pause time of the sink, respectively*

Controlled Sink Mobility:

Since the randomly moving sinks are inefficient to collect data, dedicated robots may be employed in certain applications to improve the data collection efficiency. In this section, we adopt the most straightforward strategy to control the multiple robots: The whole region $\mathbb{R}^2$ is divided into m subregions with equal area. Each robot moves inside one of the m

subregions with fixed loop route covering the whole subregion, as shown in Fig. 15(b). Note that the control strategy of the multiple robots may be further optimized. However, it is out the scope of this thesis.

In the proposed sink control strategy, each mobile sink collects data from a different subregion with equal area. Hence, the m sinks are uniformly allocated in the whole monitored region. In each subregion, the loop route with minimum length l_{route} is designed for each sink to cover every position.

$$l_{route} \leq \frac{S_{\mathbb{R}^2}/m}{R_{AG-UG}(z_{max})} \ . \tag{83}$$

where $R_{AG-UG}(z_{max})$ is the communication range of AG-UG channel when the sensor is buried at the maximum depth.

Assuming that all the mobile sink moves at a constant velocity v_{robot} without pause. Then the time duration for a sink to complete one loop route in the subregion is $T_{route} = \frac{l_{route}}{v_{robot}}$. That means the UG sensor at any position inside the monitored region can be covered by a sink at least once in every period T_{route}. Therefore, for controlled AG sinks,

$$P(\boldsymbol{y} \overset{\text{direct}}{\longleftrightarrow} \text{sink at } \Delta t \leq \Delta t') \tag{84}$$
$$\begin{cases} \geq \frac{\Delta t'}{T_{route}} \geq \frac{\Delta t' m\, v_{robot} R_{AG-UG}(z_{max})}{S_{\mathbb{R}^2}} \ , & \text{if } 0 \leq \Delta t' < \frac{S_{\mathbb{R}^2}}{m\, v_{robot} R_{AG-UG}(z_{max})} \\ = 1 \ , & \text{otherwise} \end{cases} .$$

2.4.1.6 Sensor Density Optimization Solution

Substituting (75), (81), and (84) into (71)-(74) yields the upper bound of the average monitoring error in WUSNs with random or controlled sink mobility, which is denoted as $\overline{E}[\sigma^2]$,

$$\begin{aligned} \overline{E}[\sigma^2] =& 2C(0) - \frac{2\lambda \cdot C(0)}{\Delta t_{max} S_{\mathbb{R}^2}} \iiint\limits_{\substack{\boldsymbol{x},\boldsymbol{x}' \in \mathbb{R}^2 \\ \Delta t' \in [0, \Delta t_{max}]}} \rho(\boldsymbol{x}', \boldsymbol{x}, \Delta t') \cdot e^{-\lambda S_{\mathbb{C}^2 \cap \mathbb{R}^2}} \\ & \cdot \Big\{ F_2(\boldsymbol{x}', \Delta t') + \lambda \cdot [1 - F_2(\boldsymbol{x}', \Delta t')] \\ & \cdot \int_{\boldsymbol{x}_S \in \mathbb{R}^2 - \mathbb{C}^2} F_2(\boldsymbol{x}_S, \Delta t') \cdot F_1(\boldsymbol{x}', \boldsymbol{x}_S, \boldsymbol{x})\, d\boldsymbol{x}_S \Big\} \cdot d\boldsymbol{x} d\boldsymbol{x}' d\Delta t' \ , \end{aligned} \tag{85}$$

where the function $F_1(\boldsymbol{x}', \boldsymbol{x}_S, \boldsymbol{x})$ and $F_2(\boldsymbol{y}, \Delta t')$, $\boldsymbol{y} = \boldsymbol{x}'$ or $\boldsymbol{x}_S$, are expressed as

$$F_1(\boldsymbol{x}', \boldsymbol{x}_S, \boldsymbol{x}) = \left(1 - e^{-\frac{2}{5}\lambda R_{UG-UG}^2}\right)^{\epsilon(\boldsymbol{x}', \boldsymbol{x}_S, \boldsymbol{x})} \quad ; \tag{86}$$

$$F_2(\boldsymbol{y}, \Delta t') = \begin{cases} 1 - \gamma^{\lfloor \frac{\Delta t'}{T_D} \rfloor}, & \text{if random sink mobility} \\ \frac{\Delta t' m \, v_{robot} \, R_{AG-UG}(z_{max})}{S_{\mathbb{R}^2}}, & \text{if controlled sink mobility} \end{cases} .$$

By substituting (85) into (61), the optimal sensor density in WUSNs with random or controlled sink mobility is derived:

$$\lambda^{opt} = \min \left\{ \lambda : \overline{E}[\sigma^2] > \sigma_{max}^2 \right\} . \tag{87}$$

2.4.2 Numerical Analysis

In this section, we numerically analyze the effects of multiple system configurations and environmental conditions on the optimal sensor density in WUSNs.

Except studying the effect of certain parameters, the default values are set as follows: The monitored region is a $100\,m \times 100\,m$ square. The UG sensors are deployed according to a homogeneous Poisson point process of spatial intensity λ with random burial depths. The burial depths are uniformly distributed in the interval $[0.3, 0.7]$ m (i.e. the mean burial depth is 0.5 m). There are M mobile sinks in the field. For the randomly moving sinks, the velocity of each flight is uniformly chosen from $[0.5, 3]$ m/s. The pause duration is uniformly chosen from $[0, 5]$ min. For the controlled robots, the constant moving velocity is set to be 0.5 m/s. All the transceivers in sensors and sinks are assumed to be the same. The transmitting power is 10 mW at 900 MHz. The minimum received power for correct demodulation is -90 dBm. The antenna gains $g_t = g_r = 5$ dB. The antenna height of the AG mobile sinks is 1 m. The mean volumetric water content (VWC) in the soil is 5%. The monitoring error is represented by the normalized error, which is calculated by $E[\sigma_{norm}^2] = E(\sigma^2)/C(0)$. The normalized maximum tolerable error $(\sigma_{max}^2)_{norm}$ is set as 10%. Each simulation result is averaged over 500 iterations. We use the soil moisture as an example of the monitored physical quantity, since the spatio-temporal model of the soil

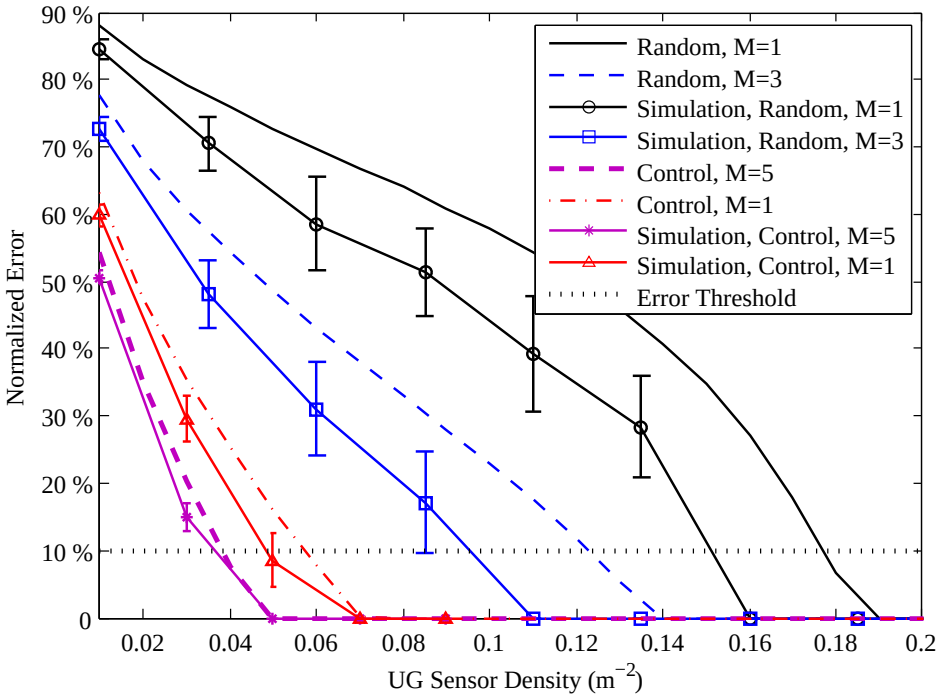


Figure 16: Normalized monitoring error in WUSNs as a function of the UG sensor density.

moisture has been well analyzed thoroughly. However, the soil moisture is much more spatio-temporally correlated than other physical quantities (e.g. vibrations). To reveal the general characteristics of the data collection scheme in most WUSN applications, we select the scenarios where the soil moisture is not so highly correlated. The related parameters in the spatio-temporally correlation model are set as follows: the normalized soil water loss $a = 20$ /day; the irrigation cell radius $1/r_R = 5$ m; and the irrigation duration $1/\eta = 60$ sec.

The theoretical bound of the monitoring error derived in Section 2.4.1 is first validated by simulations. In Fig. 16, the normalized monitoring error is given as a function of UG sensor density with different number and different mobility model of AG mobile sinks. It shows that the error bound is tight enough to serve as the optimization objective function under various system configurations. The optimal sensor density can be read from Fig. 16 by checking the x-coordinate of the intersection point of the error upper bound and the error threshold.

Next, we analyze the effects of multiple system and environmental parameters on the

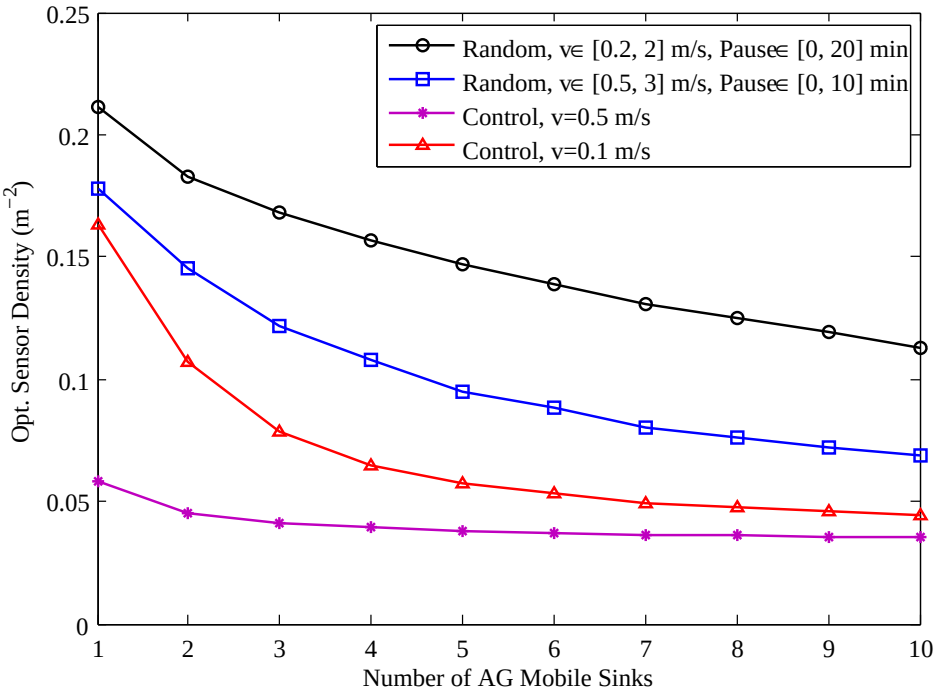


Figure 17: Optimal sensor density in WUSNs as a function of the number of mobile sinks.

optimal sensor density in WUSNs, including the number and mobility model of the mobile sinks, the mean burial depth of the UG sensors, the soil water content, and the spatio-temporal correlation model.

2.4.2.1 Number and Mobility Model of Mobile Sinks

The effect of the number and mobility of mobile sinks on the optimal sensor density is captured in Fig. 17. The optimal sensor density is given as a function of the number of mobile sinks with different mobility models. Both random and controlled mobility models with different velocities and pause times are considered. It indicates that the optimal sensor density can be significantly reduced by three ways: 1) introducing more mobile sinks, 2) increasing the sink velocity and reducing the pause time, and 3) employing controlled mobile sink instead of the randomly moving sinks. As shown in Fig. 17, thousands of UG sensors can be saved in a 100 m^2 region by the three ways. The reason behind this phenomenon is explained as follows. On the one hand, when the number of mobile sinks is

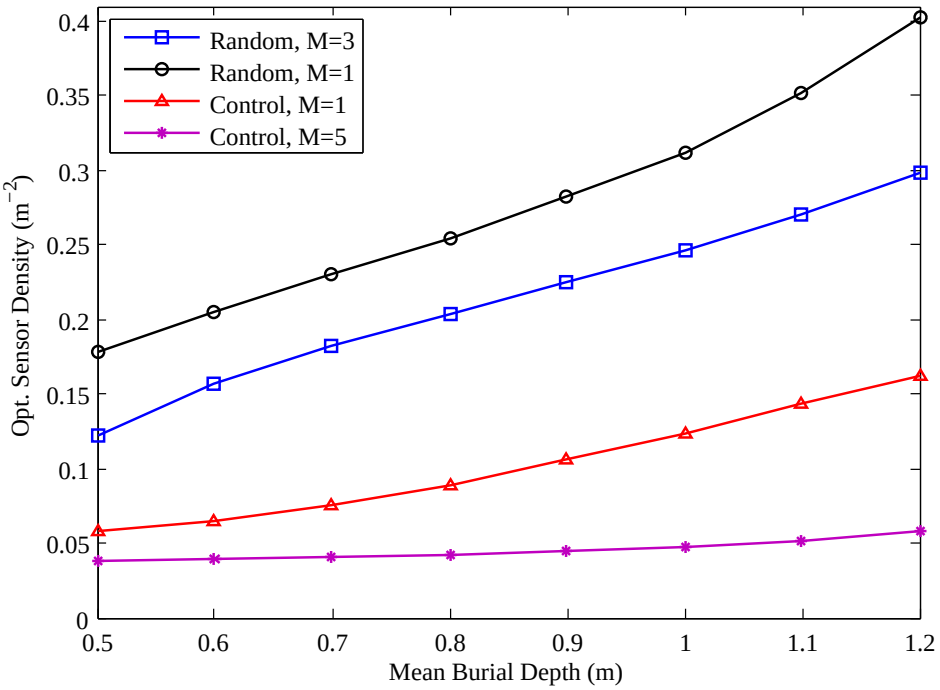


Figure 18: Optimal sensor density in WUSNs as a function of the mean burial depth.

small or their moving velocity is low, a large portion of the UG sensors in the WUSN need to be connected to guarantee the temporal sampling rate to achieve the required monitoring accuracy, which results in a high optimal sensor density. On the other hand, when the number of mobile sinks is large or their moving velocity is high, the mobile sinks can collect the data from each unconnected UG sensor cluster on time. Hence, the request on the UG sensor connectivity is lowered and the optimal sensor density is also reduced. Since the controlled sink mobility can enhance the data collection efficiency by guaranteeing the temporal sampling rate of every UG sensor cluster, the optimal sensor density can be further reduced by using the controlled mobile sinks. Note that when the moving velocity of the controlled mobile sink is high, the effect of the number of mobile sinks is not obvious since the data collection efficiency is already high enough.

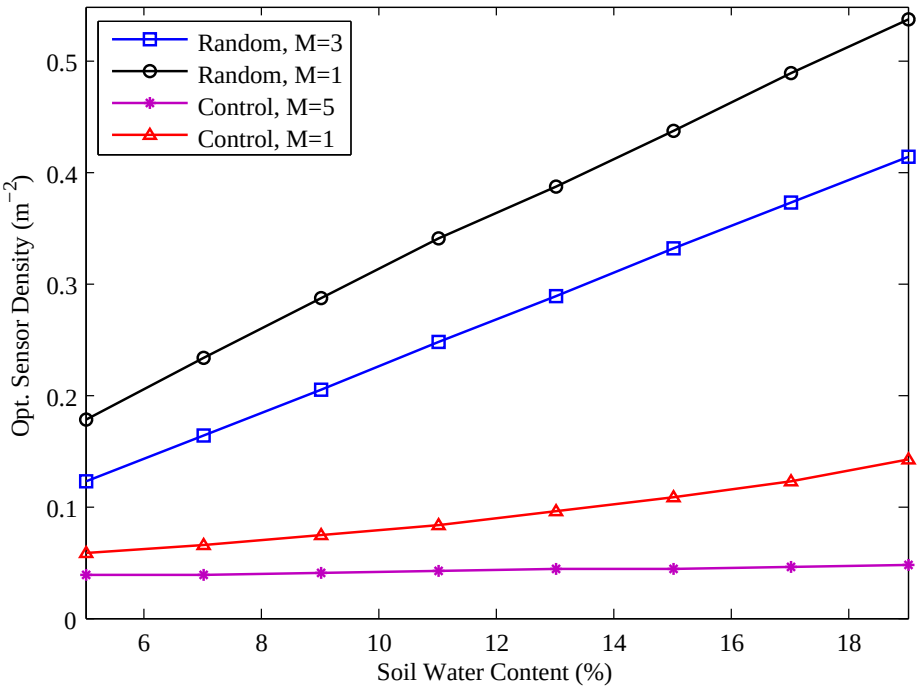


Figure 19: Optimal sensor density in WUSNs as a function of the volumetric soil water content.

2.4.2.2 Mean Burial Depth and Soil Water Content

The effects of the UG sensor burial depth and soil water content are analyzed in Fig. 18 and Fig. 19, respectively. In Fig. 18, the optimal sensor density is give as a function of the mean burial depth. When changing the mean burial depth, we assume that the span of the random depths remains the same, which is 0.4 m. In Fig. 19, the optimal sensor density is given as a function of mean volumetric soil water content. As discussed in Section 2.2, the communication ranges of the three types of channels in WUSNs significantly decrease as the UG sensor burial depth and the soil water content increase. Therefore, the optimal sensor densities of WUSNs dramatically increase as the mean burial depth or the soil water content increases, especially when the mobile sink moves randomly or the number of sinks is small.

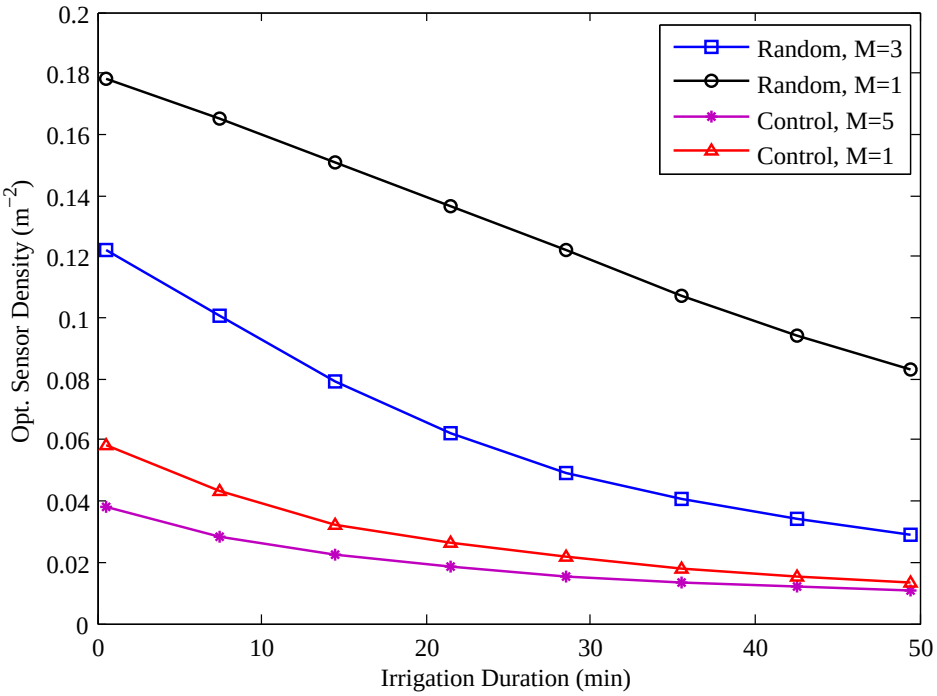


Figure 20: Optimal sensor density in WUSNs as a function of the irrigation duration.

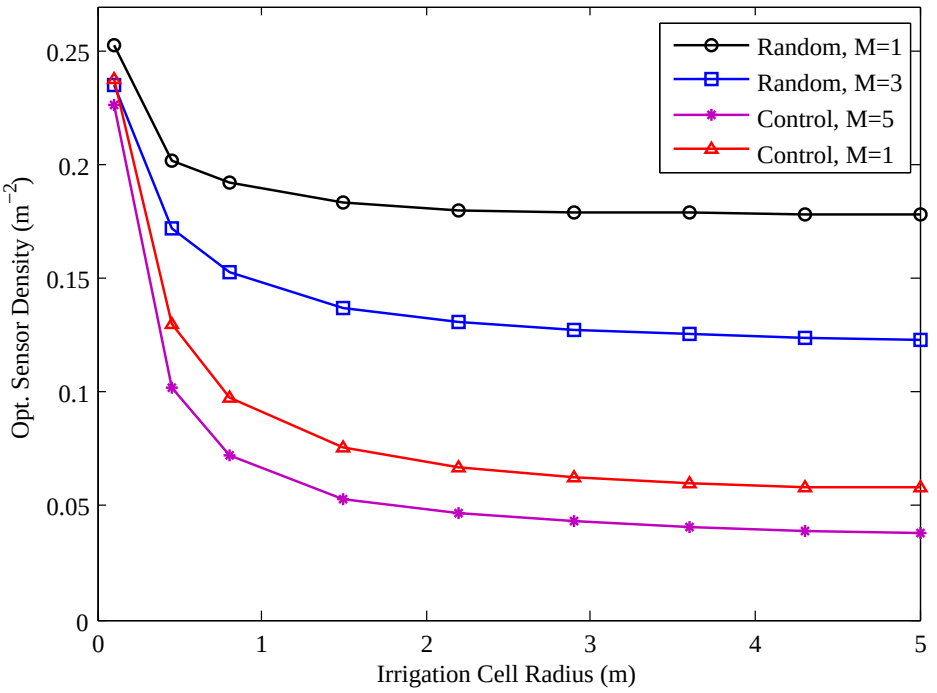


Figure 21: Optimal sensor density in WUSNs as a function of the irrigation cell radius.

2.4.2.3 Spatio-Temporal Correlation Model

In Fig. 20 and Fig. 21, we analyze the effects of the different parameters of the spatio-temporal model on the WUSN optimal sensor density. Note that we ignore the change of the soil water content due to the change of the spatio-temporal model parameters, since the spatio-temporal model does not affect the soil water content in most WUSN applications. In Fig. 20, we change the temporal correlation parameter, the mean irrigation duration $1/\eta$. As the irrigation duration increases, the temporal correlation between the monitored data also increases. The increased temporal correlation can help the WUSN estimate the unavailable data more accurately using the available data. As a result, the optimal sensor density decreases as the mean irrigation duration increases. In Fig. 21, we change the spatial correlation parameter, the mean irrigation cell radius $1/r_R$. As the irrigation cell radius increases, the spatial correlation between the monitored data also increases. Similar to the analysis on the temporal correlation, the optimal sensor density decreases as the mean irrigation cell radius increases.

CHAPTER III

MI-BASED WUSNS IN SOIL MEDIUM

3.1 Motivation and Related Work

Traditional wireless communication techniques using EM waves encounter two major problems in underground environments: the high path loss and the dynamic channel condition [4, 53]. In particular, first, EM waves experience high levels of attenuation due to absorption by soil, rock, and water in the underground. Second, the path loss is highly dependent on numerous soil properties such as water content, soil makeup (sand, silt, or clay) and density, and can change dramatically with time (e.g., increased soil water content after a rainfall) and location (soil properties change dramatically over short distances). The unreliable channel brings design challenges for the sensor devices and networks to achieve both satisfying connectivity and energy efficiency.

If the sensors of WUSNs are buried in the shallow depth, sensor can communicate with the aboveground data sinks directly using EM waves. This is because the underground path is short in this case. Hence the impacts of the additional path loss and the dynamic channel caused by the soil medium are much smaller. However, many WUSN applications, such as underground structure monitoring, require the sensors buried deep underground, where only underground-to-underground channel is available.

Magnetic induction (MI) is a promising alternative physical layer technique for WUSNs in deep burial depth. It can address the problems on the dynamic channel condition. Specifically, the underground medium such as soil and water cause little variation in the attenuation rate of magnetic fields from that of air, since the magnetic permeabilities of each of these materials are similar [2, 44, 80]. This fact guarantees that the MI channel conditions

remain constant for a certain path in different times. In addition, since the radiation resistance of coil is much smaller than electric dipole, very small portion of energy is radiated to the far field by the coil. Hence, the multi-path fading is not an issue for MI communication. However, MI is generally unfavorable for terrestrial wireless communication. As the transmission distance r increases, magnetic field strength falls off much faster ($1/r^3$) than the EM waves ($1/r$) in terrestrial environments. In underground environments, although it is known that the soil absorption causes high signal attenuation in the EM wave systems but does not affect the MI systems, it is not clear whether the total path loss of the MI system is lower than the EM wave system or not. Additionally, since the MI communication involves reactance coils as antenna, the system bandwidth needs to be analyzed.

The magnetic induction has been introduced as a new physical layer technique for wireless communication in recent years. In [80], MI communication is employed in the mine warfare (MIW) operations to provide a more reliable wireless command, control and navigation channel. The EM channel is qualitatively analyzed and the low data rates of 100 to 300 bit/s are achieved in various MI communication experiments carried out in coastal areas. The authors notice that the high path loss limits the transmission range. They suggest to place more MI transceivers to mitigate the high path loss, which is not feasible for underground wireless networks due to cost/energy constraint and deployment difficulty. In [67, 9, 18], the MI is utilized as an alternative personal communication technique to the Bluetooth. In the near-field communication applications (such as the link between a cell phone or an MP3 player and a headset), the rapid fall off of the MI signal strength is exploited to provide each user with his own private bubble, without having to worry about mutual interference among multiple users, and permitting bandwidth reuse. However, in the underground communication applications, the high path loss is obviously not an advantage.

In [44], the MI is first introduced to the field of wireless underground communication. It shows that the MI transmission is not affected by soil type, composition, compaction, or

moisture content, and requires less power and lower operating frequencies than RF transmission. However, the theoretical/experimental results show that the communication range is no more than 30 inches (0.76 m). Moreover, the bandwidth of the MI system is not considered in the paper. Besides underground, the MI communication can also be used in other RF-impenetrable environments, such as human body. In [83], a body network is built to collect data from, and transport information to, implanted miniature devices at multiple sites within the human body. The MI technique is employed to link information between a pair of implants, and to provide electric power to these implants. In [102], a new magnetic material is analyzed to guide magnetic information to the receiver coil, permitting a clear image deep within the body.

In this chapter, we first provide a detailed analysis on the path loss and the bandwidth of the MI communication channel in underground environments. Then based on the analysis, we develop the MI waveguide technique for WUSNs, which can significantly reduce the path loss, enlarge the transmission range and achieve practical bandwidth for MI communication in underground environments. In particular, the MI transmitter and receiver are modeled as the primary coil and secondary coil of a transformer. Multiple factors are considered in the analysis, including the soil properties, coil size, the number of turns in the coil loop, coil resistance, operating frequency. The analysis shows that the ordinary MI systems have larger transmission range but lower bandwidth than the EM wave systems. However, neither the ordinary MI system nor the EM wave system is able to provide enough communication range for practical WUSNs applications. Motivated by this fact, we develop the MI waveguide technique to enlarge the communication range. In this case, some small coils are deployed between the transmitter and the receiver as relay points, which form a discontinuous waveguide.

Up to now, the MI waveguide has been designed and used as artificial delay lines and filters, dielectric mirrors, distributed Bragg reflectors, slow-wave structures in microwave tubes, coupled cavities in accelerators, modulators, among others [93, 45, 94]. However,

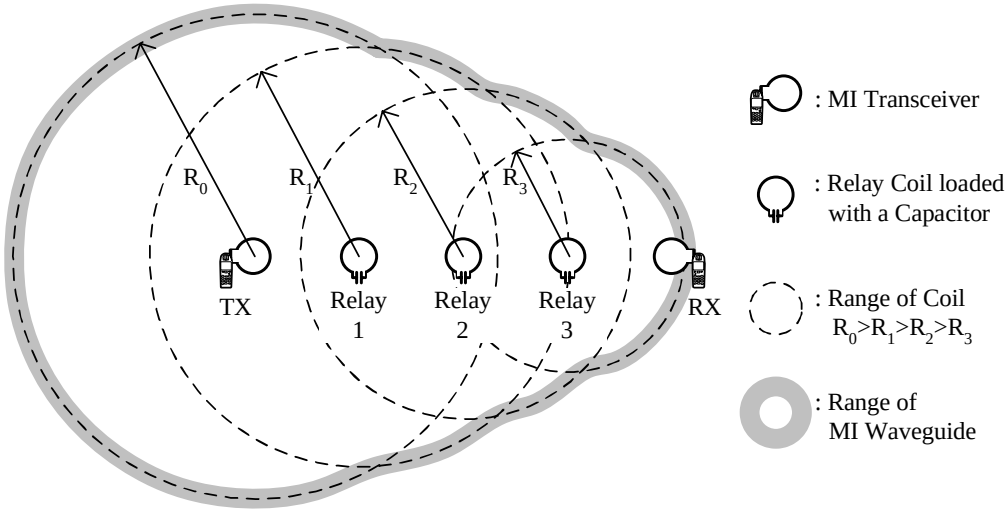


Figure 22: The structure and the communication range of a MI waveguide.

there is no attempt to utilize the MI waveguide in the wireless communication field. The theoretical analysis of the MI waveguide in [45] is validated by experiments in [103]. Note that we adopt similar theoretical analysis method as [45] in this paper.

The MI waveguide has three advantages in underground wireless communications: first, by appropriately designing the waveguide parameters, the total path loss can be greatly reduced. The maximum communication range between two transceivers can achieve several hundreds meters. Second, MI waveguide is not a continuous structure like traditional waveguide. It is only required to deploy one relay coil every 5 meters (or even longer) between the transceivers. Hence it is very flexible and easy to deploy and maintain. Third, the relay coils do not consume any energy and the cost is very small. The bandwidth of the MI waveguide systems is several KHz. Although it is much smaller than the EM wave system, it is enough for the low data rate monitoring applications of WUSNs.

Despite of the potential advantages, the deployment of the MI waveguides to connect the underground sensors is challenging due to the following reasons. First, on the one hand, a non-trivial number of relay coils are required to guarantee the network connectivity and robustness. On the other hand, the intensive deployment of the coils in underground soil medium cost a great amount of labor. Therefore the optimal number of relay coils needs

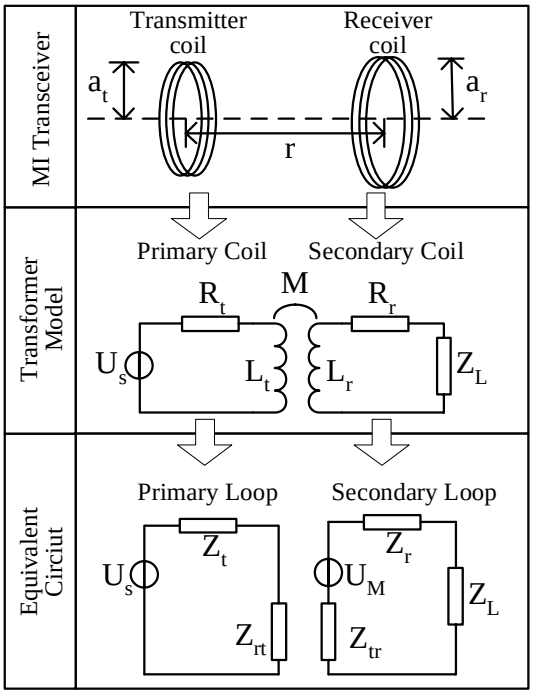


Figure 23: MI communication channel model

to be found out. Second, the communication range of the MI relay coil is not the same as each other, as shown in Fig. 22. Consequently, the shape of the communication range of the MI waveguide is much more complex than the disk communication range of the traditional wireless devices. Current sensor deployment strategies [101, 113] are based on the disk communication range, hence cannot be utilized to deploy the MI waveguides in the WUSNs. Therefore, at the end of this chapter, we analyze relay coil deployment strategies for the WUSNs using MI waveguides. The optimal deployment algorithms to use the MI relay coils to connect the underground sensors is developed.

3.2 Channel Modeling

In MI communications, the transmission and reception are accomplished with the use of a coil of wire, as shown in the first row in Fig. 23, where a_t and a_r are the radii of the transmission coil and receiving coil, respectively; r is the distance between the transmitter and the receiver.

Suppose the signal in the transmitter coil is a sinusoidal current, i.e., $I = I_0 \cdot e^{-j\omega t}$, where ω is the angle frequency of the transmitting signal. $\omega = 2\pi f$ and f is the system operating frequency. This current can induce another sinusoidal current in the receiver then accomplish the communication. The interaction between the two coupled coils is represented by the mutual induction. Therefore, the MI transmitter and receiver can be modeled as the primary coil and the secondary coil of a transformer, respectively, as shown in the second row in Fig. 23, where M is the mutual induction of the transmitter coil and receiver coil; U_s is the voltage of the transmitter's battery; L_t and L_r are the self inductions; R_t and R_r are the resistances of the coil; Z_L is the load impedance of the receiver. We use its equivalent circuit to analyze the transformer, as shown in the third row in Fig. 23, where,

$$\begin{aligned} Z_t &= R_t + j\omega L_t; \\ Z_{rt} &= \frac{\omega^2 M^2}{R_r + j\omega L_r + Z_L}; \\ Z_r &= R_r + j\omega L_r; \\ Z_{tr} &= \frac{\omega^2 M^2}{R_t + j\omega L_t}; \\ U_M &= -j\omega M \frac{U_s}{R_t + j\omega L_t}. \end{aligned} \tag{88}$$

where Z_t and Z_r are the self impedances of the transmitter coil and the receiver coil, respectively; Z_{rt} is the influence of the receiver on the transmitter while Z_{tr} is the influence of the transmitter on the receiver; U_M is the induced voltage on the receiver coil.

In the equivalent circuit, the transmitting power is equal to the power consumed in the primary loop. The receiving power is equal to the power consumed in the load impedance Z_L. Both received power and transmitting power are functions of the transmission range r:

$$\begin{aligned} P_r(r) &= Re\left\{\frac{Z_L \cdot U_M^2}{(Z_r' + Z_r + Z_L)^2}\right\} \\ P_t(r) &= Re\left\{\frac{U_s^2}{Z_t + Z_t'}\right\} \end{aligned} \tag{89}$$

According to the transmission line theory, the reflections take place unless the line is terminated by its matched impedance. In the equivalent circuit described in Fig. 23, to

maximize the received power, the load impedance is designed to be equal to the complex conjugate of the output impedance of the secondary loop, i.e.

$$Z_L = \overline{Z_r + Z_r'} \tag{90}$$

The following task is to find the analytical expression for the resistance, self and mutual induction of the transmitter and receiver coils. The resistance is determined by the material, the size and the number of turns of the coil:

$$\begin{aligned} R_t &= N_t \cdot 2\pi a_t \cdot R_0; \\ R_r &= N_r \cdot 2\pi a_r \cdot R_0 \end{aligned} \tag{91}$$

where, N_t and N_r are the number of turns of the transmitter coil and receiving coil, respectively; R_0 is the resistance of a unit length of the loop. According to American Wire Gauge (AWG) standard, R_0 can be a value from $2{\times}10^{-4}\ \Omega/m$ to $3\ \Omega/m$ with different wire diameter [6].

Since the coil is modeled as a magnetic dipole, the self induction and mutual induction can be deduced by the magnetic potential A of the magnetic dipole, which is provided in polar coordinate system by [36],

$$\mathbf{A}(r, \theta, \phi) = \frac{\mu}{4\pi r} \pi a_t^2 I_0 e^{-j\omega t} \sin\theta \left(\frac{1}{r} - j\frac{2\pi}{\lambda} \right) \cdot \hat{\mathbf{a}}_\phi \tag{92}$$

where μ is the permeability of the transmission medium; λ is the wavelength of the signal. By using Stokes' theorem [36], the mutual induction of the two coils can be calculated:

$$M = \frac{N_r \oint_{l_r} \mathbf{A} \cdot d\vec{\mathbf{l}}_r}{dI} \simeq \mu\pi N_t N_r \frac{a_t^2 a_r^2}{2r^3} \tag{93}$$

The self induction can be derived in the same way:

$$\begin{aligned} L_t &\simeq \frac{1}{2}\mu\pi N_t^2 a_t; \\ L_r &\simeq \frac{1}{2}\mu\pi N_r^2 a_r \end{aligned} \tag{94}$$

Consequently, by substituting (88), (90), (91), (93) and (94) into (89), the received power and the transmitting power can be calculated.

It should be noted that, the underground transmission medium contains different type of soil, water, rocks and etc. It is necessary to analyze the differences between the permeabilities of these materials. According to [97], the substances of the underground medium can be categorized into four main groups including organic materials, inorganic materials, air, and water, where organic materials come from plants and animals; inorganic materials include sand, silt and clay. The relative permeabilities of the plants, animals, air and water are very close to 1. If the sand, silt, and clay do not consist of magnetite, their permeabilities are also close to 1. An example is that the average value for sedimentary rocks is given in [97] as 1.0009. Since most soil in the nature does not contain magnetite, we can assume that the permeability of the underground transmission medium is a constant based on the above discussion.

3.2.1 Path Loss

For wireless communication using EM waves, the Friis transmission equation [49] gives the power received by one antenna, given another antenna some distance away transmitting a known amount of power. Since the radiation power is the major consumption of the EM wave transmitter, the transmitting power of the EM wave system is a constant and not influenced by the position of the receiver, i.e. for EM waves, P_r is a function of distance r while P_t is a constant. Hence the path loss is measured by the ratio of the received power to the radiation power. The path loss L_{EM} of the EM wave propagation in soil medium is given by [4, 53]:

$$\begin{aligned} L_{EM}(r) &= -10 \lg \frac{P_r(r)}{P_t} \\ &= 6.4 + 20 \lg r + 20 \lg \beta + 8.69 \alpha r \end{aligned} \tag{95}$$

where the transmission distance r is given in meters; the attenuation constant α is in $1/m$ and the phase shifting constant β is in $radian/m$. The values of α and β depend on the

dielectric properties of soil, and is derived in [53] using the Peplinski principle [68]. Note that the reflection from the air-ground interface is neglected since the burial depth is large, which has been explained in [4, 53].

Unlike the EM wave transmitter, the radiation power of the MI communication system can be neglected since the radiation resistance is very small. Meanwhile, the induced power consumed at the MI receiver is the major power consumption since the MI communication is achieved by coupling in the non-propagating near-field. The transmitting power of the MI system consists of the induced power consumed at the MI receiver and the power consumed in the coil resistance. If the coil resistance is small, the ratio of the received power to the transmitting power will be close to 1 since the receiving power and transmission power decrease simultaneously as the transmission distance increases. The advantage of this feature is that the limited transmission power won't be wasted on radiation to the surrounding space. Most power is transmitted to the receiver, which is favorable to the energy constrained WUSNs. However, as the transmission distance increases, less and less power is transmitted to the receiver. Hence there still exists a so called *Path loss*. It should be noted that the power is not really lost but in fact not transmitted. To fairly compare the performance of the EM wave system and MI system, the path loss of the MI system with transmission distance r is defined as $L_{MI}(r) = -10\lg\frac{P_r(r)}{P_t(r_0)}$, where $P_r(r)$ is the received power at the receiver that is r meters away from the transmitter; $P_t(r_0)$ is the reference transmitting power when the transmission distance is a very small value r_0. We can consider that $P_t(r_0) \simeq U_s^2/R_t$ if r_0 is small enough. In case of low coil resistances and high operating frequency ($R_0 << \omega\mu$), the path loss of the MI communication system can be simplified as

$$\begin{aligned} L_{MI}(r) &= -10\lg\frac{P_r(r)}{P_t(r_0)} \simeq -10\lg\frac{N_r a_t^3 a_r^3}{4N_t r^6} \\ &= 6.02 + 60\lg r + 10\lg\frac{N_t}{N_r a_t^3 a_r^3} \end{aligned} \tag{96}$$

We compare (95) with (96) to analyze the path loss of MI and EM wave systems in

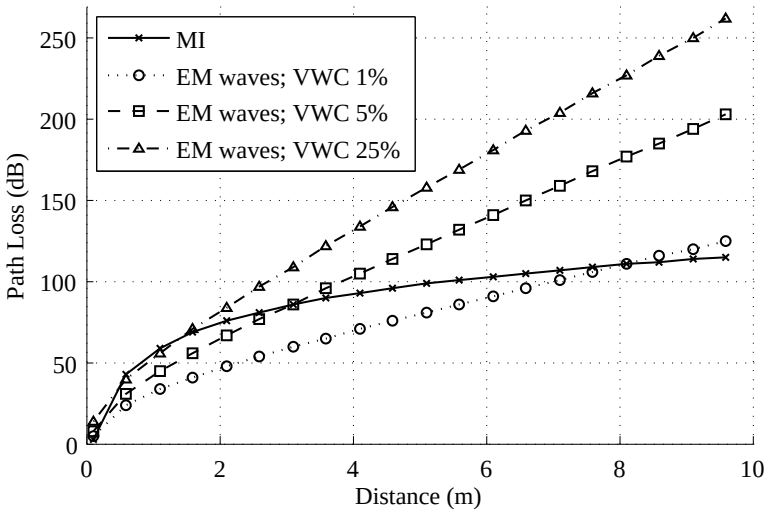


Figure 24: Path loss of the EM wave system and the MI system with different soil water content

underground environments. In (95), there are two terms in the path loss that are determined by the distance r, where the term $(20 \lg r)$ is due to the space spread and the term $(8.69\alpha r)$ is due to the material absorption. The transmission medium has significant influence in the path loss since it determines the propagation constants α and β. In (96), only one term $(60 \lg r)$ is determined by the distance r, which is due to the spread of the magnetic field. The transmission medium has no obvious influence on the MI path loss since we assume that the permeability of the medium is a constant as discussed in the beginning. Although the path loss term $(60 \lg r)$ in MI case is much higher than the term $(20 \lg r)$ in EM waves case, it is not clear whether the total path loss of MI system is larger than that of the EM wave system or not, since the material absorption term $(8.69\alpha r)$ in EM wave path loss varies a lot in different transmission medium.

3.2.2 Numerical Analysis

3.2.2.1 Path Loss

The path losses of the MI system and the EM wave system shown in (95) and (96) are evaluated using MATLAB. The results are shown in Fig. 24. According to [4, 53], the propagation of the EM waves in soil medium is severely affected by the soil properties,

especially, the volumetric water content (VWC) of soil. Hence in the evaluations, we set the VWC of the soil medium as 1%, 5% and 25%. The permittivity and conductivity of soil medium is calculated by the Peplinski principle [3, (8)-(12)], which are functions of VWC and soil composition. In our simulations, besides VWC, the soil composition is set as follows, the sand particle percent is 50%, the clay percent is 15%, the bulk density is 1.5 $grams/cm^3$, and the solid soil particle density is 2.66 $grams/cm^3$, which are typical values in nature. As discussed in the beginning, the permeability of the underground transmission medium is a constant and is the same as that in the air, which is $4\pi \times 10^{-7}$ H/m. Other simulation parameters are set as follows: for EM wave system, the operating frequency is set to 300 MHz. The reason for this choice is as follows: on the one hand, lower frequency bands are necessary for acceptable path loss. On the other hand, decreasing operating frequency below 300 MHz increases the antenna size, which can also prevent practical implementation of WUSNs. For MI system, the transmitter and the receiver coil have the same radius of 0.15 m and the number of turns is 5. The coil is made of copper wire with a 1.45 mm diameter. Hence the resistance of unit length R_0 can be looked up in AWG standard [6] as 0.01 Ω/m. The operating frequency is set to 10 MHz. This low operating frequency together with the small number of turns can effectively mitigate the influence of the parasitic capacitance [22].

In Fig. 24, the path losses of the MI system and EM wave system are shown in dB versus the transmission distance r with different soil VWC. As expected, the path loss of the MI system is not affected by the environment since the permeability μ remains the same. On the other hand, the path loss of the EM wave system dramatically increases as the VWC increases. When the soil is very dry (VWC=1%), the path loss of the EM waves is smaller than that of the MI system. When the soil is very wet (VWC=25%), the path loss of the EM waves is significant larger than that of the MI system. When VWC=5%, the path losses of these two systems are similar. It can be seen that the path loss of the MI system is a lg function of the distance r while the path loss of the EM wave system is an

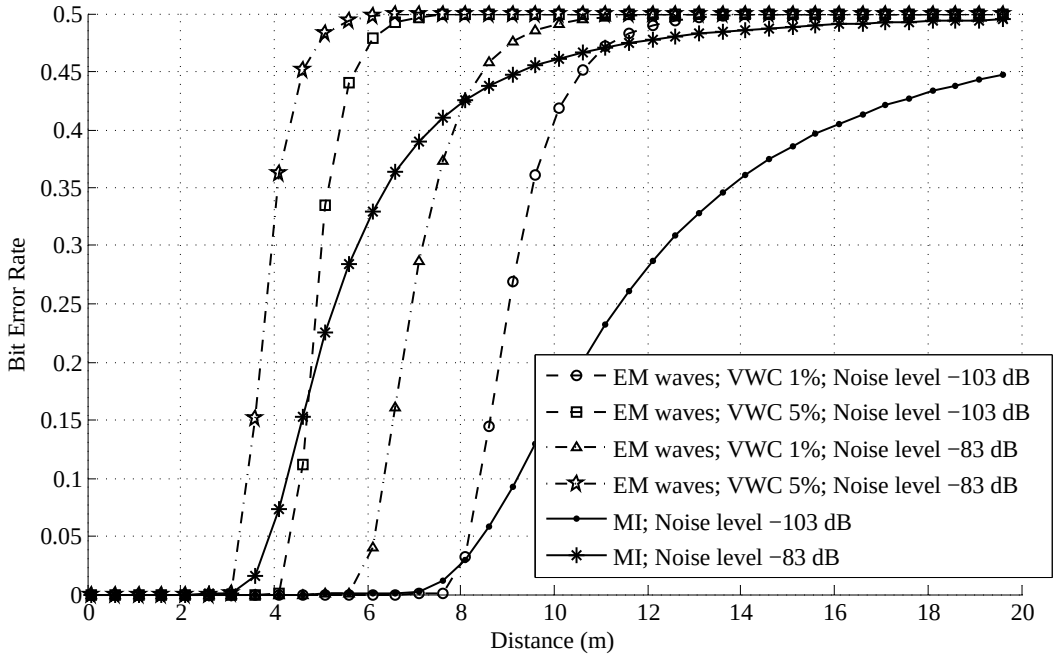


Figure 25: Bit error rate of the EM wave system and the MI system with different soil water content and noise level

approximately linear function of the distance r. This is because that the path loss caused by material absorption is the major part in the EM waves' propagation. When VWC=5%, in the near region between 0.5 m and 3 m, the EM wave system has smaller path loss; in the relatively far region ($r > 3$ m), the MI system has smaller path loss than the EM wave system. Even in the very dry soil medium (VWC=1%), the MI system can achieve smaller path loss than the EM wave system after a sufficient long transmission distance.

3.2.2.2 Bit Error Rate

Furthermore, we investigate the bit error rate (BER) characteristics of the two propagation techniques. The results are shown in Fig. 25. The BER characteristic depends mainly on three factors: 1). the path loss, 2). the noise level and 3). the modulation scheme used by the system. The path loss of the MI system and the EM wave system has been given in (95) and (96). The noise power in soil is measured using the BVS YellowJacket wireless spectrum analyzer [12] in [4, 53]. The average noise level P_n is found to be -103 dBm. Besides the experiment measurement, we also assume a high noise scenario where the

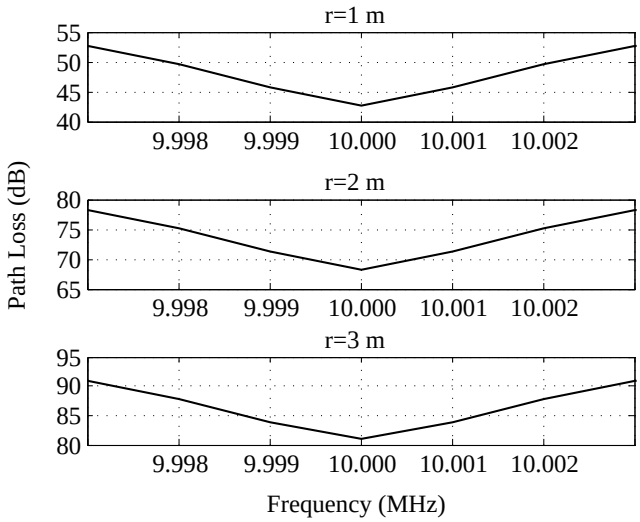


Figure 26: Frequency response of the MI system with different transmission range

average noise level P_n is set to be -83 dBm. Then the signal to noise ratio (SNR) can be calculated by $SNR = P_t - L - P_n$, where P_t is the transmitting power and L is the path loss given in (95) and (96). We set P_t as 10 dBm in the simulation. Considering the modulation scheme as the simple but widely used $2PSK$, the BER can be derived as a function of SNR: $BER = 0.5erfc(\sqrt{SNR})$, where $erfc(\cdot)$ is the error function [71].

In Fig. 25, the BERs of the MI system and EM wave system are shown as a function of the transmission distance r with different soil VWC. In low noise scenario, the transmission range of the MI system is larger than the EM wave system no matter what VWC is, which can be explained by the following reasons: 1) path loss below 100 dB cannot influence the BER performance when the noise is low. 2) The MI system has higher path loss than the EM wave system at the near region where the path losses of both systems are below 100 dB; while in the far region where the path losses are higher than 100 dB, the MI system has lower path loss. 3) It is the path loss in the far region that determines the transmission range. In the high noise scenario, the transmission range of MI system is between the range of EM wave system in dry soil and the system in wet soil, since this time the path loss above 80 dB can influence the BER performance.

3.2.2.3 Bandwidth

It should be noted that, the path loss of the MI system derived above is based on the assumption that the load impedance is designed to be equal to the complex conjugate of the output impedance of the secondary loop. However, since the output impedance of the secondary loop consists of not only resistance but also reactance, only one central frequency can realize this load matching. Any deviation from the central frequency will cause the power reflections and increase the path loss. Hence it is necessary to analyze the bandwidth of the MI system. In Fig. 26, the frequency response of the MI system described above is shown with different transmission distance. It indicates that the 3-dB bandwidth of the MI system is around 2 KHz when the operating frequency is 10 MHz. The bandwidth is not affected by the transmission distance. Although the 2 KHz bandwidth is much smaller than the EM wave system, it should be enough for the WUSNs considering that the underground sensing and monitoring applications do not require very high data rate [2].

To sum up, the MI system provides larger transmission range (around 10 m) than that of the EM wave system (around 4 m). The MI system also has the advantage that its performance is not influenced by the soil medium properties, especially the water content. Although the bandwidth of the MI system is smaller than that of the EM wave system, it should to a large extent fulfill the requirements of the WUSNs applications. However, the transmission ranges of both systems are still too short for a practical applications in underground medium.

3.3 MI Waveguide Technique

Although the ordinary MI system has constant channel condition and relatively longer transmission range than that of the EM wave system, its transmission range is still too short for practical applications. One solution is to employ some relay points between the transmitter and the receiver. Different from the relay points using the EM wave technique, the MI relay point is just a simple coil without any energy source or processing device.

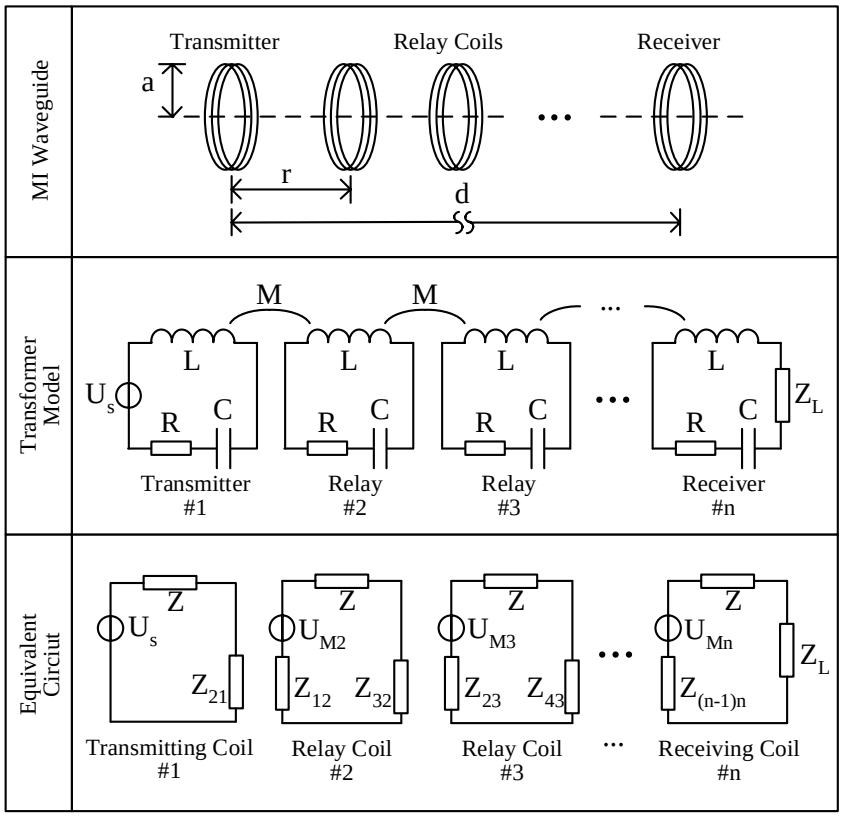


Figure 27: MI waveguide communication channel model

The sinusoidal current in the transmitter coil induces a sinusoidal current in the first relay point. This sinusoidal current in the relay coil then induces another sinusoidal current in the second relay point, and so on and so forth. Those relay coils form an MI waveguide in underground environments, which act as a waveguide that guides the so-called *MI waves*.

A typical MI waveguide structure is shown in the first row in Fig. 27, where $n-2$ relay coils equally spaced along one axis between the transmitter and the receiver, hence the total number of coils is n; r is the distance between the neighbor coils; d is the distance between the transmitter and the receiver and $d = (n-1)r$; a is the radius of the coils. Each relay coil (including the transmitter coil and the receiver coil) is loaded with a capacitor C. By appropriately designing the capacitor value, resonant coils can be formed to effectively transmit the magnetic signals. There exists mutual induction between any pair of the coils. The value of the mutual induction depends on how close the coils are to each other. In

underground communication, we set the distance between two relay coils to 5 m, which is larger than the maximum communication range of the EM wave system. Hence the MI waveguide system do not cost more on deploying the underground device than the traditional EM wave system. A lot of money can be saved by replacing the expensive relay sensor devices using EM waves by the relay coils that have very low cost. In the later part of this section, we vary the relay distance of the MI waveguide to analyze the influence. We assume that the radius of the relay coil is around 0.15 m. Comparing to the coil radius, the relay distance is large enough to validate the fact that the coils are sufficiently far from each other and only interact with the nearest neighbors. Hence, only the mutual induction between the adjacent coils needs to be taken into account in this thesis.

3.3.1 System Modeling

Similar to the strategy in section 3.2, the MI waveguide is modeled as a multi-stage transformer, where only adjacent coils are coupled, as shown in the second row in Fig. 27. Since in practical applications, the transceivers and the relay points usually use the same type of coils, we assume that all the coils have the same parameters (resistance, self and mutual inductions). M is the mutual induction between the adjacent coils; U_s is the voltage of the transmitter's battery; L is the coil self induction; R is the resistances of the coil; C is the capacitor loaded in each coil; Z_L is the load impedance of the receiver. The equivalent circuits of the multi-stage transformer is shown in the third row in Fig. 27, where

$$
\begin{aligned}
&Z = R + j\omega L + \frac{1}{j\omega C}; \\
&Z_{i(i-1)} = \frac{\omega^2 M^2}{Z + Z_{(i+1)i}},\ (i = 2, 3, ...n\text{–}1 \text{ and } Z_{n(n\text{–}1)} = Z_L); \\
&Z_{(i-1)i} = \frac{\omega^2 M^2}{Z + Z_{(i-2)(i-1)}},\ (i = 3, 4, ...n \text{ and } Z_{12} = \frac{\omega^2 M^2}{Z}); \\
&U_{Mi} = -j\omega M \frac{U_{M(i-1)}}{Z + Z_{(i-2)(i-1)}},\ (i = 2, 3, ...n \text{ and } U_{M1} = U_s).
\end{aligned}
\tag{97}
$$

where $Z_{i(i-1)}$ is the influence of the i^{th} coil on the $(i-1)^{th}$ coil and vice versa; U_{Mi} is the induced voltage on the i^{th} coil. Then the received power at the receiver can be calculated

as:

$$P_r = Re\left\{\frac{Z_L \cdot U_{Mn}^2}{(Z_{(n-1)n} + Z + Z_L)^2}\right\} \tag{98}$$

3.3.2 System Optimization

To maximize the received power is equal to maximize the induced voltage U_{Mn} at the receiver coil. According to (97), if the coils are resonant, then the impedance of each coil consists of only resistance and the absolute value becomes much smaller. Hence we design the capacitor to fufill $j\omega L + \frac{1}{j\omega C} = 0$, then using the expression of the self induction in (94), the value of the capacitors should be:

$$C = \frac{2}{\omega^2 N^2 \mu \pi a} \tag{99}$$

In case that the coils are resonant, the expression of the received power U_{Mn} in (97) can be developed as:

$$\begin{aligned} U_{Mn} &= U_s \cdot \frac{-j\omega M}{R} \cdot \frac{-j\omega M}{R + \frac{\omega^2 M^2}{R}} \cdot \frac{-j\omega M}{R + \frac{\omega^2 M^2}{R + \frac{\omega^2 M^2}{R}}} \cdots \frac{-j\omega M}{R + \frac{\omega^2 M^2}{R + \frac{\omega^2 M^2}{R + \cdots}}} \\ &= U_s \cdot (j)^{n-1} \cdot \frac{1}{x_1} \cdot \frac{1}{x_2} \cdots \frac{1}{x_{n-1}} \end{aligned} \tag{100}$$

where

$$x_i = \frac{R}{\omega M} + \frac{1}{x_{i-1}}, \quad (i = 2, 3, ...n-1 \quad \text{and} \quad x_1 = \frac{R}{\omega M})$$

Basing on the above equations, it can be shown that the multiplication $x_1 \cdot x_2 \cdot x_3 \cdots x_{n-1}$ is in fact an $(n-1)$ order polynomial of $x_1 = \frac{R}{\omega M}$, which is denoted as $\zeta(\frac{R}{\omega M},\ n-1)$ and:

$$\begin{aligned} \zeta(\tfrac{R}{\omega M},\ n-1) = b_{n-1}(\tfrac{R}{\omega M})^{n-1} + b_{n-2}(\tfrac{R}{\omega M})^{n-2} + \\ \cdots + b_2(\tfrac{R}{\omega M})^2 + b_1(\tfrac{R}{\omega M}) + b_0 \end{aligned} \tag{101}$$

where $\{b_i, i = 0, 1, 2, ..., n-1\}$ are the coefficients of the polynomial, which is fixed for a certain n and not affected by other parameters.

Since the coils are all resonant, the matched load impedance is pure resistance, which is $Z_L = \overline{Z_{(n-1)n} + R}$. Finally, in the MI waveguide system, if the receiver is d m away from the transmitter and there are $n-2$ relay coils between them, the received power can be expressed as:

$$P_r(d) = \frac{1}{4(Z_{(n-1)n} + R)} \cdot \frac{U_s^2}{\zeta^2(\frac{R}{\omega M},\ n-1)} \tag{102}$$

where d is the total transmission range and $d = (n-1)r$.

The same as the ordinary MI system, the transmission power and the receiving power of the MI waveguide system decrease simultaneously as the transmission distance increases. Hence, the path loss of the MI waveguide L_{MIG} is defined in the same way:

$$\begin{aligned} L_{MIG}(d) &= -10\lg\frac{P_r(d)}{P_t(r_0)} \\ &\simeq 10\lg\frac{4(Z_{(n-1)n}+R)}{R} + 20\lg\zeta(\tfrac{R}{\omega M},\ n-1) \\ &= 10\lg 4\left[1 + \frac{1}{(\frac{R}{\omega M})^2 + \frac{1}{1+\frac{1}{(\frac{R}{\omega M})^2+\frac{1}{1+\cdots}}}}\right] + 20\lg\left[b_{n-1}(\tfrac{R}{\omega M})^{n-1} + \cdots + b_1(\tfrac{R}{\omega M}) + b_0\right] \end{aligned} \tag{103}$$

where $P_t(r_0)$ is defined as the transmission power when the transmitter is very close to the receiver and no relay coil exists.

According to (103), the path loss of the MI waveguide system is actually a function of $\frac{R}{\omega M}$. It is the polynomial $\zeta(\frac{R}{\omega M},\ n-1)$ that has the major influence on the path loss. Therefore the path loss is a monotone increasing function of the variable $\frac{R}{\omega M}$. Consequently, to minimize the path loss is equal to minimize the variable $\frac{R}{\omega M}$. By using the expressions of the wire resistance R and the mutual induction M in (91) and (93) respectively, the variable $\frac{R}{\omega M}$ can be expressed as:

$$\frac{R}{\omega M} = \frac{4R_0}{\omega N\mu\pi} \cdot \left(\frac{r}{a}\right)^3 \tag{104}$$

Note that here the relay distance r is only $\frac{1}{n-1}$ of the total transmission range d. By this means the influence of the cubic function of the distance on the path loss can be significantly mitigated. Using this scheme, we can reduce the path loss by:

- Reduce the ratio of the relay distance to the coil radius $\frac{r}{a}$;
- Increase the operating frequency ω and the number of turns of the coils N;
- Reduce the wire resistance R_0.

However, there are other factors that constrain the path loss minimization:

- To ease the device deployment, the ratio of the relay distance to the coil radius is expected to be as large as possible, which conflicts with the requirements of the low total path loss. In our work, to keep the incontrovertible advantage over the underground EM wave system, the relay distance is set to at least the maximum transmission range of EM wave system, which is 4 m. Considering the coil radius is 0.15 m, the ratio of the relay distance to the coil radius is over 27.
- It is also impossible to unlimitedly increase the operating frequency and the number of turns of the coils, since these two parameters are constraint by (99). The loaded capacitors in each resonant coil should be larger than 10 pF, otherwise it is comparable to the coil parasitic capacitance. To achieve a practical value of the loaded capacitors in each resonant coil, the ω and N cannot be too large. Moreover, extreme high operating frequency and large number of turns may induce severe performance deterioration caused by the parasitic capacitance [22]. In our work, we use 10 MHz operating frequency and the each coil contains 5 loops of wire. The loaded capacitor is around 35 pF in this case.
- Although reducing the wire resistance can reduce the total path loss, it may cause two problems: 1) lower wire resistance require larger wire diameter, which cost more and cause the coils heavier; 2) low wire resistance can also cause dramatical in-band signal fluctuation, which may create difficulties on equalization of the received signal. In our work, the coil is made of copper wire with a 1.45 mm diameter. According

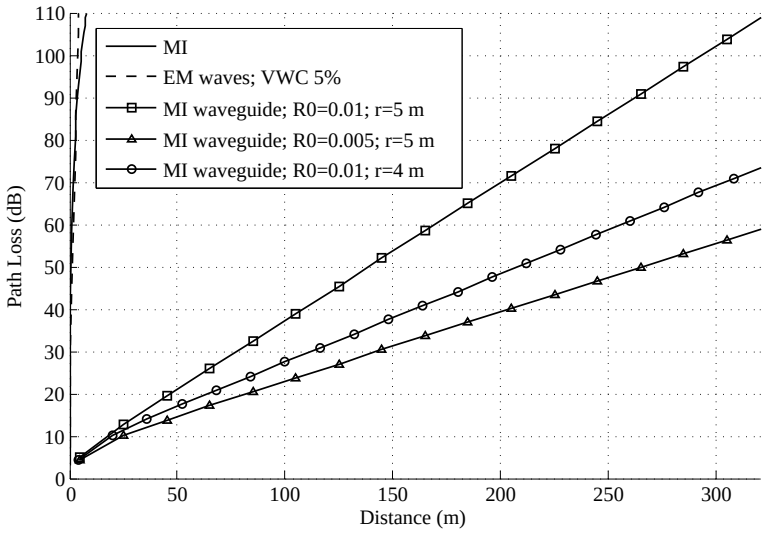


Figure 28: Path loss of the the MI waveguide system with different wire resistance and relay distance

to AWG standard [6], the resistance of unit length R_0 is 0.01 Ω/m. The influence of different wire resistances will be analyzed in the later part of this section.

3.3.3 Numerical Analysis

3.3.3.1 Path Loss

The path losses of MI waveguide system shown in (103) are evaluated using MATLAB. The results are shown in Fig. 28. For better comparison, the path loss of the 300 MHz EM wave system in 5% VWC soil and the path loss of the 10 MHz ordinary MI system are also plotted. According to the previous discussion, the performance of the MI system is not affected by the soil properties and the soil medium has the same permeability as that in the air, which is $4\pi\times10^{-7}$ H/m. Hence in the evaluation of the MI waveguide, we do not need to consider the environment parameters. Except studying the effects of certain parameters, the default values are set as follows: all the coils including the transmitter, receiver and relay points have the same radius of $a = 0.15$ m and the number of turns is $N = 5$. The resistance of unit length is $R_0 = 0.01\ \Omega/m$ for normal coil and $R_0 = 0.005\ \Omega/m$ for low resistance coil. The operating frequency is set to 10 MHz. The relay distance r is 5 m. The total number

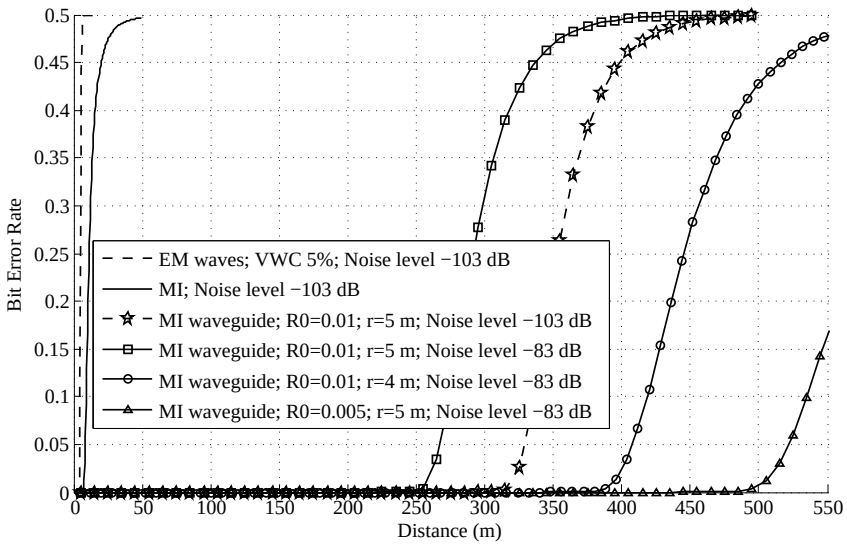


Figure 29: Bit error rate of the the MI waveguide system with different wire resistance, relay distance and noise level

of coils n is determined by the transmission distance d, where $d = (n-1)r$. In Fig. 28, the path losses of the MI waveguide system are shown in dB versus the transmission distance d with different relay distances r and different wire resistances R_0. It can be found that the MI waveguide can greatly reduce the signal path loss comparing with the EM wave system and the ordinary MI system. The path loss of the MI waveguide is less than 100 dB even after 250 m transmission distance, while the path loss of the EM wave system and the ordinary MI system becomes larger than 100 dB when the transmission distance is larger than 5 m. In addition, the path loss can be further reduced by reducing the relay distance and the wire resistance.

3.3.3.2 Bit Error Rate

In Fig. 29, we investigate the bit error rate (BER) characteristics of the MI waveguide. The same as the analysis in section 3.2, $2PSK$ is selected as the modulation scheme. Two noise level are considered, where the average noise level P_n in low noise scenario is -103 dBm while P_n in high noise scenario is -83 dBm. The transmission power P_t is set to 10 dBm.

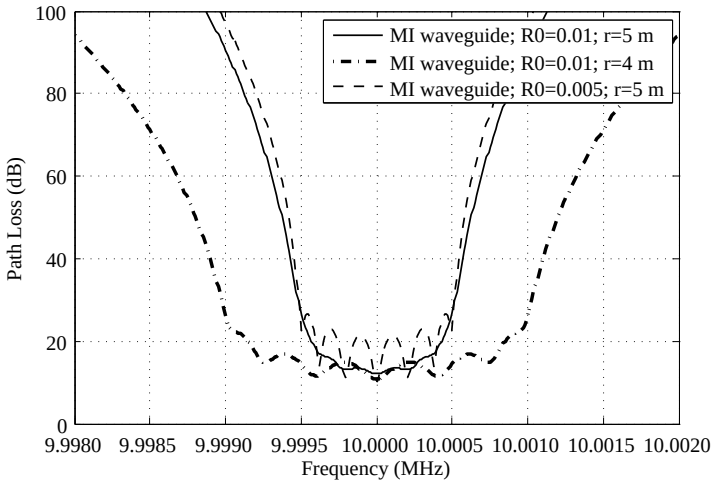


Figure 30: Frequency response of the the MI waveguide system with different wire resistance and relay distance

In Fig. 29, the BER of the MI waveguide system are shown as a function of the transmission distance d with different relay distances r and different wire resistances R_0. The BER of the EM wave system and the ordinary MI system are also plotted for comparison. Comparing with the small transmission range of the other two techniques (less than 10 m), the transmission range of the waveguide system is above 250 meters even in the high noise scenario. It means that the transmission range of the MI waveguide system is increased for more than 25 times compared with the other two systems. In accord with the analysis on the path loss, the transmission range of the MI waveguide can be extended by reducing the relay distance and the wire resistance.

3.3.3.3 Bandwidth

The above path loss and the transmission range of the MI waveguide system is calculated under the assumption that the transmitted signal has only one frequency. Under this central frequency, all the coils can achieve the resonant status. However, if there is any deviation from the central frequency, the resonant status of each coil will disappear and the load at the receiver also becomes unmatched with the system. Hence we need to analyze the bandwidth of the MI waveguide system. In Fig. 30, the frequency response of the MI

waveguide system is shown with different relay distances r and different wire resistances R_0. The number of relay coils n are fixed to 7. The results indicate that, when the operating frequency is 10 MHz, the 3-dB bandwidth of the MI waveguide system is in the same range with the ordinary MI system, which is 1 KHz to 2 KHz. Although lower wire resistance can reduce the path loss in the central frequency, the fluctuation of the in-band frequency response becomes so serious that may cause difficulties in the equalization at the receiver. The bandwidth can be enlarged by reducing the relay distance. However, for a certain transmission range, reducing the relay distance means that more relay coils needs to be deployed hence more effort is cost in the deployment. Two practical parameter sets maybe: 1) the relay distance $r = 5$ m and the unit length resistance $R_0 = 0.01\ \Omega/m$. In this case, the 10 MHz operating MI waveguide system can accomplish the communications within 250 m range and achieve 1 KHz bandwidth. And 2) the relay distance $r = 4$ m and the unit length resistance $R_0 = 0.01\ \Omega/m$. In this case, the 10 MHz operating MI waveguide system has 400 m transmission range and 2 KHz bandwidth.

3.3.3.4 Influence of Position Deviation

It should be noted that the above performance of MI waveguide system is derived in the ideal deployment case, where all the relay coils are accurately deployed so that the $n-2$ relay coils are uniformly distributed between the transceivers. The transmission range is divided into $n-1$ exactly equal intervals hence the mutual inductions between each relay coil are the same. However, in the practical applications, this requirements may not be precisely satisfied due to the following two reasons: on the one hand, in the initial deployment stage, the relay coils can not be set in the exact position as planned because of deployment constraints, such as rocks or pipes in the soil; on the other hand, the positions of the coils may change while the network is operating due to the above ground pressure or the movement of the soil. Hence, in Fig 31 and Fig 32, the influence of the non-ideal deployment is analyzed.

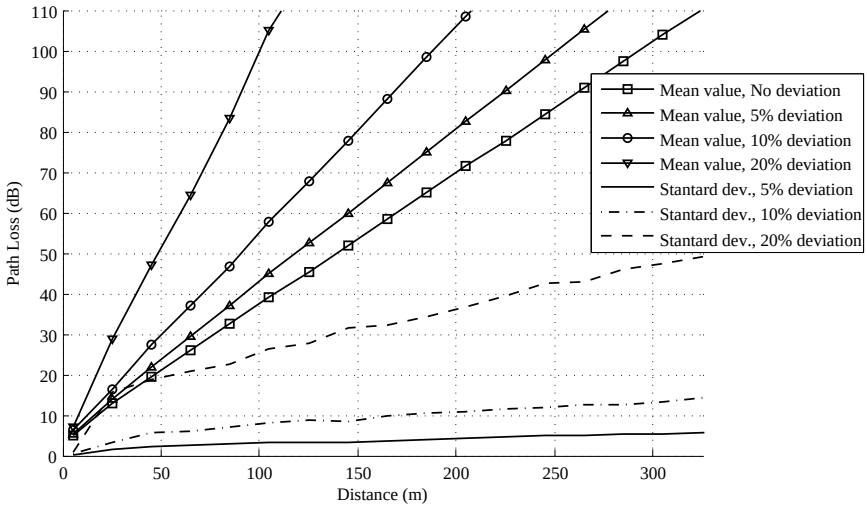


Figure 31: Path loss of the the MI waveguide system with different deviation from the designed relay distance

We assume that the relay coils are not deployed at the exact planed positions but may not deviate a lot. There are $n-2$ relay coils deployed between the transceivers and the transmission. Their designed positions are $\{i \cdot d/(n-1), i = 1, 2, \cdots n-2\}$. The position x_i of relay coil i is a gaussian random variable with mean value $i \cdot d/(n-1)$ and standard deviation σ_r. Then the transmission distance d is divided into $n-1$ intervals with length: $r_1,\ r_2,\ \ldots\ r_{n-1}$, where $r_i = x_i - x_{i-1}$. x_0 and x_{n-1} are the positions of the transmitter and the receiver, respectively. We assume that the standard deviation are either 5%, 10% or 20% of the designed relay distance. Other simulation parameters are set to the default value. The results are the average of 100 iterations. Both mean value and the standard deviation of the results are plotted.

It is shown that there exists additional path loss in practical deployment. Moreover, the bandwidth decreases dramatically when the standard deviation is 20%. The level of the additional path loss and the bandwidth decrease are determined by the standard deviation. Higher standard deviation can cause larger performance deterioration. Moreover, the additional path loss also increases as the transmission distance increase, which is because

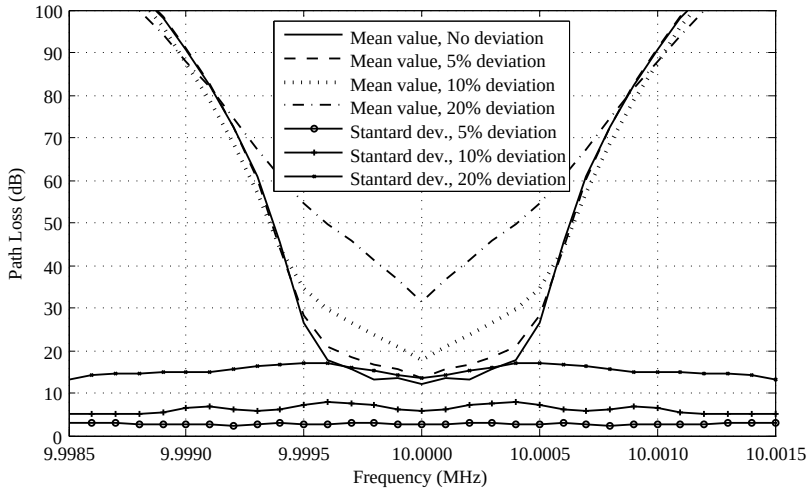


Figure 32: Frequency response of the the MI waveguide system with different deviation from the designed relay distance

that more relay coils are deployed with longer transmission distance hence more deployment deviation may occur. The standard deviation of the path loss and the bandwidth also increases dramatically as the deployment deviation increases, which indicates that the reliability of the MI waveguide system also decreases if deployment deviation occurs. It should be noted that the influence of the deployment deviation on the performance of the MI waveguide system can be neglected if the standard deviation is less than 10%.

3.4 Optimal Deployment

So far we derive the MI waveguide technique to connect two underground sensors. In this section, we analyze the deployment strategy to use MI waveguides to construct a connected and reliable underground network with low cost. In particular, we first consider the one-dimensional (1D) WUSNs. The optimal number of relay coils between two sensors are analyzed according to the required bandwidth and the distance between two sensors. Then based on the analysis of the 1D WUSNs, the optimal MI waveguide deployment strategy is developed for the two-dimensional (2D) WUSNs. Two coil deployment algorithms, the *MST* algorithm and the *TC* algorithm are proposed. To minimize the number of relay coils, we provide the MST algorithm, where the MI waveguides are deployed along the minimum

spanning tree of the WUSN. The weight of each link of the network is the optimal relay coil number. Since the WUSN constructed by MST algorithm is not robust to sensor failure, we propose the TC algorithm. In the TC algorithm, the MI waveguides are deployed around the centroids of the triangle cells that are constructed by the Voronoi diagram [11]. The WUSN constructed by the TC algorithm is robust to sensor failure but requires more relay coils.

3.4.1 MI Waveguide Deployment in 1D WUSNs

In this section, the deployment of the MI waveguides in a 1D WUSN is analyzed. The underground sensors are buried along a line or a polygonal line. This 1D network topology is applicable in the underground pipeline monitoring system. Moreover, the analysis results lay the foundation of the MI waveguide deployment strategy in 2D WUSNs.

The 1D WUSN can be divided into multiple links that starts at one sensor and ends at the next sensor. The goal of the optimal deployment of the MI waveguide in 1D WUSNs is to use as few relay coils as possible to connect the two sensors in each link. The optimal number of relay coils for each link is determined by the length of the link and the required bandwidth. We assume that the length of each link and the bandwidth have been determined by the requirements of the specific applications.

To minimize the deployment cost while maintaining the proper network functionality of the WUSNs, a MI waveguide should use the minimum number of relay coils to connect the two sensors on the link. According to (103), the path loss increases monotonically when the signal frequency deviates from the central frequency ω_0. Therefore, if the signal with the frequency $\omega = \omega_0 + 0.5B$ can be correctly received, a communication channel with bandwidth of B can be established between the two sensors. Assuming that transmission power is P_t and the minimum power for a sensor to correctly receive a signal is P_{th}. Using the path loss given in (103), the received power can be calculated. Then the optimal number

of relay coils for this link is:

$$n_{opt}(d, B) = \arg\min_{n}\{P_t - L_{MI}(d, n, \omega_0 + 0.5B) \geq P_{th}\}. \quad (105)$$

According to (105), the optimal number of relay coils is the function of the link length and the required bandwidth. Since the required bandwidth can be viewed as a constant, it is the link length that determines the optimal number of relay coil.

By using the parameters of the MI waveguide developed in [89], we can numerically analyze the optimal number of relay coils with different link length. In the following analysis, the transmission power is set to be 2.5 mW (4 dBm). The threshold of the power for correctly reception is set to be -80 dBm. Due to the resonant characteristics of the MI waveguide, the bandwidth of the system is much smaller than the terrestrial wireless networks. However, the small bandwidth is acceptable for WUSNs since the underground sensing and monitoring applications do not require very high data rate [2]. Therefore, the system bandwidth of the MI waveguide is set to be 1 KHz. The operating frequency is set to 10 MHz. The relay coils have the same radius of 0.15 m and the number of turns is 20. The coil is made of copper wire with a 1.45 mm diameter. The cost and weight of coils made of this kind of wire is neglectable. The wire resistance of unit length can be looked up in AWG standard [6] as 0.01 Ω/m. This relatively high wire resistance also effectively mitigates the in-band signal fluctuation. The permeability of the underground soil medium is a constant and is similar to the permeability of the air, since most soil in the nature does not contain magnetite. Therefore, $\mu = 4\pi \times 10^{-7}$ H/m. The soil moisture and the soil composition do not affect the MI communication as discussed perviously.

In Fig. 33, the received power of the 10 $MHz + 0.5$ KHz signal using MI waveguides with different relay coil numbers is shown as the function of the link length d. The axial communication range of a MI waveguide with a certain relay coil number is shown as the intersection point of the received power and the -80 dBm threshold. Fig. 33 shows that the axial communication range increases as the relay coil number increases. However, the increment of the communication range caused by additional relay coils decreases as

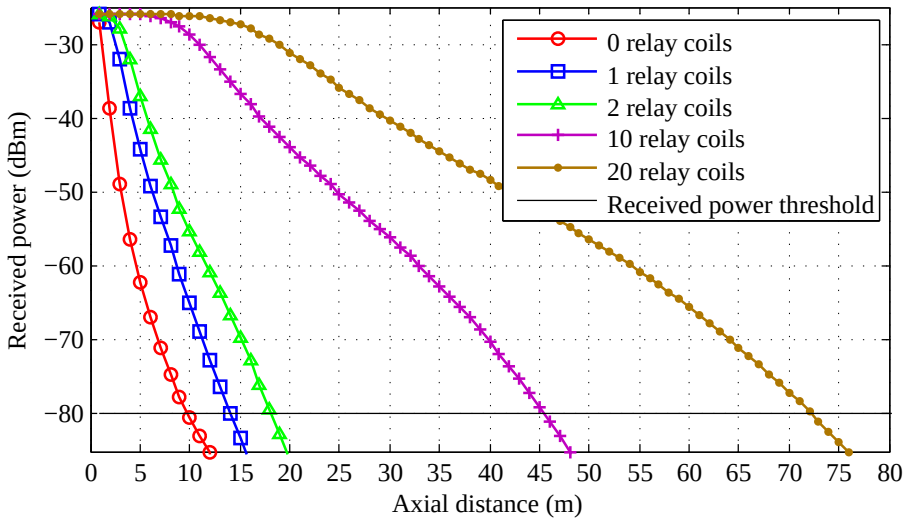


Figure 33: Received power of a 10 MHz + 0.5 KHz signal using MI waveguides with different relay coil numbers.

the relay coil number increases. For example, the axial communication range of a MI transceiver pair can be increased by 36 meters by adding the first 10 relay coils but can be only increased by 27 meters by adding another 10 relay coils. This phenomenon is due to the fact that the coils relay the signal in a passive way and there is no extra power added at each relay coil.

According to (105), the optimal relay coil number for the link with a certain length can be read from Fig. 33 by finding out the curve with the minimum relay coil number that has the axial communication range larger than the link length. We summarize the optimal number of relay coils and the corresponding link length in Table. 1. It shows that the optimal number of the relay coils increases faster than the link length increases. Consequently, the required interval between two adjacent coils decreases as the link length increases.

3.4.2 MI Waveguide Deployment in 2D WUSNs

In most WUSN applications, the network has a 2D topology. In this section, we investigate the deployment strategies of the MI waveguides to connect the underground sensors in a 2D WUSN. Compared with the MI waveguide deployment in 1D WUSNs, the deployment

Table 1: Optimal Number of Relay Coils and Corresponding Link Length

Link Length (m)	**Optimal Number of Relay Coils**	**Coil Interval (m)**
(0, 10]	0	10
(10, 14.5]	1	7.3
(14.5, 18.5]	2	6.2
(18.5, 22.5]	3	5.6
(22.5, 26]	4	5.2
(26, 29.5]	5	4.9
⋮	⋮	⋮
(43, 46]	10	4.2
⋮	⋮	⋮
(70, 73]	20	3.5
⋮	⋮	⋮

in 2D WUSNs is much more complicated due to the following reason: 1) in 1D WUSNs, the route connecting the sensor nodes are determined, while in 2D WUSNs, the optimal route to connect all the sensors needs to be found out; and 2) it is possible in a 2D WUSN that some common relay coils can be shared by multiple links.

Note that the MI waveguide deployment is also influenced by the topology of the sensors in the WUSNs. The topology of the sensors is determined by specific applications. If full sensor coverage is required in a sensing area where underground sensors can be buried at any desired positions, the hexagonal tessellation topology is preferred due to its efficiency and simplicity. If only some specific positions need to be monitored by sensors or some positions in the sensing area are not suitable to bury underground sensors, the WUSN has a random topology. In the hexagonal tessellation topology, the underground sensors of the WUSN are set in all vertexes of a hexagonal tessellation. The length of each tessellation edge is determined by specific applications. In the random topology, the positions of the sensors can be viewed as random distributed. Therefore, the hexagonal tessellation topology can be viewed as a special case of the random topology. In this section, we start the analyze of the MI waveguide deployment in WUSNs with the hexagonal tessellation

topology. Then we extend our research to the deployment strategy in WUSNs with random topologies.

3.4.2.1 Deployment in WUSNs with Hexagonal Tessellation Topology

Hexagonal tessellations have been widely used for the wireless network topologies, such as the base station placement of the cellular networks [104]. Due to the disk shape of the sensing range of the sensor devices, using hexagonal tessellation topology is the most efficient way to cover the whole sensing area. Different from the terrestrial wireless sensor networks, the communication range of the underground sensors is very limited. Hence, the MI waveguides are used to connect the sensors on the vertexes of the hexagonal tessellation. In the following analysis, we assume that the sensor density of the WUSN with the hexagonal tessellation topology is λ_{hex} (m^{-2}).

Minimum Spanning Tree (MST) Algorithm:

If the network robustness is not considered, the optimal deployment goal is to connect all the sensors in a WUSN with minimum number of relay coils. Therefore, the minimum spanning tree [105] can be used to find the optimal routes of MI waveguides. If the sensor number is K, the number of edges of the minimum spanning tree is $K-1$. The weight of each edge in the spanning tree is the optimal number of the relay coil. As discussed in Section 3.4.1, the optimal number of relay coils for a link is determined by the length of the link. The edges of the hexagonal tessellation have the same length e^{hex}, which is determined by the sensor density of the hexagonal tessellation λ_{hex}. Hence,

$$e^{hex} = 2 \cdot 3^{-\frac{1}{4}} \cdot \lambda_{hex}^{-\frac{1}{2}} , \tag{106}$$

Then the required number of the relay coils to connect K sensors based on the MST algorithm can be calculated as

$$N_{mst}^{hex} = (K-1) \cdot n_{opt}(2 \cdot 3^{-\frac{1}{4}} \cdot \lambda_{hex}^{-\frac{1}{2}}, B) , \tag{107}$$

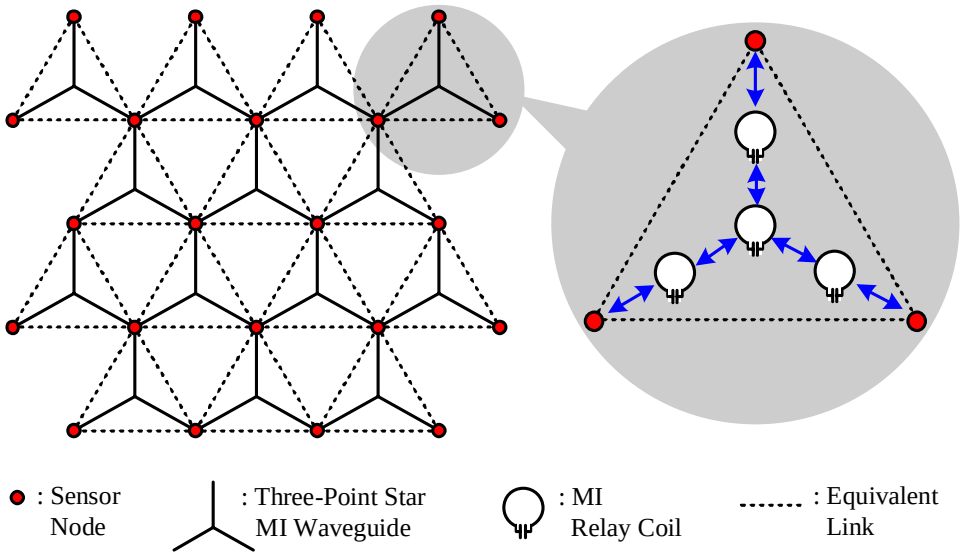


Figure 34: The MI waveguide deployment using TC algorithm in the WUSN with hexagonal tessellation topology.

where $n_{opt}(2 \cdot 3^{-\frac{1}{4}} \cdot \lambda_{hex}^{-\frac{1}{2}}, B)$ is the optimal coil number for each edge in the tessellation, which can be calculated by (105).

It should be noted that the WUSN constructed by the MST algorithm is only 1-connected. Consequently, the failure of any one sensor can disconnect the network.

Triangle Centroid (TC) Algorithm:

To enhance the robustness of the network, more edges should be established. If the MI waveguides are deployed along all the edges in the hexagonal tessellation, every sensor in the WUSN is connected to all the 6 neighbors in the tessellation. Consequently, the network becomes 6-connected. We define this deployment strategy as the *full-deployment*. However, in the full deployment strategy, the required number of relay coils for K sensors is doubled at the same time:

$$N_{full}^{hex} \simeq 2K \cdot n_{opt}(2 \cdot 3^{-\frac{1}{4}} \cdot \lambda_{hex}^{-\frac{1}{2}}, B) \,, \tag{108}$$

To reduce the number of relay coils, we change the positions of the MI waveguides so that multiple links can share one set of the MI waveguide. In particular, the three MI waveguides along the three edges of one triangle cell can be replaced by one MI waveguide

with a shape of the three-pointed star, as shown in Fig. 34. The center of the three-pointed star is located at the centroid of the triangle so that the sensors on all the three vertexes can use the same waveguide to communicate with each other directly. It can be proved that the total edge length of the three-pointed star is minimized if its center is located in the triangle centroid. Hence, the number of the relay coils to form the three-pointed star MI waveguide is minimized. To connect all the sensors in the WUSN, the three-pointed star MI waveguides are deployed in every other triangle in the tessellation, as shown in Fig. 34. The total number of triangles in the tessellation is approximately the same as the number of all sensors. Hence, the three-pointed star MI waveguides are deployed in half of the triangles. The edge length of the three-pointed star is $\sqrt{3} \cdot e_{hex}$. Then, the total required number of the relay coils to connect K sensors based on the TC algorithm is:

$$N_{tc}^{hex} \simeq \frac{K}{2} \cdot n_{opt}(2 \cdot 3^{\frac{1}{4}} \cdot \lambda_{hex}^{-\frac{1}{2}}, B) \; . \tag{109}$$

The WUSN constructed by the TC algorithm is 6-connected, the same as the full-deployment strategy. By comparing (108) with (109), we find that the required relay coil number of the TC algorithm is much smaller than that of the full deployment if the sensor density is not too low. Detailed numerical analysis is given in Section 3.4.3.

3.4.2.2 Deployment in WUSNs with Random Topology

Based on the analysis on the WUSNs with the hexagonal tessellation topology, we investigate the deployment algorithms for WUSNs with random topology in this section. Assuming that the underground sensors are uniformly distributed with the spatial density λ_{rand} (m^{-2}). Similar to the strategy in hexagonal tessellation, the MST algorithm are provided to achieve the minimum relay coil number, while the TC algorithm are implemented to provide the robustness to sensor failure with acceptable relay coil number.

MST Algorithm:

The MST algorithm for WUSN with random topology is similar to the MST algorithm in hexagonal tessellation. First, the edge lengths between any two underground sensor

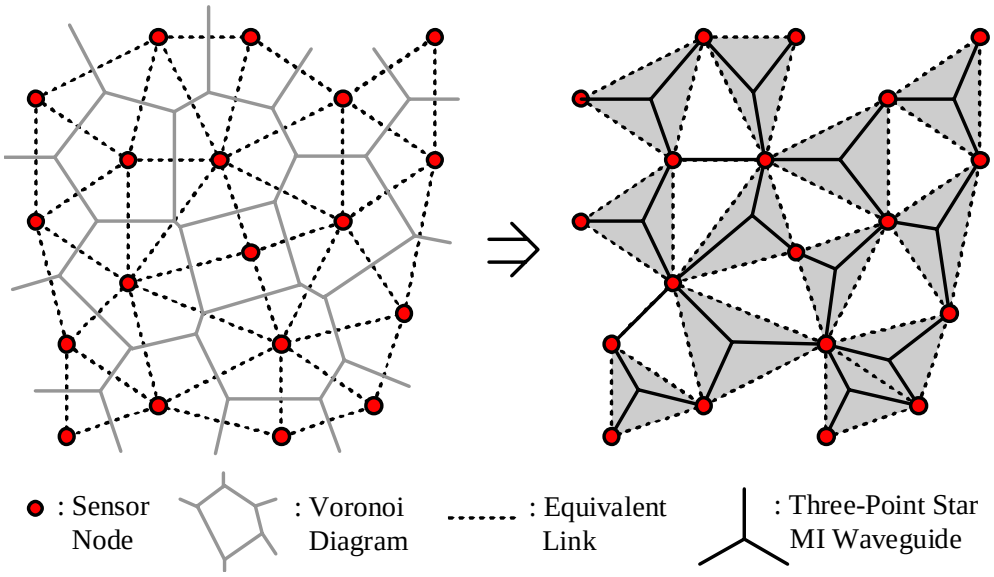


Figure 35: The MI waveguide deployment using TC algorithm in the WUSN with random topology.

nodes are calculated. Second, the optimal number of relay coils for each edge is calculated by (105), which is the weight of each edge. Third, the minimum spanning tree of the WUSN is found out by the Boruvka's algorithm [105]. Finally, the MI waveguides with the optimal relay coil number are deployed along each edge of the minimum spanning tree.

TC Algorithm:

As discussed previously, the TC algorithm needs to find out the centroid in each triangle cell of the network. In the hexagonal tessellation topology, the network is well partitioned into numerous equilateral triangle cells. Therefore the centroid in each triangle cell is easy to be located. However, in the random topology, the TC algorithm encounters two problems: 1) how to partition the random network into non-overlapped triangle cells; and 2) how to deploy the three-pointed star MI waveguide in those randomly distributed triangle cells.

To solve the above problems, we introduce the Voronoi diagram [101]. As shown in the left of Fig. 35, the Voronoi diagram of the sensors partitions the whole area into polygons (Voronoi cells). Each Voronoi cell contains only one sensor. All the points in one Voronoi

cell are closer to the sensor in this Voronoi cell than to any other sensors. By connecting the sensors that are in the adjacent Voronoi cells, the sensing area can be partitioned into non-overlapped triangle cells. Then in every other triangle cell, the MI waveguide is deployed along the three lines connecting the triangle vertexes and the centroid, which forms the three-pointed star MI waveguide, as shown in the right of Fig. 35. The detailed procedure of the TC algorithm in WUSNs with random topology is described in Algorithm 1.

Create the Voronoi diagram of the K sensors, and derive K Voronoi cells $\mathbf{VC} = \{VC_1, VC_2, ..., VC_K\}$.
Keep a subset $\mathbf{G}$ of $\mathbf{VC}$; $\mathbf{G}$ initially contains VC_1.
while (Not all Voronoi cells are in $\mathbf{G}$) **do**
  Find a Voronoi cell VC_x in $\mathbf{G}$ that has the neighbor Voronoi cells $\{VC_x^1, VC_x^2, ..., VC_x^j\}$ which are not in $\mathbf{G}$.
  Connect the adjacent sensors in $\{VC_x^1, VC_x^2, ..., VC_x^j\}$ and VC_x, and derive the non-overlapped triangle cells $\{Tr_1, Tr_2, ..., Tr_{j-1}\}$.
  if (j is odd) **then**
    In triangle cells $Tr_1, Tr_2, Tr_4, ..., Tr_{j-1}$, deploy the MI waveguide along the the three lines connecting the vertexes and the centroid.
  else
    In triangle cells $Tr_1, Tr_3, Tr_5, ..., Tr_{j-1}$, deploy the MI waveguide along the the three lines connecting the vertexes and the centroid.
  end if
  Add $\{VC_x^1, VC_x^2, ..., VC_x^j\}$ to $\mathbf{G}$.
end while

Algorithm 1: TC Algorithm for MI Waveguide Deployment in WUSNs with Random Topology

For the random topology, the WUSN constructed by the MST algorithm is only 1-connected. Meanwhile, the network created by the TC algorithm in random topology is k-connected, where $k \geq 3$. The required number of relay coils of the MST algorithm as well as the TC algorithm in the WUSN with random topology cannot be accurately estimated since the positions of the sensors are highly random. The simulation analysis is given in the next section.

3.4.3 Performance Evaluation

In this section, we numerically evaluate the required relay coil number and the network robustness of the MST algorithm and the TC algorithm in both WUSNs with hexagonal tessellation topology and WUSNs with random topology. The performance of the full-deployment strategy is also shown as a reference. In the following simulations, 100 sensors are deployed in a square area according to the hexagonal tessellation topology or the random topology. The size of the square area is determined by the sensor density. The MI waveguide parameters used in the simulations are the same as the parameters used in Section 3.4.1.

3.4.3.1 Hexagonal Tessellation Topology

In Fig. 36(a), the required relay coil numbers of the deployment algorithms are given as a function of the sensor density in the WUSN with hexagonal tessellation topology. Fig. 36(a) shows that the relay coil number required by the TC algorithm is slightly larger than the number required by the MST algorithm but much smaller than the number required by the full-deployment strategy. Meanwhile, the network constructed by the TC algorithm is 6-connected, the same as the full-deployment strategy and far more robust than the 1-connected network constructed by the MST algorithm.

Therefore, in the WUSNs with hexagonal tessellation topology, the TC algorithm achieves both small relay coil number and high network robustness.

3.4.3.2 Random Topology

Fig. 37 shows the deployment results of the MST algorithm, the TC algorithm and the full-deployment strategy. The network constructed by the MST algorithm is only 1-connected. Consequently, the failure of any one sensor can disconnect the network. One the other hand, the networks constructed by the TC algorithm and the full-deployment strategy have the same network topology, since the three-pointed star MI waveguide in a triangle cell is

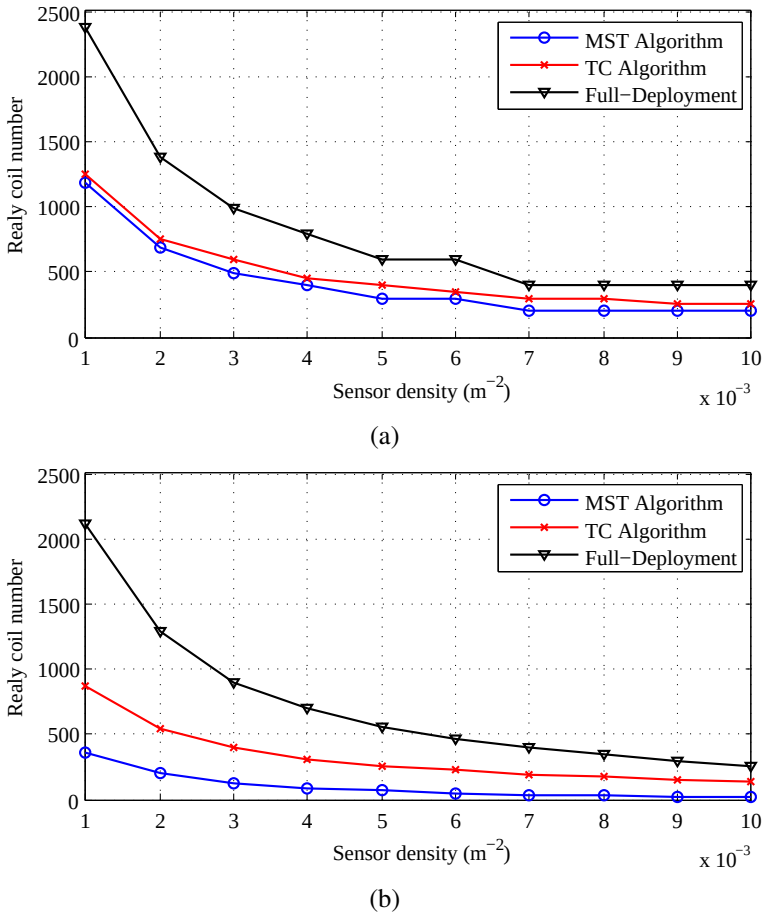


Figure 36: The number of relay coils to connect 100 sensors in WUSNs with (a) hexagonal tessellation topology and (b) random topology.

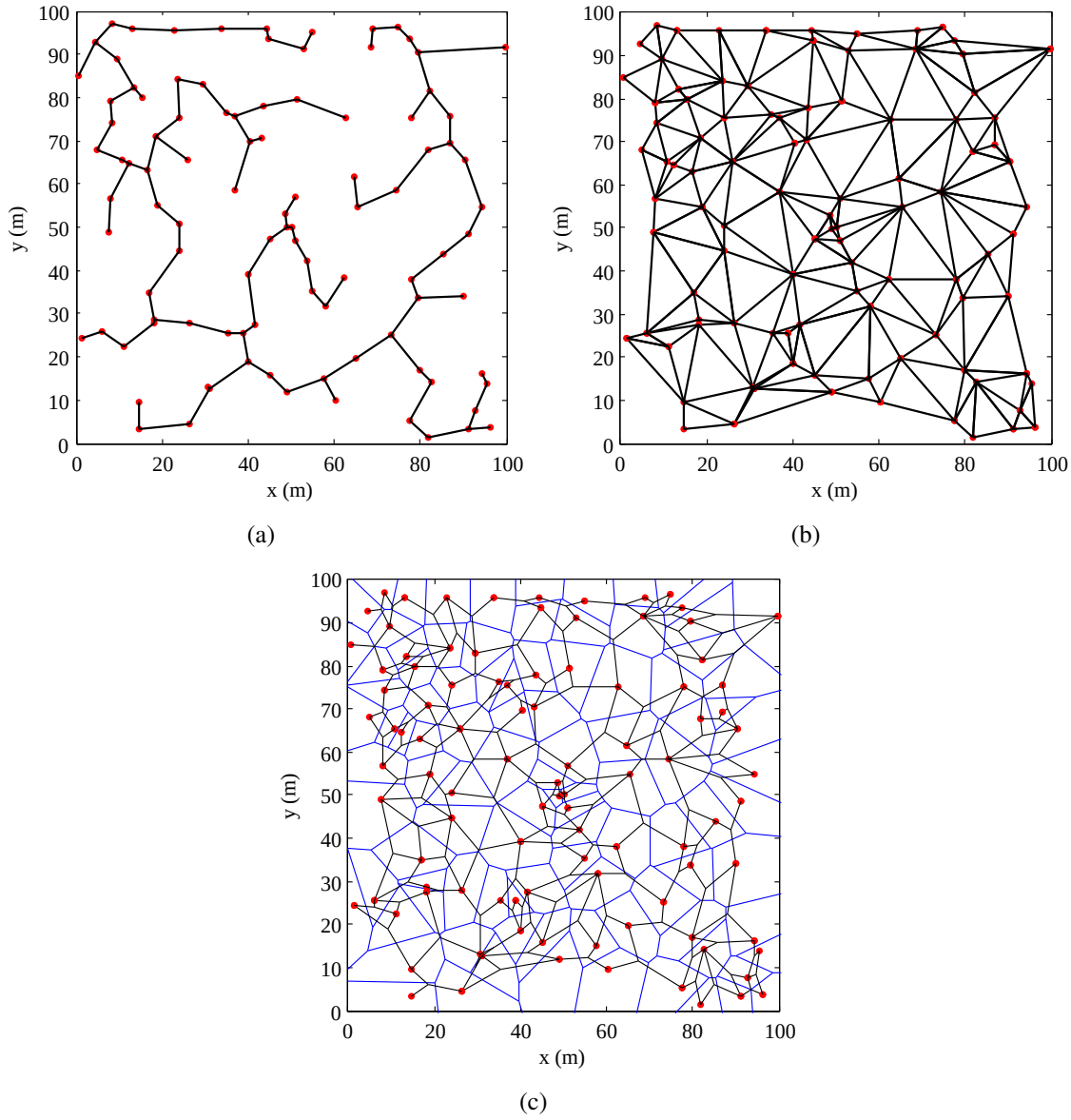


Figure 37: The deployment results of (a) the MST algorithm, (b) the TC algorithm, and (c) the full-deployment strategy. (The red dots are the sensors; the black lines represent the MI waveguides; and the blue cells are the Voronoi diagrams. 100 sensors are uniformly distributed with of a spatial intensity $\lambda_{rand} = 0.01\ m^{-2}$.)

equivalent to the three MI waveguides on the edges of the triangle cell. Except the sensors on the border, the network constructed by the TC algorithm or the full-deployment strategy is k-connected. k is determined by the sensor topology and $k \geq 3$. Therefore, the TC algorithm and the full-deployment strategy are more robust to sensor failures.

In Fig. 36(b), the required relay coil numbers of the deployment algorithms are given as a function of the sensor density in the WUSN with random topologies. It indicates that the relay coil number required by the TC algorithm is obviously larger than the number required by the MST algorithm. As the sensor density increases, the differences in terms of the coil number between the deployment algorithms become smaller. Compared with the hexagonal tessellation topology, the advantages of the MST algorithm in terms of the relay coil number is much more obvious in the random topology.

Therefore, in the WUSNs with random topology, the relay coils number required by the MST algorithm is significantly smaller than other deployment algorithms. However, the MST algorithm is not robust to sensor failures. Although the The TC algorithm requires more relay coils than the MST algorithm, it can construct a k-connected WUSN. Moreover, the required coil number of the TC algorithm is much smaller than the number required by the full-deployment strategy.

CHAPTER IV

WUSNS IN UNDERGROUND MINES AND TUNNELS

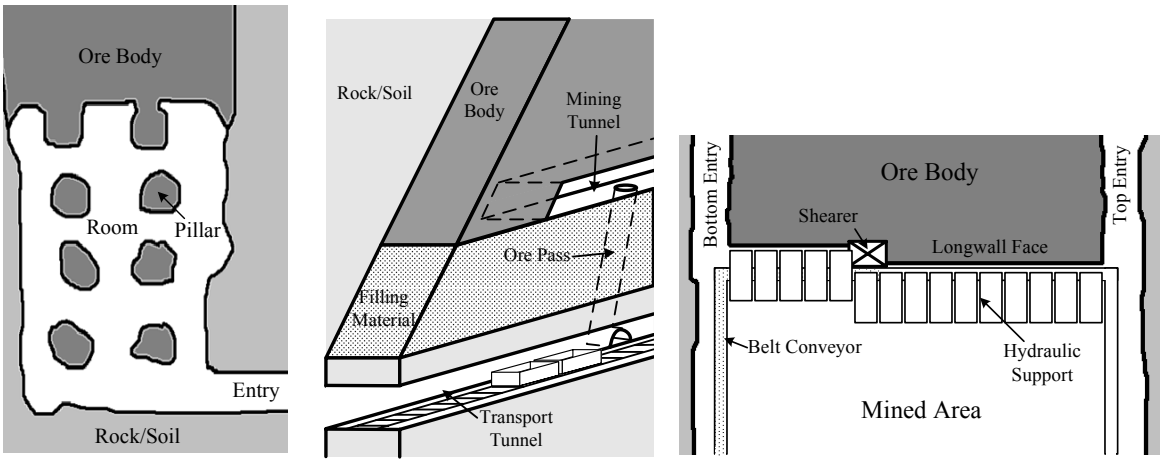
4.1 Motivation and Related Work

Reliable and efficient WUSNs in underground mines and tunnels are important to improve the safety and productivity [2, 23]. Wireless communications experience severe fading problems in underground mines and tunnels[4]. Due to the reflections of EM waves on tunnel walls, the channel characteristics in these environments are dramatically different from the terrestrial wireless channels [84, 86, 21]. Moreover, the tunnels in operation are filled with mobile vehicles with random size and positions. The reflections and the diffractions on the vehicles make the wireless channel in the tunnel even more complicated. To setup reliable and efficient wireless communication systems in underground mines and tunnels, the analytical channel model that explicitly contains the dependence on the tunnel geometry, vehicular traffic information, and other communication parameters are needed.

The underground mines have complex structures: multiple passageways are developed to connect the aboveground entrance and different mining areas. The structure of mining area is determined by mining methods, while the mining methods are dertermined by the shape and position of the ore body [38].

- If the ore body is flat and competent, room-and-pillar mining can be implemented. The mining area can be viewed as a big room with some randomly shaped pillars in it, as shown in Fig. 38(a).

- If the ore body has a steep dip, cut-and-fill mining, sublevel stoping or shrinkage stoping can be employed. Mines using those techniques have similar structures: the mining area consists of several types of tunnels, e.g. mining tunnel and transport tunnel. The sectional plan of cut-and-fill mining is shown in Fig. 38(b).



(a) Plan of room-and-pillar mining (b) Cut-and-fill mining (c) Plan of longwall mining

Figure 38: Mine structure of different mining methods

- If the ore body has a large, thin, seam-type shape, longwall mining is preferred, as shown in Fig. 38(c). Besides the entry tunnels, the mining area near the longwall face can also be modeled as a tunnel since it is encircled by the hydraulic support and the longwall face.

Therefore, underground mines require two kinds of channel models. The *tunnel channel model* is used to describe the signal propagation in passageways and mining area tunnels. On the other hand, the *room-and-pillar channel model* characterizes the wireless channel of room-and-pillar mining areas. It should be noted that the structure of road/subway tunnels is similar to that of underground mine tunnels, thus they can share the same tunnel channel model.

Existing channel models for tunnels include the GO model [59], the waveguide model [34] and full wave model [95]. In the GO model, EM waves are approximately modeled as optical rays. The EM field is obtained by summing the contributions of rays undergoing reflections on the tunnel walls. In [41, 111], the rays diffracted near tunnel wedges are considered to improve the accuracy of the GO model. Except in some very idealized situations, e.g., the waveguide with two perfectly reflecting side walls [59], the GO model depends

on computer simulations to obtain numerical solutions, and the computational burden increases dramatically as the signal path is prolonged [70, 110]. In the waveguide model, the tunnel behaves as an oversized waveguide with imperfectly lossy walls. Maxwell's equations are solved by taking consideration of the boundary conditions. The eigenfunctions and propagation constants for the EM field of all possible modes are provided in [50]. The waveguide model assumes that there is only the lowest mode signal propagation in the tunnel. However, since the operating frequency (UHF) is much higher than the cut-off frequency in tunnels, the large number of modes will be exited near the transmitter antenna [112]. Consequently, the waveguide model can not characterize the multi-mode operating channel in the near region. Full wave models can solve the Maxwell's equations with arbitrary boundary conditions using numerical methods, such as FDTD [95]. The partial differential equations are solved at discrete time and discrete points (finite grid). However, it is required that the size of the finite grid in space should be less than one tenth of the free space wavelength, and the time integration step must be less than the grid size divided by the velocity of the light. Given the large size of tunnels and the high operating frequency (UHF), the computational burden exceeds well beyond the capacity of existing computers. For curved tunnels, the additional attenuation coefficients of each propagation mode caused by the tunnel curvature are given in [69]. For the signal propagation around tunnel junctions, the coupling from the main tunnel to the sub-tunnel is analyzed in [52].

Currently, there is no existing channel model for room-and-pillar mining area. In [56, 55], some experimental measurements are provided. It is indicated that the signal experiences higher attenuation in room-and-pillar environments than in tunnels. Additionally, the multipath fading is severe in both near and far region of the transmitter.

For tunnels with vehicular traffic flow, current channel analyses are limited to either experiments [107, 46] or numerical methods (i.e., GO model [8, 24, 25, 54] and Full Wave model [5]). These experimental and numerical solutions cannot provide explicit

description on the effect of tunnel geometry, the vehicular traffic information, antenna position/polarization, operating frequency, and other environmental or communication parameters. Moreover, the numerical solutions require a great amount of input data of detailed geometric information of the vehicular traffic flow, including the exact size and position of each vehicle in the traffic flow, which is infeasible to acquire from the in-operation tunnels.

In this chapter, we first provide a new hybrid model that combines the GO model and waveguide model using Poisson sum formula. Analytical solutions for both near and far regions are developed for tunnel environments. Combined with the shadow fading model, our model can also characterize the wireless channel in the room-and-pillar mining area. Then based on the channel model in empty underground mines and tunnels, we extend our work to characterize the influence of the vehicular traffic flow on the signal propagations.

According to the channel models, the wireless link error caused by the multipath fading in underground tunnels is much more severe than the terrestrial wireless channels. To solve this problem, spatial diversity-based techniques including MIMO (Multiple Input Multiple Output) and Cooperative Communication system can be utilized. In particular, the MIMO system employs multiple antenna elements at both transmitter and receiver to achieve the spatial diversity [96], which is suitable for large devices such as the base stations and the mobile terminals on vehicles. In contrast, the cooperative communication system [75] explores the broadcast nature of the wireless channel and utilize multiple wireless nodes with single antenna to form a virtual MIMO, which is suitable for small and low-cost devices such as wireless sensors and handsets.

The MIMO capacity has been widely analyzed in terrestrial wireless communication systems. In [96] and [60], the MIMO capacity over the additive Gaussian channel with and without multipath fading is analyzed. In [77], the effect of MIMO antenna geometry on capacity is analyzed to mitigate the impact of the correlated MIMO channel. In [65], it is proved that the distribution of the MIMO capacity in terrestrial channel follows a normal distribution under the condition that the number of antenna elements is large. All

the above works are based on the terrestrial wireless channel model that is simpler and fuzzier than the tunnel channel model in two aspects: 1) The channel gain of the terrestrial channel is assumed to be a Rayleigh random variable multiplied by a power function of the transmission distance. However, the parameters of the Rayleigh fading cannot be accurately calculated. In contrast, the channel gain in underground tunnels is a weighted sum of multiple propagation modes. The intensity and the field distribution of each mode can be accurately characterized. 2) In terrestrial channel, each pair of TX and RX antenna elements in the MIMO system is assumed to have the same mean channel gain. However, in underground tunnels, the positions of the transceivers have significant influences on the channel gain. Hence the MIMO antenna geometry significantly affects MIMO channel capacity in tunnels.

In [26], current terrestrial MIMO techniques are evaluated in tunnel environments by simulations. It shows that the MIMO technique can also effectively mitigate multipath fading in underground tunnels. In [58, 64], the MIMO channel capacity in empty waveguide and cavity channels are calculated using the modal expansion technique. Those existing works on MIMO capacity in tunnels are based on the empty tunnel channel model. However, in real underground tunnels, there are a large number of random obstructions, such as vehicles and mining machines. Hence the randomness of the MIMO capacity caused by the random obstructions has significant influence on the performance of the MIMO systems in underground tunnels. To the best knowledge of the authors, the capacity distribution as well as the outage behavior of the MIMO systems in tunnels have not been investigated yet.

The cooperative communication technique has also been intensively investigated in terrestrial environments recently. In [51], several efficient cooperative schemes are proposed, and the corresponding outage behavior in terms of outage probability are investigated in high SNR regime. In [7], the outage capacity of cooperative communication system is

calculated in low SNR regime. In [76], a centralized cooperative relay assignment protocol is proposed to maximize the minimum cooperative capacity in the whole network. In [10], a distributed relay assignment protocols are proposed. However, it still requires the information exchanges among the source node, relay nodes and destination node, which is difficult to achieve in networks with high dynamic topology, such as the vehicular networks in road tunnels. In [73], a nearest neighbor relay assignment protocol is proposed based on the analysis on the outage probability, which is fully distributed and only requires the local position information. Similar to the existing works on MIMO system, the above works on cooperative communications are also based on the terrestrial wireless channel. Currently, there is no existing paper on either the outage analysis or the relay assignment for cooperative communications in underground tunnels.

At the end of this chapter, we analyze the capacity distribution and outage behavior of the MIMO and cooperative communication system in underground tunnel environments. The parameters of the capacity distribution and the outage probability are explicitly expressed as functions of the tunnel environmental conditions, the antenna geometry (for MIMO), and the relay assignment strategy (for cooperative communications). Then, based on the capacity and outage analysis, the optimal MIMO antenna geometry and the optimal cooperative relay assignment protocol are developed for wireless communications in underground tunnels.

4.2 Channel Modeling

To settle the problems of current tunnel channel models, we introduce the multimode model, which can be viewed as a multi-mode operating waveguide model. Since the modes derived by the waveguide model are actually all possible solutions for the Maxwell's equations, only the EM waves that have the same shapes as those modes are possible to exist in the tunnel. However, the intensity of each mode depends on the excitation, which cannot be given by the waveguide model. Hence, the GO model is involved to analyze the EM field

distribution for the excitation plane, i.e., the tunnel cross-section that contains the transmitter antenna. This field distribution can be viewed as the weighted sum of the field of all modes. The mode intensities are estimated by a mode-matching technique. Once the mode intensity is determined in the excitation plane, the mode propagation is mostly governed by the tunnel itself. Then the EM field in the rest of the tunnel can be predicted by summing the EM field of each mode.

The room-and-pillar environment can be viewed as a planar air waveguide superimposed with some random distributed and random shaped pillars in it. A simplified multimode model is able to describe the EM wave propagation in the planar air waveguide. The random distributed and random shaped pillars form an environment very similar to a terrestrial metropolitan area with many buildings. Hence, the shadow fading model can be used to describe the signal's slow fading caused by the reflection and diffraction on those pillars.

In the remainder of this section, we first develop the multimode model for tunnel channels. Then the multimode model is extended to cover the room-and-pillar case.

4.2.1 Multimode Model in Tunnel Environments

Actual tunnel cross sections are generally in-between a rectangle and a circle. However, the EM field distribution and attenuation of the modes in rectangle waveguide are almost the same as the circular waveguide [56]. Hence, in our model, the tunnel cross section is treated as an equivalent rectangle with a width of $2a$ and a height of $2b$. A Cartesian coordinate system is set with its origin located at the center of the rectangle tunnel. k_v, k_h and k_a are the complex electrical parameters of the tunnel vertical/horizontal walls and the air in the tunnel, respectively, which are defined as: $k_v = \varepsilon_0\varepsilon_v + \frac{\sigma_v}{j2\pi f_0}$, $k_h = \varepsilon_0\varepsilon_h + \frac{\sigma_h}{j2\pi f_0}$ and $k_a = \varepsilon_0\varepsilon_a + \frac{\sigma_a}{j2\pi f_0}$, where ε_v, ε_h and ε_a are the relative permittivity for vertical/horizontal walls and the air in the tunnel; ε_0 is the permittivity in vacuum space; σ_v, σ_h and σ_a are their conductivity; f_0 is the central frequency of the signal. The three areas are assumed

to have the same permeability μ_0. The wave number in the tunnel space is given by $k = 2\pi f_0 \sqrt{\mu_0 \varepsilon_0 \varepsilon_a}$. We define the relative electrical parameter $\overline{k_v}$ and $\overline{k_h}$ for concise expression, which are $\overline{k_v} = k_v/k_a$ and $\overline{k_h} = k_h/k_a$. We assume that the transmitter antenna is an X-polarized electrical dipole. The results for Y-polarized antenna can be obtained simply by interchanging the x- and y-axes. The major polarized field plays a dominant role inside the tunnel and the coupling term can be omitted. Hence, in our multimode model, we only consider the major polarized field.

4.2.1.1 Multiple Mode Propagation in Tunnels

The propagation of EM waves in tunnels can be viewed as the superposition of multiple modes with different field distribution and attenuation coefficients. By solving the Maxwell's equations, the field distribution of each mode can be derived in the form of eigenfunctions [29, 50, 27]:

$$E^{eign}_{m,n}(x,y) \simeq \sin\left(\frac{m\pi}{2a}x + \varphi_x\right) \cdot \cos\left(\frac{n\pi}{2b}y + \varphi_y\right) \tag{110}$$

where $\varphi_x = 0$ if m is even; $\varphi_x = \frac{\pi}{2}$ if m is odd; $\varphi_y = 0$ if n is odd; $\varphi_y = \frac{\pi}{2}$ if n is even.

The field at any position (x, y, z) inside the tunnel can be obtained by summing up the field of all significant modes, which is given by:

$$E^{Rx}(x,y,z) = \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} C_{mn} \cdot E^{eign}_{m,n}(x,y) \cdot e^{-(\alpha_{mn}+j\beta_{mn})\cdot z} \tag{111}$$

where C_{mn} is the mode intensity on the excitation plane; α_{mn} and β_{mn} are the attenuation coefficient and the phase-shift coefficient, respectively, which is given by [29, 34, 50]:

$$\begin{aligned}
\alpha_{mn} &= \frac{1}{a}\left(\frac{m\pi}{2ak}\right)^2 Re\frac{\overline{k_v}}{\sqrt{\overline{k_v}-1}} + \frac{1}{b}\left(\frac{n\pi}{2bk}\right)^2 Re\frac{1}{\sqrt{\overline{k_h}-1}} \\
\beta_{mn} &= \sqrt{k^2 - \left(\frac{m\pi}{2a}\right)^2 - \left(\frac{n\pi}{2b}\right)^2}
\end{aligned} \tag{112}$$

The waveguide model considers that only the lowest order mode exists in the tunnel, i.e. $C_{11} = 1$ and $C_{mn} = 0$ if $(m,n) \neq (1,1)$. However, in the near region of the transmitter,

there exist multiple modes. The intensity of each modes need to be determined. In the next step, we first analyze the field distribution of the excitation plane by the GO model. Then a mode matching technique is utilized to convert the sum of rays of the GO model to the sum of modes. Consequently, the mode intensity C_{mn} on the excitation plane can be obtained.

4.2.1.2 Field Analysis of the Excitation Plane by the GO Model

The total field in the tunnel is equal to the sum of ray contributions from all reflection images and the source. The reflection images and the source on the excitation plane are located as Fig. 39 shows. Due to the geometry characteristic of rectangle cross section shape, the images and the reflection rays have the following properties:

- The ray coming from image $I_{p,q}$ experiences $|p|$ times reflection from vertical wall and $|q|$ times reflection from horizontal ceiling/floor.
- Suppose that α is the incident angle on the ceiling/floor, and β is the incident angle on the wall. For a certain ray, these angles remain the same.

Consider that the transmitter is located at the coordinate $(x_0, y_0, 0)$, and the observation point is set at the coordinate (x, y, z). The field at the transmitter is E_0. The field at the observation point is the sum of the rays coming from all the images:

$$E^{Rx}(x, y, z) = E_0 \cdot \sum_{p=-\infty}^{\infty} \sum_{q=-\infty}^{\infty} \left[\frac{\exp(-jkr_{p,q})}{r_{p,q}} \right] \cdot S(\overline{k_v})^{|p|} \cdot R(\overline{k_h})^{|q|} \tag{113}$$

where $r_{p,q}$ is the distance between image $I_{p,q}$ and the receiver:

$$r_{p,q} = \sqrt{(2pa \pm x_0 - x)^2 + (2qb \pm y_0 - y)^2 + z^2}; \tag{114}$$

where "+" sign is for the case when p or q is even, while "−" sign is for that case when p or q is odd. $R(\overline{k_h})$ and $S(\overline{k_v})$ are the reflection coefficients on the horizontal and vertical walls.

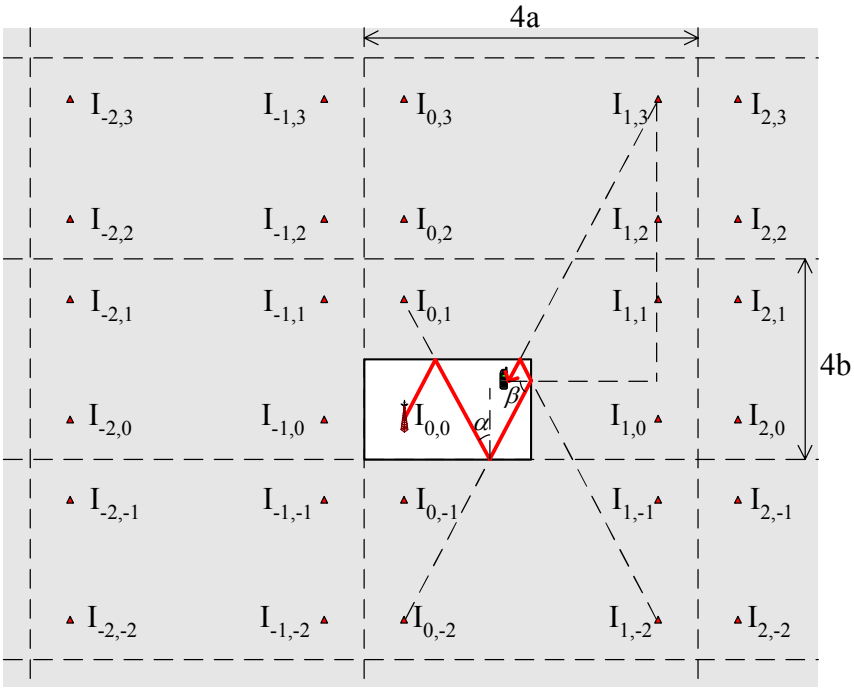


Figure 39: The set of images in the excitation plane in a rectangular cross section tunnel.

If the tunnel size is much larger than the free-space wavelength of the incidence wave, the reflection coefficients are given by [36]:

$$R(\overline{k_h}) = \frac{\cos\alpha - \sqrt{\overline{k_h} - \sin^2\alpha}}{\cos\alpha + \sqrt{\overline{k_h} - \sin^2\alpha}}; \qquad S(\overline{k_v}) = \frac{\overline{k_v}\cos\beta - \sqrt{\overline{k_v} - \sin^2\beta}}{\overline{k_v}\cos\beta + \sqrt{\overline{k_v} - \sin^2\beta}} \tag{115}$$

where α is the incident angle of rays on the horizontal ceiling/floor; and β is the incident angle of rays on the vertical walls. Since we only consider the rays with small grazing angle (otherwise the path loss is huge), $R(\overline{k_h})$ and $S(\overline{k_v})$ can be approximated as:

$$\begin{aligned} R(\overline{k_h}) &= -\exp\left(\frac{-2\sin\alpha}{\sqrt{\overline{k_h}-1}}\right) = -\exp\left(\frac{-2}{\sqrt{\overline{k_h}-1}} \cdot \frac{|2qb \pm y_0 - y|}{r_{p,q}}\right); \\ S(\overline{k_v}) &= -\exp\left(\frac{-2\overline{k_v}\sin\beta}{\sqrt{\overline{k_v}-1}}\right) = -\exp\left(\frac{-2\overline{k_v}}{\sqrt{\overline{k_v}-1}} \cdot \frac{|2pa \pm x_0 - x|}{r_{p,q}}\right) \end{aligned} \tag{116}$$

4.2.1.3 Mode-Matching in the Excitation Plane

By rearranging the ray sum in (113), we can divide the ray sum into four parts:

$$E^{Rx}(x,y,z) = \sum_{p,q=-\infty}^{\infty} f(4qa + x_0 - x, 4pb + y_0 - y) + \sum_{p,q=-\infty}^{\infty} f(4qa + x_0 - x, 4pb + 2b - y_0 - y)$$
$$+ \sum_{p,q=-\infty}^{\infty} f(4qa + 2a - x_0 - x, 4pb + y_0 - y) + \sum_{p,q=-\infty}^{\infty} f(4qa + 2a - x_0 - x, 4pb + 2b - y_0 - y) \tag{117}$$

where $f(u,v)$ is the function defined as:

$$f(u,v) = E_0 \frac{\exp(-jk\sqrt{u^2+v^2+z^2})}{\sqrt{u^2+v^2+z^2}} \cdot (-1)^{p(v)+q(u)} \tag{118}$$
$$\cdot \exp\left[\frac{-2}{\sqrt{u^2+v^2+z^2}}\left(\frac{|v|p(v)}{\sqrt{k_h}-1} + \frac{|u|\overline{k_v}q(u)}{\sqrt{k_v}-1}\right)\right]$$

where $p(v)$ and $q(u)$ are discontinuous functions that takes values of $0, \pm1, \pm2, \cdots$. To facilitate the mode matching, we approximately transform $p(v)$ and $q(u)$ to continuous functions. Then,

$$p(v) = \frac{|v|}{2b}; \quad q(u) = \frac{|u|}{2a} \tag{119}$$

Note that each part in (117) is a periodic function of $4a$ and $4b$. We first consider the first part in (117). According to 2-dimension Poisson Summation Formula [108], the sum can be converted to:

$$\sum_{p,q=-\infty}^{\infty} f(4qa + x_0 - x, 4pb + y_0 - y) = \frac{1}{4a}\frac{1}{4b}\sum_{m=-\infty}^{\infty}\sum_{n=-\infty}^{\infty} \mathscr{F}_1(m,n) \cdot e^{j\frac{m\pi}{2a}x} e^{j\frac{n\pi}{2b}y} \tag{120}$$

The coefficient $\mathscr{F}_1(m,n)$ is the 2-dimension Fourier transform of the function $f(x_0-x, y_0-y)$ in the first part in (117):

$$\mathscr{F}_1(m,n) = \iint_{-\infty}^{\infty} f(x_0-x, y_0-y) e^{-j\frac{m\pi}{2a}x} e^{-j\frac{n\pi}{2b}y} dxdy \tag{121}$$

We utilize the saddle-point method [28] to derive the closed-form result of the integration. 2-dimensional saddle point method provides the approximate integration results of

the form $\int_a^b \int_c^d g(u,v)e^{h(u,v)}dudv$. The integration in (121) has exactly the same form, where

$$g(u,v) = \frac{E_0}{\sqrt{u^2+v^2+z^2}} \tag{122}$$

$$h(u,v) = -jk\sqrt{u^2+v^2+z^2} - j\frac{m\pi}{2a}(x_0-u) - j\frac{n\pi}{2b}(y_0-v) - \frac{2}{\sqrt{u^2+v^2+z^2}}\Big(\frac{|v|p(v)}{\sqrt{k_h}-1} + \frac{|u|\overline{k_v}q(u)}{\sqrt{k_v}-1}\Big) \tag{123}$$

Note that the last term in (123) can be omitted since $\sqrt{u^2+v^2+z^2}$ is much larger than 1.

The saddle point of the integration is (u_0, v_0) so that $\frac{\partial h(u,v)}{\partial u}|_{u=u_0} = 0$ and $\frac{\partial h(u,v)}{\partial v}|_{v=v_0} = 0$. Hence the saddle point for (121) can be calculated as:

$$u_0 = |z| \cdot \tan\theta_m; \qquad v_0 = |z| \cdot \tan\theta_n \tag{124}$$

where

$$\theta_m = \arcsin\frac{m\pi}{2ka}; \qquad \theta_n = \arcsin\frac{n\pi}{2kb} \tag{125}$$

Then the approximate results of the integration can be expressed as:

$$\iint_{-\infty}^{\infty} g(u,v)e^{h(u,v)}dudv \simeq g(u_0,v_0)\cdot e^{h(u_0,v_0)}\cdot\frac{\pi}{\sqrt{\left|\frac{\partial^2 h(u_0,v_0)}{\partial u^2}\right|\left|\frac{\partial^2 h(u_0,v_0)}{\partial v^2}\right| - \left(\frac{\partial^2 h(u_0,v_0)}{\partial u\partial v}\right)^2}} \tag{126}$$

By this way, the approximate result of the integration in (121) can be obtained by substituting (122), (123) and (124) into (126). Note that here we only care about the field $E^{Rx}(x,y,z)$ on the excitation plane where $z = 0$. Therefore, the coefficient $\mathscr{F}_1(m,n)$ on the excitation can be expressed as:

$$\mathscr{F}_1(m,n) \simeq E_0\frac{\pi}{\sqrt{1-(\frac{m\pi}{2ak})^2-(\frac{n\pi}{2bk})^2}}\cdot e^{-j(\frac{m\pi}{2a}x_0+\frac{n\pi}{2b}y_0)} \tag{127}$$

By this way, the first part of the ray sum in (117) can be converted to the sum of complex modes in (120). In the same way, the Poisson sum formula can be utilized in the rest parts

in (117), and the coefficients $\mathscr{F}_2(m,n)$, $\mathscr{F}_3(m,n)$, $\mathscr{F}_4(m,n)$ can also be derived by the saddle point method. Therefore, the field in the excitation plane can be expressed as:

$$\begin{aligned}
E^{Rx}(x,y,0) &= \frac{1}{4a}\frac{1}{4b}\sum_{m=-\infty}^{\infty}\sum_{n=-\infty}^{\infty}\left[\mathscr{F}_1(m,n)+\mathscr{F}_2(m,n)+\mathscr{F}_3(m,n)+\mathscr{F}_4(m,n)\right]\cdot e^{j\frac{m\pi}{2a}x}e^{j\frac{n\pi}{2b}y} \\
&= \sum_{m=-\infty}^{\infty}\sum_{n=-\infty}^{\infty}\frac{E_0\pi}{16ab\sqrt{1-(\frac{m\pi}{2ak})^2-(\frac{n\pi}{2bk})^2}}\cdot e^{j\frac{m\pi}{2a}x}e^{j\frac{n\pi}{2b}y} \\
&\qquad\cdot\left(e^{-j\frac{m\pi}{2a}x_0}e^{-j\frac{n\pi}{2b}y_0}+e^{j\frac{m\pi}{2a}x_0-m\pi}e^{j\frac{n\pi}{2b}y_0-n\pi}-e^{j\frac{m\pi}{2a}x_0}e^{j\frac{n\pi}{2b}y_0-n\pi}-e^{j\frac{m\pi}{2a}x_0-m\pi}e^{j\frac{n\pi}{2b}y_0}\right) \\
&= \sum_{m=1}^{\infty}\sum_{n=1}^{\infty}\frac{E_0\pi}{ab\sqrt{1-(\frac{m\pi}{2ak})^2-(\frac{n\pi}{2bk})^2}}\sin\left(\frac{m\pi}{2a}x_0+\varphi_x\right)\cos\left(\frac{n\pi}{2b}y_0+\varphi_y\right)\sin\left(\frac{m\pi}{2a}x+\varphi_x\right)\cos\left(\frac{n\pi}{2b}y+\varphi_y\right)
\end{aligned} \tag{128}$$

Note that (128) is exactly the weighted sum of the eigenfunction of each propagation mode in (110). The weight of each eigenfunction is the mode intensity C_{mn} in the excitation plane:

$$C_{mn} = \frac{E_0\pi}{ab\sqrt{1-(\frac{m\pi}{2ak})^2-(\frac{n\pi}{2bk})^2}}\sin\left(\frac{m\pi}{2a}x_0+\varphi_x\right)\cos\left(\frac{n\pi}{2b}y_0+\varphi_y\right) \tag{129}$$

By substituting (110), (112) and (129) into (111), the field of any position in the tunnel can be analytically calculated.

Then suppose the transmitting power is P_t; G_t and G_r are the antenna gains of the transmitter and the receiver, respectively. The predicted received signal power at the coordinate (x,y,z) is given by:

$$P_t(x,y,z) = P_tG_tG_r\left(\frac{1}{E_0}\sum_{m,n}C_{mn}\cdot E^{eign}_{m,n}(x,y)\cdot e^{-(\alpha_{mn}+j\beta_{mn})\cdot z}\right)^2 \tag{130}$$

4.2.1.4 Power Delay Profile for Wideband Signal

If the transmitting signal is wideband, significant signal distortion may happen due to the dispersion effect of the tunnel waveguide, which will cause severe inter symbol interference (ISI). We characterize this channel effect by calculating the power delay profile (PDP).

We assume that the wideband signal $s(t)$ has a bandwidth of B around the central frequency f_0, i.e. $f \in [f_0 - B/2, f_0 + B/2]$. The frequency spectrum of the signal is characterized by its fourier transform $S(f)$. This signal can be viewed as the sum of all the sinusoidal waves whose frequencies fall into the band. The intensity of each sinusoidal wave is determined by the fourier transform $S(f)$. In addition, if the signal $s(t)$ is real, then its fourier transform $S(f)$ is an even function of the frequency f. Hence,

$$s(t) = \int_{f_0 - B/2}^{f_0 + B/2} S(f) \cdot 2\cos(2\pi f \cdot t)\, df \tag{131}$$

Different frequency elements in (131) have different wave number $k(f)$. Consequently, the mode intensity $C_{mn}(f)$, field distribution $E_{m,n}^{eign}(x, y, f)$, attenuation coefficients $\alpha_{mn}(f)$ and phase-shift coefficients $\beta_{mn}(f)$ become the functions of the frequency f.

Moreover, the propagation delay of a certain mode also varies with the frequency. For a sinusoidal wave signal with a single frequency f, the propagation delay of EH_{mn} mode can be calculated by $\tau_{mn}(f) = z/v_{mn}(f)$, where $v_{mn}(f)$ is the group velocity that is given by:

$$v_{mn}(f) = c\sqrt{1 - \left(\frac{c\sqrt{\left(\frac{m\pi}{2a}\right)^2 + \left(\frac{n\pi}{2b}\right)^2}}{2\pi f}\right)^2} \tag{132}$$

According to (132), the group velocity is a function of both the operating frequency f and the mode's order (m, n). For the same mode, different frequency signals have different propagation delay. For a single frequency, different modes also have different delay. Hence, both the dispersion among modes and the dispersion among frequency elements should be considered when calculating the power delay profile of a wideband signal. At a certain time t and position (x, y, z) in the tunnel, the received power of a wideband signal P_{WB} can be calculated by summing up the contributions of all the arrived significant modes of all frequency elements, which is given by:

$$P_{WB}(x, y, z, t) = P_t G_t G_r \left\{ \frac{1}{E_0} \sum_{m,n} \int_{f_0 - \frac{B}{2}}^{f_0 + \frac{B}{2}} \Big[C_{mn}(f) \cdot E_{m,n}^{eign}(x, y, f) \cdot e^{-\alpha_{mn} \cdot z} \cdot S(f) \cdot \delta\left(t - \frac{z}{v_{mn}(f)}\right) \cdot \cos(2\pi f t - \beta_{mn} z) \Big] df \right\}^2 \tag{133}$$

where

$$\delta(x) = \begin{cases} 1, & \text{if } x \geq 0 \\ 0, & \text{otherwise} \end{cases} \tag{134}$$

Then the power delay profile can be derived by calculating (133) in a continuous time slot.

4.2.2 Multimode Model in the Room-and-pillar Environment

As discussed in the beginning of Section III, simplified multimode model combined with shadow fading model is implemented to characterize the wireless channel in room-and-pillar environment.

4.2.2.1 The Simplified Multimode Model

Because the room of the room-and-pillar channel in underground mines is usually very large, the influence of the reflection on the vertical walls is very limited. However, the reflection on the ceiling and floor cannot be omitted. Hence, the room without pillars is modeled as a planar air waveguide. It can be viewed as a simplified rectangular waveguide with dependence on only one coordinate. Hence, we use the same procedure as in the tunnel case to develop the multimode model in room-and-pillar environment.

First, we utilize the GO model to analyze the excitation area. Because the planar air waveguide has dependence on only one coordinate, the excitation plane is degenerated to a line that is perpendicular to the ceiling and floor plane and contains the point of the transmission antenna. The geometry of the cross section is just the same as that of tunnels but with only y-coordinate. The properties of the images and the reflection rays in the tunnel case is still valid. The difference lies in: 1) only y-coordinate takes effect; and 2) the incident angle on the ceiling and floor is a constant – 0°, hence the reflection coefficient is $(1 - \sqrt{k_h})/(1 + \sqrt{k_h})$ for X-polarized field and $(\sqrt{k_h} - 1)/(\sqrt{k_h} + 1)$ for Y-polarized field. In the following derivation, we assume the transmission antenna is X-polarized. The result for Y-polarized antenna can be derived in the similar way. Consider that the transmitter is

located at the height y_0, and the observation point is set at the height y. The major field at the observation point is given by:

$$E^{Rx} = E_0 \cdot \sum_q \left[\frac{\exp(-jky_q(y))}{y_q(y)} \right] \cdot \left(\frac{1 - \sqrt{k_h}}{1 + \sqrt{k_h}} \right)^{|q|} \tag{135}$$

where $y_q(y)$ is the distance between image I_q and the receiver, which is given by:

$$y_q(y) = \begin{cases} |2qb - y_0 - y|, & \text{if } q \text{ is odd} \\ |2qb + y_0 - y|, & \text{if } q \text{ is even} \end{cases} \tag{136}$$

Second, we express the field on the excitation line obtained above into the weighted sum of planar air waveguide modes, and then derive the mode intensity. The eigenfunctions of X-polarized modes in planar air waveguide is given by [70]:

$$\begin{aligned} E_n^x(y) &= E_0 \cdot \cos\left[\left(\frac{n\pi}{2b} - j \cdot \frac{n\pi}{2b^2 k} \frac{k_h}{\sqrt{k_h - 1}} \right) y + \varphi_y \right] \\ &\simeq E_0 \cdot \cos\left(\frac{n\pi}{2b} y + \varphi_y \right) \end{aligned} \tag{137}$$

where $\varphi_y = \frac{\pi}{2}$ if n is even; $\varphi_y = 0$ if n is odd.

The mode intensity C_n is derived by converting the ray sum in (135) into mode sum using the Poisson sum formula. By using the same saddle point method as in the tunnel case, the mode intensity C_n is:

$$C_n(z) = \frac{E_0 \pi}{bz \sqrt{1 - (\frac{n\pi}{2bk})^2}} \cdot \cos\left(\frac{n\pi}{2b} y_0 + \varphi_y \right) \tag{138}$$

Note the intensity C_n is now a function of the distance z. With the intensity and eigenfunction of each mode, the field at any position can be predicted for the case without pillars.

4.2.2.2 Shadow Fading Model and the Combined Result

The pillars in the room-and-pillar mining area are randomly distributed and have random shapes. Signals may experience many reflection and diffraction on those pillars before reaching the receiver. It is very similar to the terrestrial metropolitan area with many buildings. Hence, the shadow fading model can be used to describe the signal's slow fading

caused by the reflection and diffraction on those pillars. The amplitude change caused by shadow fading is often modeled using a log-normal distribution [82]. Since one mode can be viewed as a cluster of rays with the same grasping angle, we assume that each mode experiences identically distributed and independent shadow fading when it goes through the pillars. Therefore, the predicted field at any position ($b + y$ m above the floor, z m apart the transmitter) can be obtained by summing up the field of all modes, which is given by:

$$E^{Rx}(y,z) = E_0 \cdot \sum_n C_n(z) \cdot E_n^x(y) \cdot e^{-(\alpha_n + j\beta_n)\cdot z} \cdot \chi_n \tag{139}$$

where $\{\chi_n\}$ are identically distributed and independent log-normal random variables; the field is divided by $2\pi z$ because the plane wave in the room-and-pillar environment spreads in all horizontal directions; α_n is the attenuation coefficient and β_n is the phase-shift coefficient, which is given by [29, 70]:

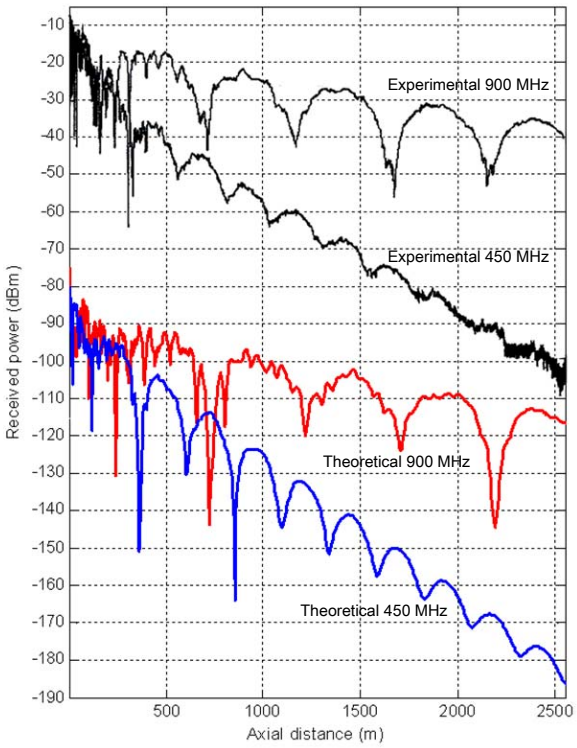
$$\alpha_n = \frac{1}{b}\left(\frac{n\pi}{2bk}\right)^2 Re\frac{1}{\sqrt{\overline{k_h} - 1}}; \quad \beta_n = \sqrt{k^2 - \left(\frac{n\pi}{2b}\right)^2} \tag{140}$$

In the room-and-pillar environment, since the shape, number and position of the pillars are random and vary from case to case, it is not possible to derive a general analytical solution to calculate the power delay profile. Consequently, to characterize the signal distortion of wideband signals in the room-and-pillar environments, field experiments are needed to measure the power delay profile in such mining areas.

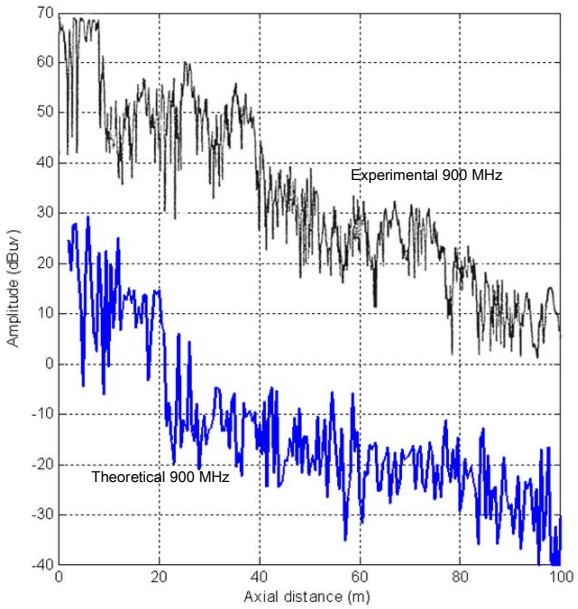
4.2.3 Comparison with Experimental Measurements

To validate the multimode model, we compare our theoretically predicted received power with the experimental measurements in both tunnel and room-and-pillar environments provided in [33] and [56]. Additionally, we also compare our calculated power delay profile with the experimental measurements in a tunnel shown in [41].

In [33], the experiments were conducted in a concrete road tunnel. The tunnel is 3.5 km long and has an equivalent rectangle (7.8 m wide and 5.3 m high) cross section shape. The transmitting and receiving antennas are vertical polarized dipoles at the same height



(a) Received power of 450 *MHz* and 900 *MHz* signals in a road tunnel (the theoretical result is displaced 75 dB downward).



(b) Received power of 900*MHz* signal in a room-and-pillar mining area (the theoretical one is displaced 40 dB downward).

Figure 40: Experimental and theoretical received power

(2 m). Both antennas are placed at the same horizontal position of one-quarter of the tunnel width. Using the same parameters stated above, we calculated the received power by the multimode model. In Fig. 40(a), the calculated results at the frequency of 450 MHz and 900 MHz are compared with the measurements shown in [32, Fig. 18]. The theoretical curves are vertically displaced from the experimental curves for better comparison. It is shown that the curves of the theoretical and experimental results are close to each other. Our multimode model accurately predicts the attenuation velocity, the fast fading in the near region, the flat fading in the far region and the effects of different operating frequency in the tunnel environment.

In [56], the experiments were conducted in a room-and-pillar mining area with an average height of 6 m. The ceiling and floor are made of rocks and the typical values of electrical parameters are $\varepsilon_h = 10\varepsilon_0$ and $\sigma_h = 0.01\ S/m$. The air in this mining area has the same electrical parameters as the atmosphere. The transmitting and receiving antennas are vertical polarized dipoles placed at the same height (2 m). In Fig. 40(b), the calculated results in the frequency of 900 MHz are compared with the measurements shown in [55, Fig. 8 (b)]. As it can be seen, the theoretical results have a good agreement with the experimental measurements in the room-and-pillar environment in underground mines.

In [41], wide-band propagation measurements were performed in a rectangular concrete subway tunnel that is 3.43 m wide, 2.6 m high and 258.7 m long. The signal has a 400 MHz bandwidth at the central frequencies of 900 MHz. The noise floor is 84 dBm, which yields a threshold of 74 dBm for an input 10 dB SNR. The transceiver antennas are horizontally polarized at the center of the tunnel and 50 m apart. In our estimation, we use 900 MHz carrier to modulate a 10 ns wide raised-cosine pulse, which has the same signal bandwidth and central frequency as the experiments. In Fig. 4.2.3, the calculated power delay profile is compared with the measurements shown in [40, Fig. 8 (a)]. It indicates that the theoretical results have a good match with the measurements in signal shape, delay spread and signal strength.

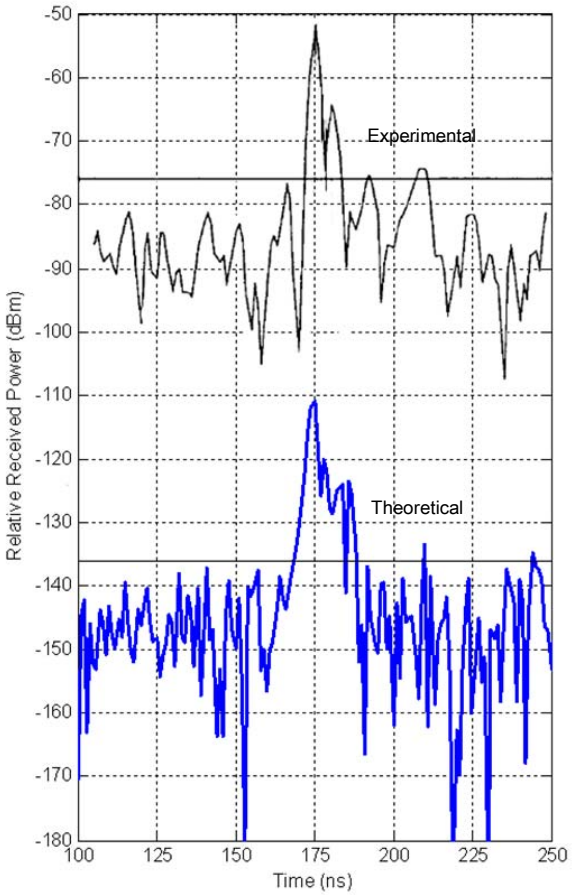


Figure 41: Power delay profile in a Tunnel. The signal bandwidth is 400 MHz and the central frequency is 900 MHz. Transceivers are 50 m apart. The theoretical result is displaced 60 dB downward.

4.2.4 Numerical Evaluation

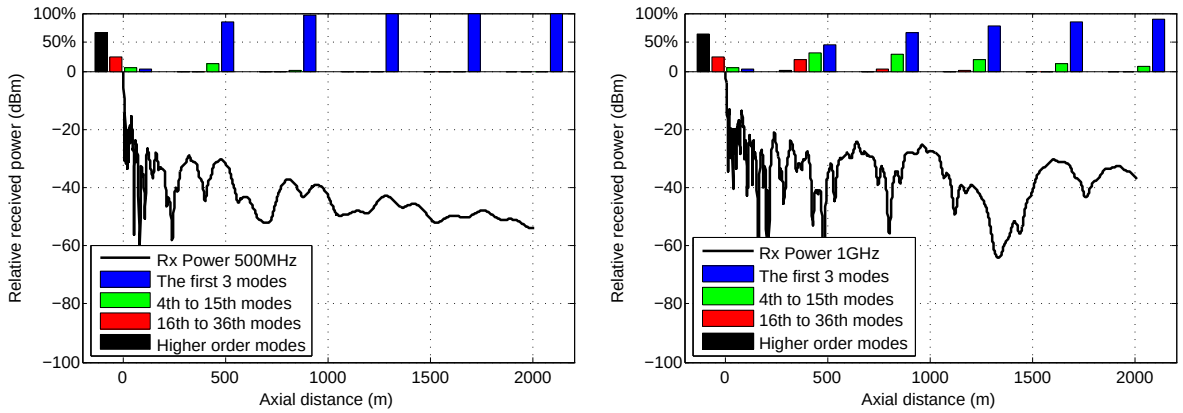
In this section, we first implement the multimode model to analyze the path loss and delay spread under various tunnel conditions. Then we extend our analysis on path loss to the room-and-pillar case.

4.2.4.1 The Tunnel Environment

Except studying the effects of certain parameters, the default values are set as follows: The tunnel cross section shape is a rectangle with a height of 6 m and a width of 10 m; the tunnel wall, ceiling and floor are made of the same material with electrical parameters $\varepsilon = 5\varepsilon_0$, $\sigma = 0.01\ S/m$; the tunnel interior is filled with air ($\varepsilon = \varepsilon_0$, $\sigma = 0\ S/m$). The operating frequency (carrier frequency) is set to 1 GHz. The wideband signal is a 20 ns wide raised-cosine pulse, which has a bandwidth of 200 MHz. The noise floor for the wideband signal is set to 90 dBm according to the strategy in [41]. It yields a threshold of 80 dBm for an input 10 dB SNR. The transmitting power is assumed to be 0 dBm. The transmitting and receiving antennas are horizontal polarized dipoles at the same height (one-third of the tunnel height). Both antennas are placed at the same horizontal position of one-quarter of the tunnel width.

The Operating Frequency:

In Fig. 42 we illustrate the effects of operating frequency on the path loss in tunnels. Specifically, in Fig. 42(a) and Fig. 42(b) we give the signal power and the corresponding power distribution among significant modes as a function of axial distance at the frequency of 500 *MHz* and 1.0 *GHz*. The curve of the signal power can be divided into two regions. In the near region, the power attenuates fast and fluctuates very rapidly. This is attributed to the combined effect of multiple modes. While in the far region, the fall in the signal power is gradual. This is due to the fact that the higher order modes attenuate very fast as the distance increases. Hence, the field in the far region is governed by the few low-order modes left. The relationship between power distribution among modes and the received



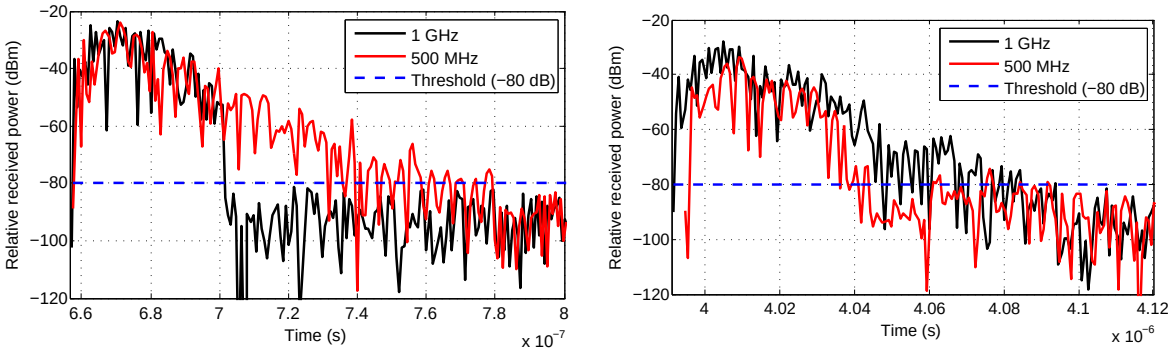
(a) Received signal power and the power distribution among modes at 500 MHz.

(b) Received signal power and the power distribution among modes at 1 GHz.

Figure 42: Received power in tunnels at different operating frequencies.

power is clearly shown in Fig. 42(a) and Fig. 42(b). Although the operation frequency does not affect the power distribution of modes significantly, it has an obvious influence on the propagation constants. Signals with higher frequency attenuate slower. Thus, as the frequency increases, the signal attenuation decreases and the fast fluctuating region is prolonged, as shown in Fig. 42(b).

In Fig. 43 we analyze the effect of operating frequency and transmission distance on the power delay profile in tunnels. It is shown that the raised-cosine pulse is widened and the pulse shape is distorted after propagation in the tunnel, which is caused by the following two reasons: 1) each frequency element and the each mode of the original signal has different propagation delay; and 2) the attenuation and phase-shift rates of different frequency elements and modes are also different. As shown in Fig. 43(a), the delay spreads of the 500 MHz signal is larger than that of the 1 GHz signal when the transmission distance is 200 m. This is because that: 1) there are multiple significant modes for both frequency in the near region; 2) for the signal with lower operating frequency, the differences of group velocity between the frequency elements are larger, which causes larger delay spread. As the transmission distance increases to 1200 m, only lower order modes of the 500 MHz signal are left due to its high attenuation rate. Meanwhile the 1 GHz signal still has several



(a) Power delay profile at 200 m.

(b) Power delay profile at 1200 m.

Figure 43: Power delay profile in tunnels at different operating frequencies.

significant modes at 1200 m. Hence the delay spread of the 1 GHz signal is larger in the far region. For the same signal with different transmission distance, it is observed that the delay spread increases as the distance increases, which is because that the difference of propagation delays among the modes and frequency elements increases as the transmission distance increases. However, after a certain distance, as higher order modes disappear in the far region, fewer modes are left and the delay spread decreases. Therefore, the delay spread is a function of distance, which is an increasing function at first and become a decreasing function after a turning point.

The Tunnel Size:

The tunnel size has similar effects on the path loss and the delay spread as the operating frequency. In larger dimension tunnels, the attenuation constant (α_{mn} in eq (112)) is smaller. Thus, more mode remains significant in far regions. Therefore, the speed of the signal attenuation decreases and the fast fluctuating region is prolonged in the larger tunnels. Moreover, since there are more significant modes in the larger dimension tunnel, the delay spread of the signal in such tunnel is also larger, as shown in Fig. 44. For horizontal polarized antennas, the tunnel width plays a more important role because the reflection coefficients on the horizontal ceiling and floor are larger than those on the vertical walls. Similarly, the tunnel height weights more for vertical polarized antenna.

The Antenna Position and Polarization:

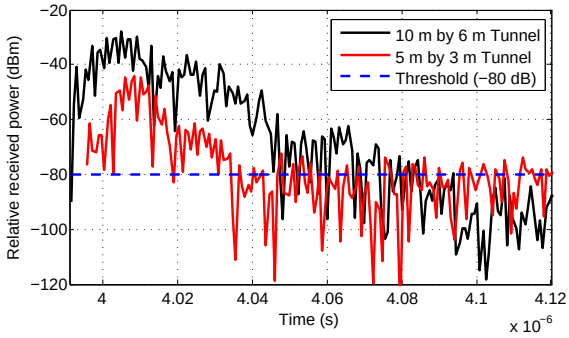
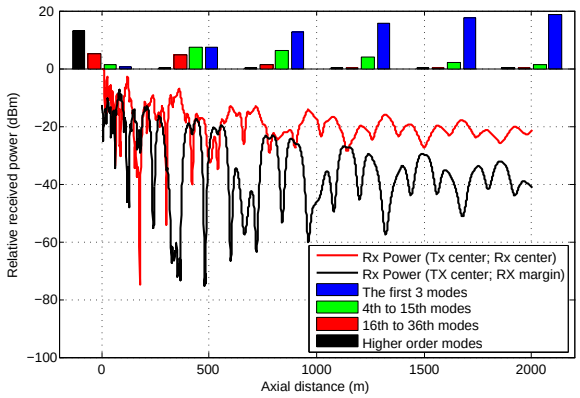
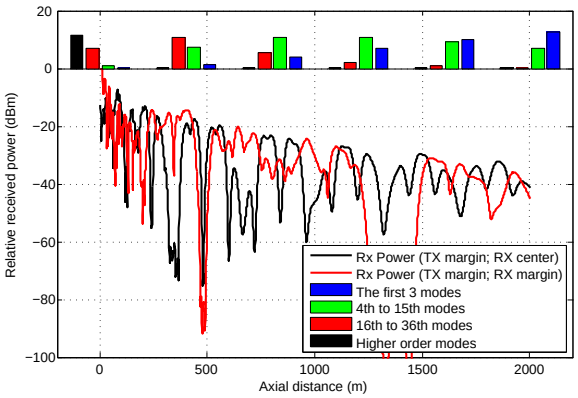


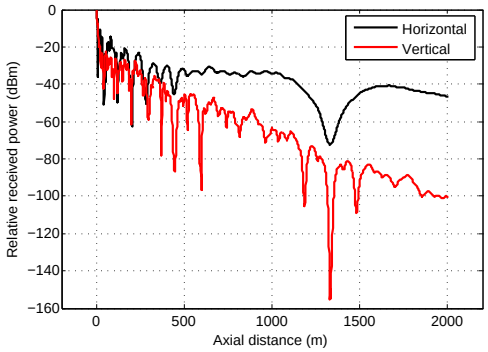
Figure 44: Power delay profile in tunnels with different tunnel sizes at 1200 m.



(a) Received signal power and the power distribution among modes when Tx antenna is placed near the tunnel center.

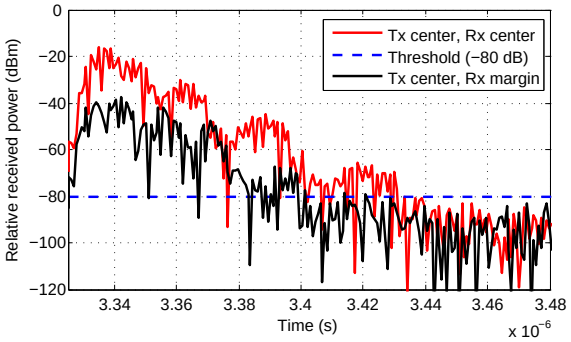


(b) Received signal power and the power distribution among modes when Tx antenna is placed near the tunnel wall.

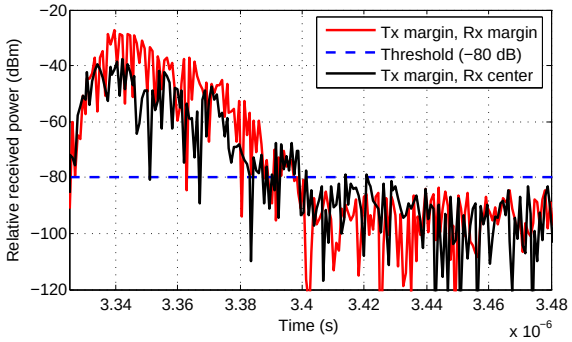


(c) Received signal power of different antenna polarization.

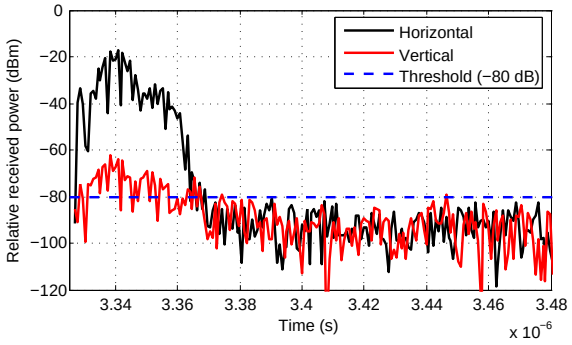
Figure 45: Path loss characteristics in tunnels with different antenna position and polarization.



(a) Power delay profile when Tx antenna is placed near the tunnel center.



(b) Power delay profile when Tx antenna is placed near the tunnel wall.



(c) Power delay profile of different antenna polarization.

Figure 46: Power delay profile in tunnels with different antenna position and polarization at 1000 m.

As discussed above, the mode attenuation is mostly determined by the tunnel size and operating frequency. The mode intensity is to a large extent governed by the position of the transmitter antenna. In Fig. 45 we give the received power and power distribution among modes with different antenna positions and polarizations. In Fig. 45(a) we show the case that the transmitter antenna is placed near the center of the tunnel cross section. The receiver antenna is placed either at the center or at the marginal position that is one-eighth of the tunnel height and one-eighth of the width. It is shown that the lowest modes are effectively excited. If the receiver is also at the center, both the signal attenuation and the fluctuation are small. If the receiver is placed near the tunnel walls, the attenuation and fluctuation are much more significant. In Fig. 45(b) we show the case that the transmitter antenna is placed near the tunnel wall (1/8 of the width and 1/8 of the height). The receiver antenna is also placed either at the center or at the margin of the tunnel. Near the excitation plane, the high order modes play the dominant role. In this case, the position of receiver antenna does not affect the received signal as much as the former case. The attenuation and fluctuation of received power is significant, no matter where the receiver is placed.

Besides the antenna position, we analyze different antenna polarization in Fig 45(c). To make the effects more obvious, we choose a wide but low tunnel ($10 \times 3\ m^2$) here. We show that the signal excited from a horizontal polarized antenna attenuates much slower than that excited from a vertical polarized one. It is consistent with the previous discussion about the relationship between the tunnel size and antenna polarization. Hence, it can be pointed out that: in wide but low tunnel, the horizontal polarized antenna is more appropriate while for narrow but high tunnel, the vertical polarized antenna is more suitable.

In Fig. 46, we investigate the effect of antenna position and polarization on the power delay profile in tunnels. In particular, we show the case that the transmitter antenna is placed near the center of the tunnel cross section in Fig. 46(a). The case that the transmitter antenna is placed near the tunnel wall is shown in Fig. 46(b). The comparison between the power delay profiles of different polarizations is given in Fig. 46(c). It is shown that the

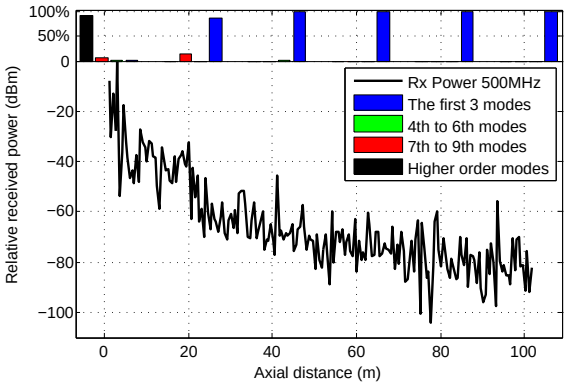
delay spread is determined by the power of existing modes and the noise level. If more modes have much higher power level than the noise, the delay spread is relatively larger. Otherwise, although the pulse is widened, most of the pulse is submerged in the noise and only a small portion of the pulse can be observed. Hence, for different antenna positions, the delay spread of the case when both transceivers are placed in the tunnel center is much larger than other cases. For different antenna polarizations, the delay spread of horizontal polarized antenna is larger than the vertical polarized one in wide tunnels where the width is larger than the height.

The Electrical Parameters:

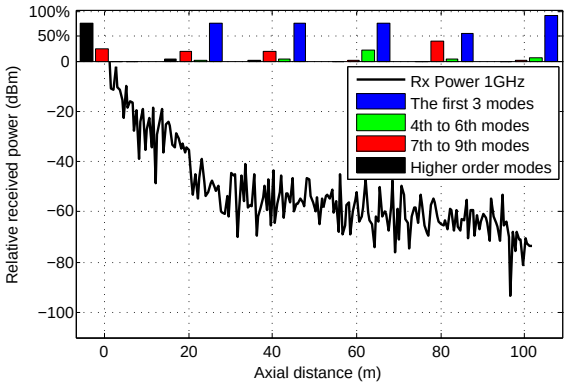
The electrical parameters consist of permittivity ε and conductivity σ. The temperature, humidity and pressure have little influence on the air permittivity but may affect the conductivity more. However, the effect of different conductivity of tunnel air may be neglectable, because it is very small compared to the permittivity. Therefore, the electrical parameters of tunnel air can be considered the same as those of atmosphere air. Tunnel walls' electrical parameters can be looked up in [29], where the permittivity of tunnel materials are in the range of $5\varepsilon_0 \sim 10\varepsilon_0$ and the conductivity is on the order of $10^{-2}\ S/m$ at the UHF frequency band. In this value range, the received power curves with different wall electrical parameters are very close to each other. Hence it can be concluded that the electrical parameters of either tunnel wall or tunnel air do not considerably influence the signal propagation inside the tunnel.

4.2.4.2 The Room-and-Pillar Environment

We conduct similar simulations in the room-and-pillar environment under the condition that: 1) the average height of the room is 6 m; 2) the electrical parameters of the ceiling and floor are $\varepsilon_h = 7\varepsilon_0$ and $\sigma_h = 0.01\ S/m$; 3) the air in this mining area has the same electrical parameters as the atmosphere ($\varepsilon_a = \varepsilon_0$ and $\sigma_h = 0\ S/m$); and 4) the transmitting and receiving antennas are vertical polarized dipoles placed at the same height (2 m). The



(a) Received signal power and the power distribution among modes at 500 MHz.



(b) Received signal power and the power distribution among modes at 1 GHz.

Figure 47: Path loss characteristics in room-and-pillar environments with different operating frequencies.

transmitting power is assumed to be 0 dBm.

We give the received power in dB and the corresponding power distribution among significant modes as a function of axial distance at the frequency of 500 *MHz* and 1.0 *GHz* in Fig. 47(a) and Fig. 47(b), respectively. It can be observed that the signal attenuates faster in the near region of the source than in the far region. Due to the shadow fading and multiple-mode operating, signals experience significant fluctuations in both near and far region. Since higher frequency signal has lower attenuation coefficients, more number of modes remain significant in the far region. However, due to the shadow fading caused by the pillars and the path loss caused by the plane wave spreading, signal propagation with different operating frequency does not have significant differences.

Besides the operating frequency, other factors such as room height, antenna position/polarization and electrical parameters in the room-and-pillar environment affect signal propagation in a similar way as in the tunnel case. However, their influence is much smaller, which can be explained as follows. Compared to the tunnel case, signals in the room-and-pillar mining area experience extra multipath fading caused by the pillars. Moreover, higher path loss is experienced by the wave spreading in the room.

4.3 Influence of Vehicular Traffic Flow

4.3.1 Channel Model in Empty Tunnels

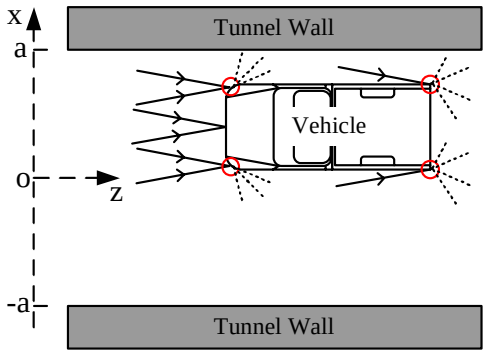
Let $\mathbf{N}_{mode}$ be the set of significant modes in the tunnel. $\mathbf{N}_{mode} = \{(m, n)\}$ and $|\mathbf{N}_{mode}| = N$. Since the mode's attenuation rate rises fast as the mode order increases, only a small number of modes are left after the signal propagates for a certain distance. The number of significant modes that need to be considered is determined by multiple parameters, including the size of the tunnel, the distance between the transceivers, and the operating frequency. In most cases, the received signal strength is convergent when the first 50 ($m + n < 10$) modes are considered. Based on the channel model in empty tunnel given in Section 4.2, the electric field $E^{RX}(x_r, y_r, z_r)$ at the position of the receiver can be obtained by by

$$E^{RX}(x_r, y_r, z_r) = \sqrt{G_t G_r} \cdot \sum_{(m,n)\in\mathbf{N}_{mode}} E^{eign}_{mn,(x_r,y_r)} \cdot e^{-\Gamma_{mn}\cdot z_r} \cdot C^{TX}_{mn}, \tag{141}$$

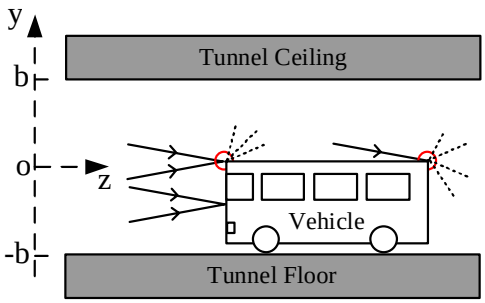
where z_r is the distance between the transmitter and the receiver; (x_r, y_r) is the coordinate of the receiver at the tunnel cross section; G_t and G_r are the TX and RX antenna gain, respectively; $E^{eign}_{mn,(x_r,y_r)}$ is the value of the eigenfunction of the mode EH_{mn} at the position of the receiver; $\Gamma_{mn} = \alpha_{mn} + j\beta_{mn}$ where α_{mn} is the attenuation coefficient of the mode EH_{mn} and βmn is the phase shift coefficient of EH_{mn}; C^{Tx}_{mn} is the intensity of the mode EH_{mn} excited by the transmitter. The expressions of $E^{eign}_{mn,(x,y)}$, α_{mn}, β_{mn}, and C^{Tx}_{mn} are given by (110), (112), and (129), respectively. If the tunnel is curved, the expressions of α_{mn} and β_{mn} should be modified according to [69].

4.3.2 Signal Propagation around a Single Vehicle

The propagation modes can travel along the tunnel without interference with each other if the tunnel is empty. Vehicles in the tunnel may cause additional propagation loss of each mode, which is regarded as *in-mode loss*. Moreover, part of the energy of each mode may be coupled to the other modes due to the existence of the vehicles, which is regarded as



(a) Top view. In-mode loss due to the reflection on the vehicle's front surface. Mode coupling due to the diffraction on the four vertical edges around the vehicle.



(b) Side view. In-mode loss due to the reflection on the vehicle's front surface. Mode coupling due to the diffraction on the two horizontal edges on the top of the vehicle.

Figure 48: Influence of a single vehicle on the mode propagation.

mode coupling. In this section, we analyze the influence of a single vehicle on the mode propagation inside the tunnel.

Without loss of generality, the vehicles are modeled as metal cubes with different sizes. According to US Federal Regulations, the width w, height h and length l of most vehicles (including cars, vans, buses and trucks) fall into the following intervals (unit is meter):

$$w \in [1.5, 2.5], \quad h \in [1.3, 4.2], \quad \text{and} \quad l \in [3.5, 16.2]\,. \tag{142}$$

4.3.2.1 Classification of Vehicle Influence

Assuming that there is a vehicle with size ($w \times h \times l$) inside the tunnel. Fig. 48 shows the top view and the side view of the tunnel with a vehicle. The vehicle is z_v meters apart from the transmitter. The middle axle of the vehicle lies parallel to the z-axes with the position $x = x_v$. In the empty tunnel between the transmitter and the vehicle, the field of the EH_{mn} mode can be calculated according to (179):

$$E_{mn}(x, y, z) = C_{mn}^{Tx} \cdot e^{-\Gamma_{mn} \cdot z} \cdot E_{mn,(x,y)}^{eign}\,, \tag{143}$$

Substituting (110) into (143) yields:

$$
\begin{aligned}
&E_{mn}(x,y,z)\\
&\simeq C_{mn}^{Tx}\cdot e^{-\Gamma_{mn}\cdot z}\cdot\sin\left(\frac{m\pi}{2a}x+\varphi_x\right)\cdot\cos\left(\frac{n\pi}{2b}y+\varphi_y\right)\\
&=C_{mn}^{Tx}\cdot e^{-(\alpha_{mn}+j\beta_{mn})\cdot z}\cdot\frac{1}{2}\left[\rho_1(m)e^{j\frac{m\pi}{2a}x}+\rho_2(m)e^{-j\frac{m\pi}{2a}x}\right]\cdot\frac{1}{2}\left[\rho_2(n)e^{j\frac{n\pi}{2b}y}+\rho_1(n)e^{-j\frac{n\pi}{2b}y}\right]\\
&=\frac{1}{4}C_{mn}^{Tx}\cdot e^{-\alpha_{mn}\cdot z}\cdot\{\rho_1(m)\rho_2(n)e^{ik[\sin(\alpha_m)\cdot x+\sin(\beta_n)\cdot y-\beta_{mn}/k\cdot z]}\\
&\quad+\rho_2(m)\rho_2(n)e^{ik[\sin(-\alpha_m)\cdot x+\sin(\beta_n)\cdot y-\beta_{mn}/k\cdot z]}+\rho_1(m)\rho_1(n)e^{ik[\sin(\alpha_m)\cdot x+\sin(-\beta_n)\cdot y-\beta_{mn}/k\cdot z]}\\
&\quad+\rho_2(m)\rho_1(n)e^{ik[\sin(-\alpha_m)\cdot x+\sin(-\beta_n)\cdot y-\beta_{mn}/k\cdot z]}\},
\end{aligned}
\tag{144}
$$

where $\rho_1(u)=-j$ if u is even; $\rho_1(u)=1$ if u is odd; $\rho_2(u)=j$ if u is even; $\rho_2(u)=1$ if u is odd. It shows that each propagation mode can be viewed as the superposition of four plane waves with four symmetric directions. $\pm\alpha_m$ and $\pm\beta_n$ are the grazing angles of the plane waves on the tunnel walls, which satisfy:

$$
\sin(\alpha_m)=\frac{m\pi}{2ak}\quad\text{and}\quad\sin(\beta_n)=\frac{n\pi}{2bk}\,. \tag{145}
$$

It indicates that the higher order modes have larger grazing angles. Since higher order mode has much higher attenuation rate [84, 86], we pay more attention on the lower order modes. For the lowest order mode EH_{11}, the grazing angle is very small (around 1°).

As illustrated in Fig. 48, the influence of the vehicle on the propagation mode EH_{mn} can be classified into two categories:

- The plane waves hitting the vehicle's front surface will be reflected back. The backward waves will travel in the opposite direction of the receiver hence will not influence the receiver. Therefore the energy of the waves hitting the vehicle's front surface can be viewed as additional loss of the EH_{mn} mode caused by the vehicle. This kind of influence is defined as *in-mode loss*.

- The plane waves hitting the edge between two surfaces of the vehicle will be diffracted. The diffracted waves has different transmission directions, which means that the energy is coupled to other modes. This kind of influence is defined as *mode coupling*.

4.3.2.2 In-Mode Loss

The in-mode loss is first calculated. The eigenfunction of the propagation mode gives the energy distribution pattern in the tunnel cross section of each mode. Based on the eigenfunctions, the energy reflected back by the vehicle's front surface can be obtained. The in-mode loss is the ratio of the energy reflected back to the total energy in the cross section:

$$\begin{aligned} L_{mn} &= \frac{\int_{x_v-\frac{w}{2}}^{x_v+\frac{w}{2}} \int_{-b}^{-b+h} [E^{eign}_{mn,(x,y)}]^2 dxdy}{\int_{-a}^{a}\int_{-b}^{b} [E^{eign}_{mn,(x,y)}]^2 dxdy} \\ &= \frac{1}{4ab}\Big[w - \frac{2a}{m\pi}(-1)^m \cos(\frac{m\pi}{a}x_v)\sin(\frac{m\pi}{a}w)\Big] \cdot \Big[h - \frac{b}{n\pi}\sin(\frac{n\pi}{b}h)\Big]. \end{aligned} \tag{146}$$

It shows that the in-mode loss is a function of the width, height and the position of the vehicle. Actually, in the interval of the tunnel where the vehicle exists, the mode attenuation coefficient Γ_{mn} increases. It is because that the tunnel cross section in this interval is smaller than the empty tunnel. The increased attenuation coefficient causes the additional in-mode loss. However, this part of in-mode loss can be ignored for low order modes, since the attenuation coefficient is small for low order mode and the vehicle length is relatively short (less than 20 m).

4.3.2.3 Mode Coupling

The diffraction on the vehicle edges causes a portion of one mode's energy coupling to other modes. According to Fig. 48, the diffraction occurs on the four vertical edges and the two horizontal edges on the vehicle. Before quantitatively calculating the mode coupling coefficient, the property of the mode diffraction on the edge is first analyzed.

Proposition 1: In a rectangular tunnel, the vertical edge on an obstruction can only cause one mode to be coupled to the modes that have the same eigenfunctions on y-axes (the eigenfunctions on x-axes can be different). In other words, when mode EH_{mn} is diffracted on a vertical edge, it can only be coupled to the modes $\{EH_{tn}|t = 1, 2, \cdots\}$. Similarly,

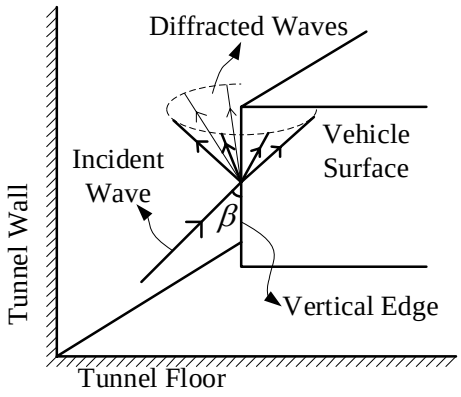


Figure 49: Diffraction of an incident wave on a vertical edge in the tunnel.

when mode EH_{mn} is diffracted on a horizontal edge, it can only be coupled to the modes $\{EH_{ms}|s = 1, 2, \cdots\}$.

Proof: Here we only prove the vertical edge case. The horizontal edge case can be proved in the same way. Consider that the incident wave of mode EH_{mn} hits a vertical edge, as shown in Fig. 49. The angle between the vertical edge and the direction of the incident wave is β, while the angle between the vertical edge and the direction of the diffracted wave is β'. β is the complementary angle of the grazing angle on the tunnel ceiling/floor β_n ($\beta_n + \beta = 90°$). According to the geometrical theory of diffraction (GTD) [17], β' should be equal to β (forming the *Keller cone*). As a result, the grazing angle of the diffracted waves on the tunnel ceiling/floor is β_n. In the meantime, the diffracted keller cone generates an equivalent radiation source only in the x-z plane. Therefore, after the diffraction on the vertical edge, the mode EH_{mn}'s grazing angle on the vertical walls can change to any value, but the grazing angle on the ceiling/floor remains the same. That is, after diffracted on a vertical edge, mode EH_{mn} can only be coupled to the modes $\{EH_{tn}|t = 1, 2, \cdots\}$.#

The proposition 1 simplifies the mode coupling analysis in tunnels to two 2-dimensional problems: the mode coupling in x-z plane for the four vertical edges; and the mode coupling in y-z plane for the two horizontal edges. In the following part, detailed calculations of the mode coupling in the y-z plane are given. The results of the mode coupling in the x-z plane can be derived in the similar way. The total mode coupling effect is the superposition of all

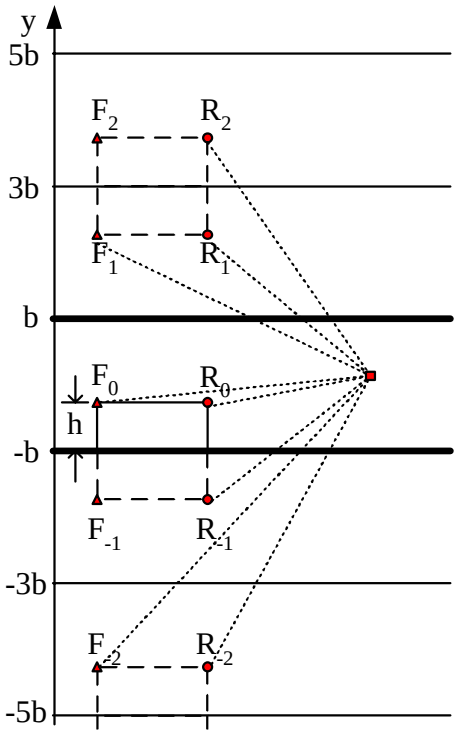


Figure 50: Side view (y-z plane) of the diffractions on the horizontal edges of a vehicle in the tunnel.

the six edges' contributions.

Fig. 50 shows the diffracted rays on the two horizontal edges in the y-z plane. The propagation mode EH_{mn} hits the front upper horizontal edge F_0 and the rear upper horizontal edge R_0. Since the attenuation coefficient of the low order mode is small and the vehicle length is usually several meters, the field attenuation and phase shift from the vehicle front to the rear can be ignored. Hence the fields at the front edge can be considered as the same as the field at the rear edge, which is:

$$\begin{aligned} E^i_{mn} &= C^{Tx}_{mn} \cdot e^{-\Gamma_{mn} \cdot z_v} \cdot E^{eign}_{mn,(x,h-b)} \\ &= C^{Tx}_{mn} \cdot e^{-\Gamma_{mn} \cdot z_v} \cdot \sin\left(\frac{m\pi}{2a}x + \varphi_x\right) \cdot \frac{1}{2}\left[\rho_2(n) e^{ik(h-b)\sin(\beta_n)} + \rho_1(n) e^{ik(h-b)\sin(-\beta_n)}\right] \\ &\triangleq E^i_{mn}(\beta_n) + E^i_{mn}(-\beta_n)\,, \end{aligned} \tag{147}$$

where we view the field as the sum of two rays with grazing angle $\pm\beta_n$ on y-z plane.

Mode Coupling on the Rear Horizontal Edge:

We first consider the diffracted field caused by the rear edge. Note that only the incident rays with direction $-\beta_n$ is considered for the diffractions on the rear edge, which is due to the reason that rays with direction β_n cannot hit the rear edge, as shown in Fig. 48(b). The edge can be viewed as a new radiation source. Due to the reflections on the tunnel ceiling/floor, the new source generates periodically positioned source images, as shown in Fig. 50. The diffracted field caused by the rear edge can be calculated by summing up all the rays coming from all the images, which is:

$$E^D(x, y, z+z_v) = \sum_{p=-\infty}^{\infty} E^i_{mn}(-\beta_n) \cdot D(\alpha_m, \pi \mp \phi_p, \beta_n) \cdot R(\phi_p)^{|p|} \cdot \frac{e^{-jkr_p}}{\sqrt{r_p}}, \tag{148}$$

where the upper sign is used for the case when p is zero, positive odd or negative even; the lower sign is for other cases; r_p is the distance from image R_p in Fig. 50:

$$r_p = \sqrt{z^2 + [2pb \pm (h-b) - y]^2} \quad \begin{cases} + \text{ if } p \text{ is even} \\ - \text{ if } p \text{ is odd} \end{cases}; \tag{149}$$

ϕ_p is the angle of the grazing angle on the tunnel celing/floor of the rays coming from image R_p; it is also the diffracted angle on the y-z plane:

$$\phi_p = \tan^{-1}\left[\frac{|2pb \pm (h-b) - y|}{z}\right] \quad \begin{cases} + \text{ if } p \text{ is even} \\ - \text{ if } p \text{ is odd} \end{cases}; \tag{150}$$

$R(\phi_p)$ is the reflection coefficient that can be simplified as an exponential function [84, 86]:

$$R(\phi_p) \simeq -\exp\left(-\frac{2\sin\phi_p}{\sqrt{\varepsilon_w}}\right); \tag{151}$$

$D(\theta_1, \theta_2, \theta_3)$ is the diffraction coefficient on an edge, where $\frac{\pi}{2} - \theta_1$ is the angle between the incident ray and the edge; θ_2 is the angle of the diffracted ray on the plane perpendicular to the edge; and θ_3 is the angle of the incident ray on the plane perpendicular to the edge. Since the GTD method cannot characterize the diffraction near the shadow boundary and the reflection boundary, we adopt unified theory of diffraction (UTD) to calculate the coefficients, which is given in [20, fomula (25)].

To calculate the energy coupled to the other modes, the Poisson summation formula [109] is introduced to translate the sum of rays in (148) to the sum of modes. The ray sum in (148) can be sorted into two parts: the sum of all the functions that the subscript p is even and the sum of the functions that the subscript p is odd:

$$E^D(x,y,z+z_v) = \sum_{q=-\infty}^{\infty} f(4qb+h-b-y) + \sum_{q=-\infty}^{\infty} f(4qd+3b-h-y)\,. \tag{152}$$

Each sum in (152) is a function with period $4b$. According to the Poisson summation formula, the sum of even subscript can be transformed to

$$\sum_{q=-\infty}^{\infty} f(4qb+h-b-y) = \frac{1}{4b} \sum_{s=-\infty}^{\infty} \mathscr{F}_e(s) \cdot e^{j\frac{s\pi}{2b}y}\,, \tag{153}$$

where the coefficient $\mathscr{F}_e(s)$ is the Fourier transform of the function $f(h-b-y)$ with even subscript:

$$\begin{aligned}\mathscr{F}_e(s) &= \int_{-\infty}^{\infty} f(h-b-y) \cdot e^{-j\frac{s\pi}{2b}y} dy \\ &= E^i_m(-\beta_n) \cdot \int_{-\infty}^{\infty} D(\alpha_m, \pi \mp \phi_0, \beta_n) \cdot (-1)^{p(y)} \cdot \exp\left(\frac{-2p(y)\sin\phi_0}{\sqrt{\varepsilon_w}}\right) \cdot \frac{e^{-jk\sqrt{z^2+(h-b-y)^2}}}{[z^2+(h-b-y)^2]^{\frac{1}{4}}} \cdot e^{-j\frac{s\pi}{2b}y}\, dy,\end{aligned} \tag{154}$$

where $\phi_0 = \tan^{-1}\left[\frac{|h-b-y|}{z}\right]$. Note that a continuous function $p(y) = |y-h+b|/2b$ is used to approximate the discrete number of reflections on the tunnel ceiling/floor. The motivation for the approximation is to fit the ray sum in (148) to the Poisson sum formula. The saddle-point method [28] is deployed to evaluate the integration in (154). The saddle-point y_{sp} is:

$$y_{sp} \simeq |z| \cdot \tan\beta_s\,, \tag{155}$$

where β_s is the grazing angle of the mode EH_{ms} on the tunnel ceiling/floor that can be calculated by (145). It is assumed that the order number s is small and the permittivity of the tunnel wall is large. Then, using the saddle point method, we can obtain the approximate result of the integration in (154):

$$\begin{aligned}\mathscr{F}_e(s) \simeq \frac{1}{2} E^i_m(-\beta_n) \cdot [D(\alpha_m, \pi-|\beta_s|, \beta_n) + D(\alpha_m, \pi+|\beta_s|, \beta_n)] \\ \cdot \mathrm{sgn}(s) \cdot \frac{\sqrt{2\pi k}}{k\cos(\beta_s)} \cdot e^{-jk\cos(\beta_s)z} \cdot e^{-j\frac{s\pi}{2b}(b-h)}\,.\end{aligned} \tag{156}$$

The coefficient $\mathscr{F}_o(s)$ for the sum of odd subscript in (152) can be calculated in the same way. Hence the total field right behind the vehicle ($z \simeq 0$) can be expressed as

$$E^D(x,y,0+z_v) = \frac{1}{4b}\sum_{s=-\infty}^{\infty}[\mathscr{F}_e(s)+\mathscr{F}_o(s)]\cdot e^{j\frac{s\pi}{2b}y} \tag{157}$$
$$= \frac{1}{4b}E^i_{mn}(-\beta_n)\sum_{s=1}^{\infty}[D(\alpha_m,\pi-\beta_s,\beta_n)+D(\alpha_m,\pi+\beta_s,\beta_n)]$$
$$\cdot\frac{\sqrt{2\pi k}}{k\cos(\beta_s)}\cdot\cos\left(\frac{s\pi}{2b}h+\varphi_y\right)\cdot\cos\left(\frac{s\pi}{2b}y+\varphi_y\right),$$

where $\varphi_y = 0$ if s is odd; $\varphi_y = \frac{\pi}{2}$ if s is even. Note that the second line in (157) is exact the sum of the propagation modes. The total contribution caused by the whole edge can be calculated by integrate (157) along the horizontal edge. Therefore we can derive the coupling coefficient $B^{r,h}_{mn\to ms}$ from mode EH_{mn} to mode EH_{ms} caused by the rear horizontal edge:

$$B^{r,h}_{mn\to ms} = \frac{a\sqrt{2\pi k}}{2m\pi kb\cos(\beta_s)}\cdot\rho_1(n)\cdot e^{ik(h-b)\sin(-\beta_n)}\cdot[D(\alpha_m,\pi-\beta_s,\beta_n)+D(\alpha_m,\pi+\beta_s,\beta_n)]$$
$$\cdot\sin(\frac{m\pi x_v}{2a}+\varphi_x)\sin(\frac{m\pi w}{4a})\cdot\cos(\frac{s\pi}{2b}h+\varphi_y), \tag{158}$$

where $\varphi_x = 0$ if m is even; $\varphi_x = \frac{\pi}{2}$ if m is odd.

Mode Coupling on the Front Horizontal Edge:

The mode coupling coefficient for the front horizontal edge can be calculated in the similar way as the rear horizontal edge. As shown in Fig. 50, the only two differences are: 1) the incident waves have both the directions $\pm\beta_n$; and 2) due to the shadow effect of the vehicle itself, not all the images can illuminate the rest of the tunnel; Only the original source, the images with positive odd subscripts and the images with negative subscripts take effects. As a result, the coupling coefficient $B^{f,h}_{mn\to ms}$ from EH_{mn} to EH_{ms} caused by the front horizontal edge can be expressed as

$$B^{f,h}_{mn\to ms} = \frac{a\sqrt{2\pi k}}{2m\pi kb\cos(\beta_s)}\cdot\sin(\frac{m\pi x_v}{2a}+\varphi_x)\sin(\frac{m\pi w}{4a})\cdot\cos(\frac{s\pi}{2b}h+\varphi_y) \tag{159}$$
$$\cdot\left\{\rho_1(n)\cdot e^{ik(h-b)\sin(-\beta_n)}\cdot D(\alpha_m,\pi-\beta_s,\beta_n)+\rho_2(n)\cdot e^{ik(h-b)\sin(\beta_n)}\cdot D(\alpha_m,\pi-\beta_s,-\beta_n)\right\}.$$

Total Mode Coupling:

The mode coupling coefficient for the four vertical edges can be derived in the same way as the horizontal edges. We only need to interchange the x-axes data and the y-axes data. Therefore the total coupling coefficient $B_{mn\to ms}^{total}$ from EH_{mn} to EH_{ms} caused by the vehicle is the sum of the coefficients of the two horizontal edges:

$$B_{mn\to ms}^{total} = \frac{a\sqrt{2\pi k}}{2m\pi kb\cos(\beta_s)}\cdot\sin(\frac{m\pi x_v}{2a}+\varphi_x)\sin(\frac{m\pi w}{4a})\cdot\cos(\frac{s\pi}{2b}h+\varphi_y) \tag{160}$$
$$\cdot\left\{\rho_2(n)\cdot e^{ik(h-b)\sin(\beta_n)}\cdot D(\alpha_m,\pi-\beta_s,-\beta_n)+\rho_1(n)\cdot e^{ik(h-b)\sin(-\beta_n)}\cdot\left[2D(\alpha_m,\pi-\beta_s,\beta_n)+D(\alpha_m,\pi+\beta_s,\beta_n)\right]\right\}.$$

The total coupling coefficient $B_{mn\to tn}^{total}$ from EH_{mn} to EH_{tn} caused by the vehicle is the sum of the coefficients of the four vertical edges:

$$B_{mn\to tn}^{total} = (-1)^{\lceil\frac{n}{2}\rceil}\cdot\frac{b\sqrt{2\pi k}}{2n\pi ka\cos(\alpha_t)}\cdot\cos(\frac{n\pi}{2b}h) \tag{161}$$
$$\cdot\left\{\cos(\frac{t\pi}{2a}(x_v+\frac{w}{2}))\cdot\left[\rho_1(m)e^{ik(x_v+\frac{w}{2})\sin(\alpha_m)}\cdot D(\beta_n,\pi-\alpha_t,-\alpha_m)\right.\right.$$
$$\left.+\rho_2(m)e^{ik(x_v+\frac{w}{2})\sin(-\alpha_m)}\cdot\left(2D(\beta_n,\pi-\alpha_t,\alpha_m)+D(\beta_n,\pi+\alpha_t,\alpha_m)\right)\right]$$
$$+\cos(\frac{t\pi}{2a}(x_v-\frac{w}{2}))\cdot\left[\rho_2(m)e^{ik(x_v-\frac{w}{2})\sin(-\alpha_m)}\cdot D(\beta_n,\pi+\alpha_t,\alpha_m)\right.$$
$$\left.\left.+\rho_1(m)e^{ik(x_v-\frac{w}{2})\sin(\alpha_m)}\cdot\left(2D(\beta_n,\pi+\alpha_t,-\alpha_m)+D(\beta_n,\pi-\alpha_t,-\alpha_m)\right)\right]\right\}.$$

4.3.2.4 Analytical Expression of a Single Vehicle's Effect

We construct the influence matrix of a vehicle using the results shown in (183), (160) and (161). If there are N significant modes inside the tunnel, the size of the influence matrix $\mathbf{I}$ is $N\times N$. The elements in the diagonal of the matrix consist of the in-mode loss of each significant mode, which is given by (183). The other elements in the matrix are the mode coupling coefficients. According to proposition 1, mode coupling can only happen between two modes that has the same field distribution either in x-z plane or y-z plane. The coupling coefficients between those modes are given by (160) and (161). The mode coupling coefficients between other modes are zero. Given the size and the position of the vehicle, its influence on the signal propagation can be analytically expressed by using the

influence matrix $\mathbf{I}$, which is:

$$\mathbf{I}=\begin{pmatrix} 1-L_1 & B_{2\to1}^{total} & \cdots & B_{N\to1}^{total} \\ B_{1\to2}^{total} & 1-L_2 & \cdots & B_{N\to2}^{total} \\ \vdots & \vdots & \ddots & \vdots \\ B_{1\to N}^{total} & B_{2\to N}^{total} & \cdots & 1-L_N \end{pmatrix}. \tag{162}$$

Hence, the intensity of all the significant modes behind a vehicle inside the tunnel are:

$$(C_1', C_2', \cdots, C_N')^T = \mathbf{I}\cdot(C_1, C_2, \cdots, C_N)^T \ , \tag{163}$$

where $\{C_1, C_2, \ldots, C_N\}$ are the intensity of all the significant modes in front of the vehicle; and $\{C_1', C_2', \ldots, C_N'\}$ are the intensity behind the vehicle.

4.3.3 Channel Modeling for Underground Tunnels with Deterministic and Random Vehicular Traffic Flows

In the real tunnels, there exist multiple vehicles between the transceivers. In this section, the wireless channels for the L-lane ($L = 1, 2, ...$) road tunnel with deterministic and random vehicular traffic flows are characterized, respectively. In particular, we first develop the *deterministic model* for the case when the geometric information of the vehicular traffic flow (the size and position of each vehicle between the transceivers) are known. Then we extend the channel model to cover the case when the statistical data of the vehicular traffic flow is given but the exact number, sizes and the positions of the vehicles of the traffic flow between the transceivers are not known. This channel model is defined as *statistical model*. It should be noted that the statistical model is suitable for most real world applications. Finally, the theoretical models are validated by comparing the predicted received power with the simulation result provided in [8].

4.3.3.1 Deterministic Model

Given the geometric information of the tunnel and all vehicles, the influence of multiple vehicles on the signal propagation can be derived basing on the results in Section 4.3.1 and

Section 4.3.2. Assuming that there are M vehicles between the transceivers in the tunnel. The tunnel geometry is the same as Section 4.3.1. The position and size of the i^{th} vehicle are (x_v^i, z_i) and (w_i, h_i, l_i), respectively. The transmitter is placed in the tunnel cross section where $z = z_0$ and the receiver is placed at (x_r, y_r, z_r). Since the mode propagation can be expressed using matrix [58, 85], the field intensity at the position of the receiver can be expressed as

$$E^{RX}(x_r, y_r, z_r) = \mathbf{E}^{eign}_{(x_r, y_r)} \cdot \mathbf{D}_M \cdot \mathbf{C}^{TX}, \tag{164}$$

where $\mathbf{E}^{eign}_{(x_r, y_r)}$ is the eigenfunction vector at the position (x_r, y_r):

$$\mathbf{E}^{eign}_{(x_r, y_r)} = \left[E^{eign}_{1,(x_r, y_r)}, E^{eign}_{2,(x_r, y_r)}, \cdots E^{eign}_{N,(x_r, y_r)} \right]; \tag{165}$$

$\mathbf{D}_M$ is the propagation matrix, which is a $N \times N$ matrix that indicates the mode propagation characteristics in the tunnel with M vehicles between the transceivers. $\mathbf{D}_M$ is defined as

$$\mathbf{D}_M = \mathbf{A}(z_r - z_M) \cdot \prod_{i=1}^{M} \left[\mathbf{I}_i \cdot \mathbf{A}(z_i - z_{i-1}) \right], \tag{166}$$

where $\mathbf{A}(z)$ is the mode attenuation matrix for transmitting all the significant modes for z meters in an empty tunnel:

$$\mathbf{A}(z) = \begin{pmatrix} e^{-\Gamma_1 \cdot z} & 0 & \cdots & 0 \\ 0 & e^{-\Gamma_2 \cdot z} & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & e^{-\Gamma_N \cdot z} \end{pmatrix}; \tag{167}$$

$\mathbf{I}_i$ is the influence matrix caused by the i^{th} vehicle that is defined in (162). $\mathbf{C}^{TX}$ is the mode intensity vector excited by the transmitter at z_0:

$$\mathbf{C}^{TX} = \left[C_1^{TX}, C_2^{TX}, \cdots C_N^{TX} \right]^T. \tag{168}$$

4.3.3.2 Statistical Model

In practical applications, the exact sizes and positions of all the vehicles in the tunnel cannot be acquired. Consequently, the propagation matrix $\mathbf{D}_M$ in (185) cannot be accurately calculated. Therefore, instead of using the deterministic vehicular traffic information, we adopt

the traffic flow theory and the vehicle size distribution model to predict the propagation matrix $\mathbf{D}_M$.

Distribution of the Propagation Matrix $\mathbf{D}_M$:

According to (162), (185), and (167), the element D_{uv} in the u^{th} row and the v^{th} column of the propagation matrix $\mathbf{D}_M$ can be calculated as

$$D_{uv} = \sum_{i_1}^{N}\sum_{i_2}^{N}\cdots\sum_{i_{M-1}}^{N}\left\{I_{1,\{i_1,v\}}\cdot e^{-\Gamma_v\cdot(z_1-z_0)}\cdot\prod_{l=2}^{M-1}\left[I_{l,\{i_l,i_{l-1}\}}\cdot e^{-\Gamma_{i_{l-1}}\cdot(z_l-z_{l-1})}\right]\cdot I_{M,\{u,i_{M-1}\}}\cdot e^{-\Gamma_{i_{M-1}}\cdot(z_M-z_{M-1})}\cdot e^{-\Gamma_{i_M}\cdot(z_r-z_M)}\right\}, \tag{169}$$

where $I_{l,\{i_l,i_{l-1}\}}$ is the element in the $i_l{}^{th}$ row and the $i_{l-1}{}^{th}$ column of the influence matrix $\mathbf{I}_l$. Since $I_{l,\{i_l,i_{l-1}\}}$ and Γ_n are complex, D_{uv} is also complex. Hence it can be expressed as

$$D_{uv} = Re(D_{uv}) + j\cdot Im(D_{uv})\,. \tag{170}$$

For clear expression, we use $\Theta_{i_1,i_2,\ldots i_{M-1}}$ to denote the addends in the summation in (169). Therefore,

$$D_{uv} = \sum_{i_1}^{N}\sum_{i_2}^{N}\cdots\sum_{i_{M-1}}^{N}\Theta_{i_1,i_2,\ldots i_{M-1}}\,. \tag{171}$$

Since the tunnel is considered to be long, we assume that the total number of vehicles in the traffic flow is correspondingly large. Moreover, the vehicles have random sizes and positions. Thus, $I_{l,\{i_l,i_{l-1}\}}$ and $(z_l - z_{l-1})$ are random variables. Therefore, $Re(D_{uv})$ and $Im(D_{uv})$ can be viewed as the summation of a sufficiently large number of independent random variables, each with finite mean and variance. According to the central limit theorem and the Lindeberg's condition [15], $Re(D_{uv})$ and $Im(D_{uv})$ are approximately normally distributed.

Mean and Variance of the Propagation Matrix $\mathbf{D}_M$:

To characterize the above normal distributions, in this section, we calculate the mean and variance of $Re(D_{uv})$ according to the traffic flow theory and the vehicle size distribution model. The mean and variance of $Im(D_{uv})$ can be derived in the similar way.

In long road tunnels, there is usually no traffic lights and the vehicular traffic flow has only one direction (separate tunnels for traffic with different directions). Hence the

vehicular traffic flow can be modeled as a Poisson flow [37]. The total number of the vehicles within the distance d is denoted as M. The probability that the number of vehicles $M = m$ is:

$$P(M = m) = \left(\lambda \frac{d}{v}\right)^m \cdot \frac{1}{m!} \cdot e^{-\lambda \frac{d}{v}}, \tag{172}$$

where λ is the average rate of vehicle arrival (vehicles/sec) in the tunnel; and v is the average velocity of the vehicles. The distance between the i^{th} and $(i-1)^{th}$ vehicle $\Delta z = z_i - z_{i-1}$ obeys independent and identical exponential distribution. The probability density function (pdf) is:

$$f(\Delta z) = \lambda \cdot e^{-\lambda \frac{\Delta z}{v}}. \tag{173}$$

Every vehicle runs in one of the L lanes in the tunnel. Hence the x-coordinate of the i^{th} vehicles x_v^i belongs to $\{a(\frac{1+2l}{L} - 1) | l = 0, 1, \ldots L-1\}$. The x-coordinate of the vehicle obeys uniform distribution with the probability $1/L$. The size of the vehicles are also assumed to have the uniform distribution in the size interval defined in (142).

According to the central limit theorem and the Lindeberg's condition [15], the mean and variance of $Re(D_{uv})$ are approximately the summation of the mean and variance of the real part of each addend in (171), respectively. Therefore,

$$E\left[Re(D_{uv})\right] \simeq \sum_{i_1}^{N} \sum_{i_2}^{N} \cdots \sum_{i_{M-1}}^{N} E\left[Re(\Theta_{i_1, i_2, \ldots i_{M-1}})\right] = Re\Big[E\Big(\sum_{i_1}^{N} \sum_{i_2}^{N} \cdots \sum_{i_{M-1}}^{N} \Theta_{i_1, i_2, \ldots i_{M-1}}\Big)\Big]; \tag{174}$$

$$\sigma^2[Re(D_{uv})] \simeq \sum_{i_1}^{N} \sum_{i_2}^{N} \cdots \sum_{i_{M-1}}^{N} \sigma^2\big[Re(\Theta_{i_1, i_2, \ldots i_{M-1}})\big] \tag{175}$$

$$= -E^2\left[Re(D_{uv})\right] + \frac{1}{2} E\Big(\sum_{i_1}^{N} \sum_{i_2}^{N} \cdots \sum_{i_{M-1}}^{N} |\Theta_{i_1, i_2, \ldots i_{M-1}}|^2\Big) + \frac{1}{2} Re\Big[E\Big(\sum_{i_1}^{N} \sum_{i_2}^{N} \cdots \sum_{i_{M-1}}^{N} \Theta^2_{i_1, i_2, \ldots i_{M-1}}\Big)\Big].$$

According to (174), the mean of $Re(D_{uv})$ is the real part of the element in the u^{th} row and the v^{th} column of the mean of the propagation matrix $\mathbf{D}_M$. Since the sizes/positions of all vehicles in the vehicular traffic flow can be viewed as independently and identically distributed, the mean of the propagation matrix $\mathbf{D}_M$ can be calculated as

$$E(\mathbf{D}_M) = E\left[E(\mathbf{D}_M | M = m)\right] = \sum_{m=0}^{\infty} E[\mathbf{A}(\Delta z)] \cdot \Big\{E(\mathbf{I}) \cdot E[\mathbf{A}(\Delta z)]\Big\}^m \cdot P(M = m), \tag{176}$$

where $E(\mathbf{I})$ is the mean of the influence matrix that can be calculated according to (142) and (162); $E[\mathbf{A}(\Delta z)]$ is the mean of the mode attenuation matrix that can be calculated using (167) and (173). Based on the Eigendecomposition method [81], the matrix $E(\mathbf{I}) \cdot E[\mathbf{A}(\Delta z)]$ can be decomposed as

$$E(\mathbf{I}) \cdot E[\mathbf{A}(\Delta z)] = \mathbf{V} \cdot \begin{pmatrix} \Lambda_1 & 0 & \cdots & 0 \\ 0 & \Lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \Lambda_N \end{pmatrix} \cdot \mathbf{V}^{-1}, \tag{177}$$

where Λ_i is the i^{th} eigenvalue of the matrix $E(\mathbf{I}) \cdot E[\mathbf{A}(\Delta z)]$; $\mathbf{V}$ is the square $(N \times N)$ matrix whose columns are the eigenvectors of $E(\mathbf{I}) \cdot E[\mathbf{A}(\Delta z)]$. Then, by substituting (193) and (177) into (176), we derive:

$$\begin{aligned} E(\mathbf{D}_M) = & \sum_{m=0}^{\infty} \left\{ E[\mathbf{A}(\Delta z)] \cdot \mathbf{V} \cdot \begin{pmatrix} \Lambda_1^m & 0 & \cdots & 0 \\ 0 & \Lambda_2^m & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \Lambda_N^m \end{pmatrix} \cdot \mathbf{V}^{-1} \cdot \left(\lambda \frac{d}{v} \right)^m \cdot \frac{1}{m!} \cdot e^{-\lambda \frac{d}{v}} \right\} \\ = & E[\mathbf{A}(\Delta z)] \cdot \mathbf{V} \cdot \begin{pmatrix} e^{\lambda \frac{d}{v}(\Lambda_1 - 1)} & 0 & \cdots & 0 \\ 0 & e^{\lambda \frac{d}{v}(\Lambda_2 - 1)} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & e^{\lambda \frac{d}{v}(\Lambda_N - 1)} \end{pmatrix} \cdot \mathbf{V}^{-1}. \end{aligned} \tag{178}$$

$E[Re(D_{uv})]$ is the real part of the element in the u^{th} row and the v^{th} column of the matrix $E(\mathbf{D}_M)$ given in (178). The variance of $Re(D_{uv})$ (i.e. $\sigma^2[Re(D_{uv})]$) can be calculated in the same way according to (175). Due to limited space, the detail deductions are not elaborated here.

The Gaussian-distributed entries in the propagation matrix $\mathbf{D}_M$ can be completely characterized after the mean and the variance are derived. Then the received power at any positions in the tunnel can be analytically predicted by (180).

4.3.3.3 Comparison with Simulation Results

To validate the deterministic model and the statistical model, we compare our theoretical results with the numerical simulation results provided in [8]. The numerical results in [8] are simulated by the geometrical optics (GO) model. Noted that we use the numerical simulation results instead of the experiment measurements due to the following reasons: 1) it is difficult to conduct experiments in the real road tunnels with running vehicles; 2) most existing experiment results are taken in empty tunnels; 3) although the GO model requires a great amount of detailed input data of the parameters to describe the tunnel environments and the vehicles, which is impossible in practical applications, the accuracy of the GO model has been widely accepted. We choose the simulation results in [8] because the simulation scenario is complicated enough and very similar to the real road tunnel.

In the simulation in [8], the tunnel is 1000 m long with 2 lanes. The tunnel cross section has a equivalent rectangle (10 m wide and 6 m high) shape. The tunnel can be viewed as straight for the first 500 m. Form 500 m to 1000 m, the tunnel turns left and forms a $\frac{1}{4}$ circular arc. The transmitter is vertical polarized dipoles at the height of 4.4 m with 1 W power, which is placed near the entrance of the tunnel. The operating frequency is 945 MHz. The receiver is also vertical polarized and the coordinate in the tunnel cross is (2.6 m, -1.5 m). There are 23 vans (1.8 m wide, 1.8 m high, and 5 m long) uniformly distributed on the left lane and 24 buses (2.2 m wide, 3 m high, and and 10 m long) uniformly distributed on the right lane.

Using the same scenario parameters stated above, we first calculated the received power along the tunnel by the deterministic model. As shown in Fig. 51, the received power calculated by the deterministic model is compared with the simulation results given in [8, Bild 6.7]. It is shown that the results derived by the deterministic model match the simulation results accurately. Then the accuracy of the statistical model is tested. The statistical model does not require the size and position information of each specific vehicle. According to the scenario parameters stated above, we can derive that the vehicular traffic

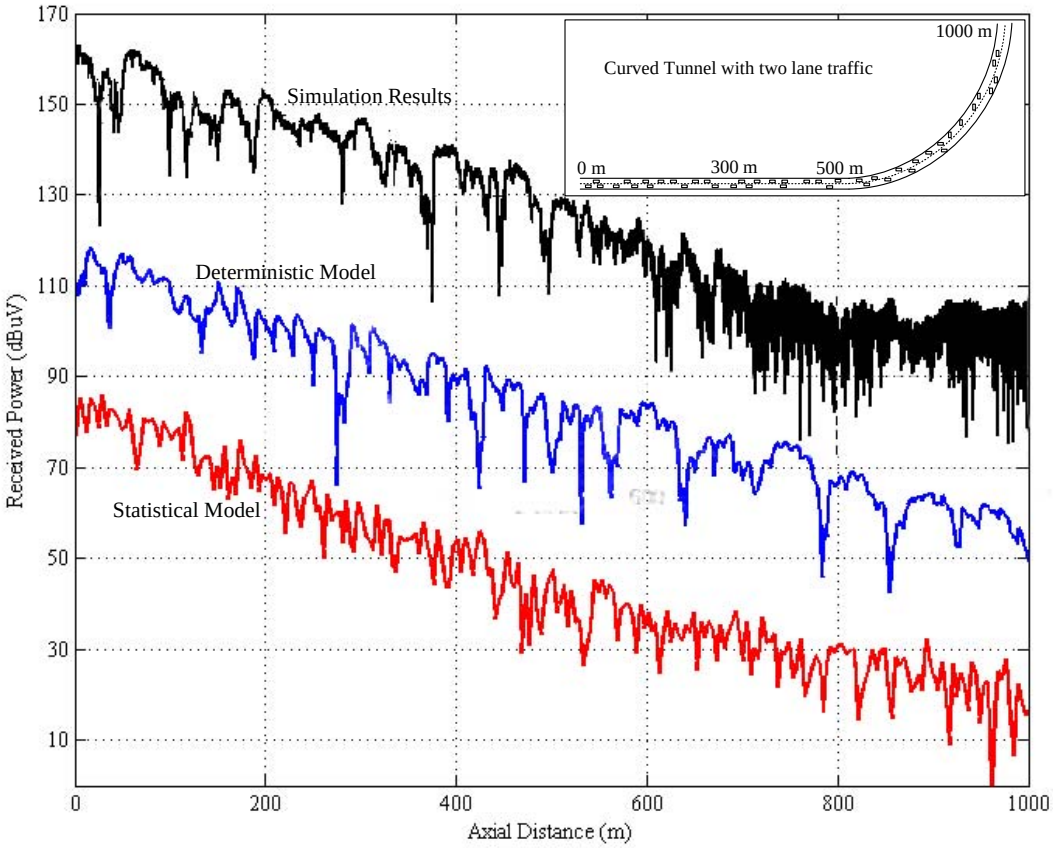


Figure 51: Numerical and theoretical received power in a curved tunnel with traffic (the numerical one is displaced 40 dBuV upward for better comparison).

flow parameter $\frac{\lambda}{v} = 0.047\ \mathrm{m}^{-1}$. Then the received power along the tunnel can be calculated by the statistical model. In Fig. 51, the received power calculated by the statistical model is compared with the simulation results given in [8, Bild 6.7]. The curve of the statistical model accurately match the simulation results, especially when the axial distance is larger than 200 m. The prediction of the statistical model is less accuracy in the very near region, since the condition of the central limit theorem is no longer valid if the distance between the transceivers is too short. It should be noted that In the curved tunnel section, the prediction accuracy of both the deterministic model and the statistical model is not as high as in the straight tunnel section, which is because that the Eigenfunction in the curved tunnel deviates from its original value. However, the error is small and tolerable.

4.3.4 Numerical Anaylsis

In this section, we utilize the proposed channel models to analyze the signal propagations in different tunnels with different vehicular traffic flows. Three types of tunnels are investigated, including a straight one-lane road tunnel, a straight two-lane road tunnel, and a curved two-lane road tunnel. First, we use the deterministic model to analyze the effects of the size, number and the positions of the vehicles on the signal propagation in the tunnels. Then, the statistical model is utilized to give the received power as a function of the transmission distance, vehicular traffic load and vehicle average velocity.

Except studying the effect of certain parameters, default values are set as follows: The tunnel cross section shape is a rectangle with a height of 6 m and a width of 10 m for two-lane tunnels and with a height of 5 m and a width of 6 m for one-lane tunnels; the tunnel wall, ceiling and floor are made of the same material with electrical parameter $\varepsilon_w = 5\varepsilon_0$, $\sigma = 0.01\ S/m$; the tunnel interior is filled with air ($\varepsilon = \varepsilon_0$, $\sigma = 0\ S/m$). The transmitting and receiving antennas are vertically polarized dipoles at the same height (one-third of the tunnel height). Both antennas are placed at the same horizontal position of one-quarter of the tunnel width. The transmitting antenna has the power of 1 W and the impedance of

50 Ω. The operating frequency is 1 GHz.

4.3.4.1 Tunnels with Deterministic Vehicular Traffic Flow

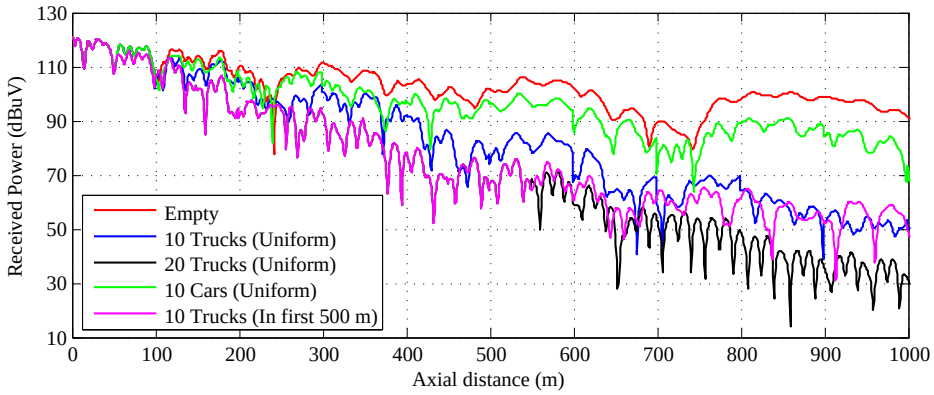
The deterministic model is first analyzed to capture the effects of the size, number and the positions of the vehicles on the signal propagation in the tunnels.

Straight One-Lane Road Tunnel:

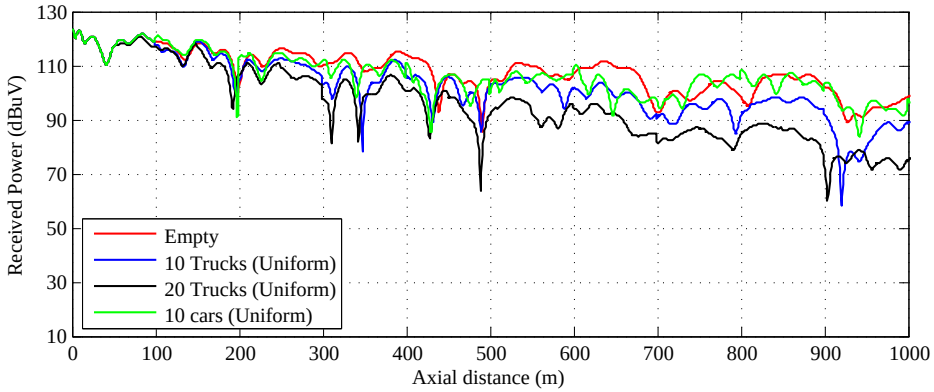
Fig. 52(a) illustrates the effects of the size, number and position of the vehicles on the received power in a straight one-lane tunnel. Compare with the signal propagation in empty tunnels, the existence of the vehicles causes the two impacts: the additional path loss and additional signal fluctuation. These effects can be clearly explained by the theory proposed in Section 4.3.2: 1) the additional path loss is caused by the in-mode loss; and 2) due to the mode coupling, more significant modes appear behind the vehicle; hence the received signal experiences more serious fluctuations. Comparing the received power when there are 10 trucks with the power when there are 20 trucks, we find that the additional loss and fluctuation are approximately proportional to the number of vehicles. The size of the vehicle has significant effects on the path loss, since the influence of 10 cars is much smaller than the influence of the 10 trucks. We compare the received power when 10 truck are uniformly placed in the 1000 m tunnel and the power when all the trucks are placed in the first 500 m of the tunnel. It indicates that the influence of the axial position of the vehicles is not significant.

Straight Two-Lane Road Tunnel:

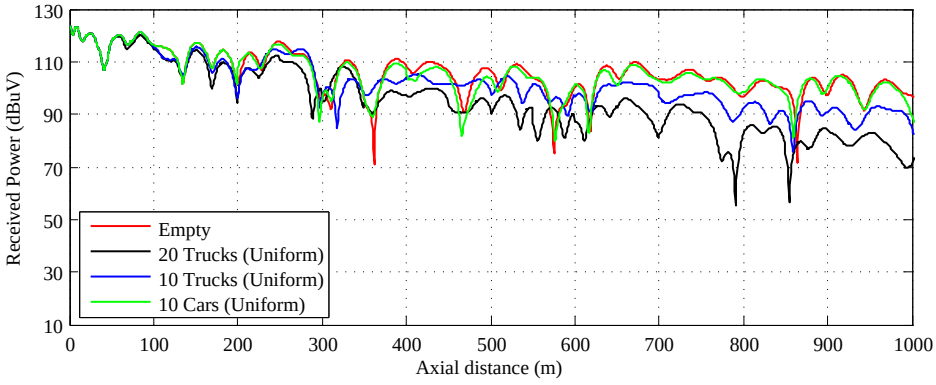
Fig 52(b) shows that the signal propagation in a straight two-lane road tunnel. It shows that the path loss and signal fluctuation are much smaller than the one-lane tunnel. The phenomenon can be explained by the following reasons: 1) the two-lane tunnel has a larger cross section. Hence the signal attenuation of each mode in the empty tunnel is smaller than the one-lane case. 2) the ratio of the vehicle's cross section area to the tunnel's cross section area is much smaller than the one-lane case, hence the in-mode loss and the mode



(a) Received signal power in straight one-lane road tunnel.



(b) Received signal power in straight two-lane road tunnel.



(c) Received signal power in curved two-lane road tunnel.

Figure 52: Signal propagation in road tunnels with determined vehicular traffic flows.

coupling become less significant.

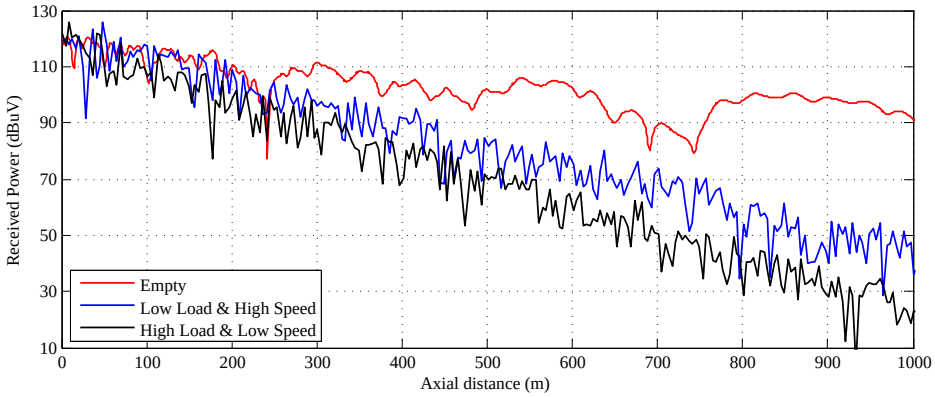
Curved Two-Lane Road Tunnel:

Fig. 52(c) illustrates the signal propagations in a curved tunnel with different types of vehicular traffic flows. The curve radius of the tunnel is 500 m. Since the attenuation rate of the signal excited by a vertically polarized antenna does not change significantly in the curved tunnels [69], the signal propagation characteristics are similar to the straight tunnel case.

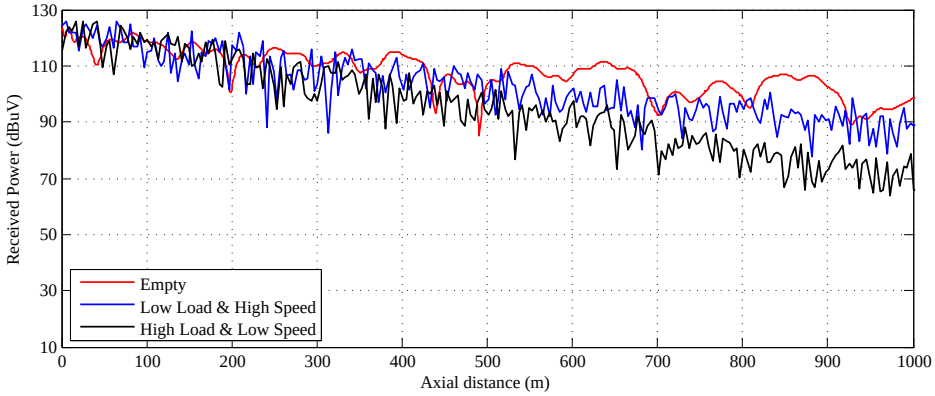
4.3.4.2 Tunnels with Random Vehicular Traffic Flow

In most practical applications, the sizes and positions of all the vehicles in the traffic flow are not deterministic. However, the traffic load and the average vehicle velocity can be estimated in the tunnel. Then, our statistical model can be utilized to predict the received power in the road tunnel. In this section, the statistical model is used to analyze the effect of traffic load and the average vehicle velocity. The traffic load is described using the average rate of vehicle arrival λ. We define $\lambda = 0.5$ as the high traffic load case and $\lambda = 0.3$ as the low traffic load case. In the same way, we define the average vehicle velocity $v = 54$ km/hour as the high speed case and $v = 36$ km/hour as the low speed case. It should be noted that the Doppler effect is not considered since the velocity of the vehicles is not high. The motivation to define the vehicle velocity is to determine the mean number of vehicles inside the tunnel.

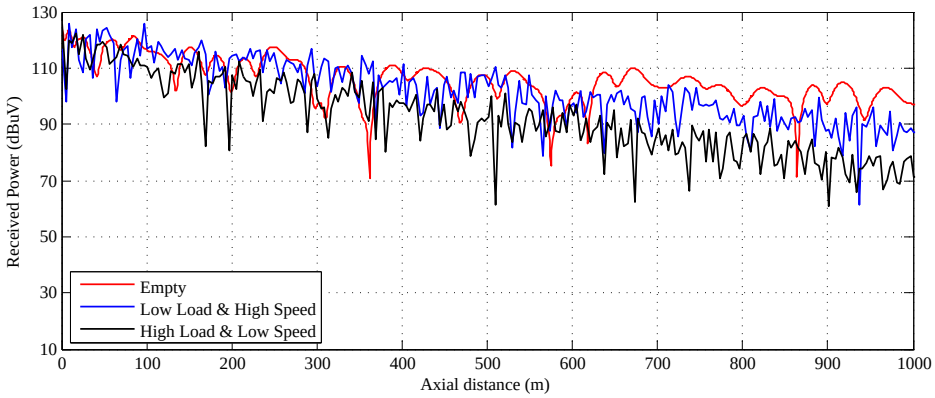
Fig. 53(a), Fig. 53(b) and Fig. 53(c) show the received power under the influence of different types of vehicular traffic flows in a straight one-lane tunnel, a straight two-lane tunnel, and a curved two-lane tunnel, respectively. The curve radius of the curved tunnel is 500 m. The received power is shown as a function of the transmission distance. According to the statistical model, the sizes of the vehicles are uniformly distributed in the size interval defined in (142). Each vehicle runs in one of the L lanes with the same probability. The number of vehicles in the traffic flow is determined by the traffic load and the vehicle



(a) Received signal power in straight one-lane road tunnel.



(b) Received signal power in straight two-lane road tunnel.



(c) Received signal power in curved two-lane road tunnel.

Figure 53: Signal propagation in road tunnels with random vehicular traffic flows

velocity. Fig. 53 shows that the impact of the random vehicular traffic flow is similar to the impact of the deterministic traffic flow. Specifically, 1) The additional loss caused by the traffic flow in the one-lane tunnel is much more significant than the additional loss in the two-lane tunnel, since the same vehicle can block higher ratio of the one-lane tunnel cross section than the two-lane tunnel cross section. 2) The traffic flow with lower load and higher speed causes smaller additional loss than the traffic flow with higher load and lower speed does, which can be explained as follows: When the traffic load is light, fewer vehicles enter the tunnel; and when the average vehicle speed is high, the vehicles in the tunnel leave the tunnel more quickly. Consequently, the number of vehicles in the tunnel is smaller. According to the previous discussion on the effects of vehicle number, the additional path loss caused by the traffic flow with lower load and higher speed should be less significant. 3) In the curved tunnel, the influence of the vehicles on the signal propagation is similar to the straight tunnel case due to the vertically polarized antenna. The signal fluctuation is more severe than the straight tunnel case, which is not caused by the vehicles but the tunnel curvature.

4.4 Capacity and Outage Analysis of MIMO and Cooperative Communication Systems in Underground Mines and Tunnels

As the most important criteria in designing MIMO and cooperative communication systems, the channel capacity and the outage behavior need to be investigated in underground tunnel environments. Specifically, according to channel model in underground mines and tunnels, the channel characteristics in underground tunnels are significantly different from the terrestrial channel. First, due to the reflections on the tunnel walls, the propagation of electromagnetic (EM) waves form regular patterns (i.e. modes) in underground tunnels. Each propagation mode has different field distribution and attenuation rate. The effectiveness of excitation and reception of those modes is determined by the position of the transmitter and the receiver [84, 86]. Second, obstructions with random sizes and positions inside the tunnel, such as vehicles and machineries, cause additional loss and coupling of

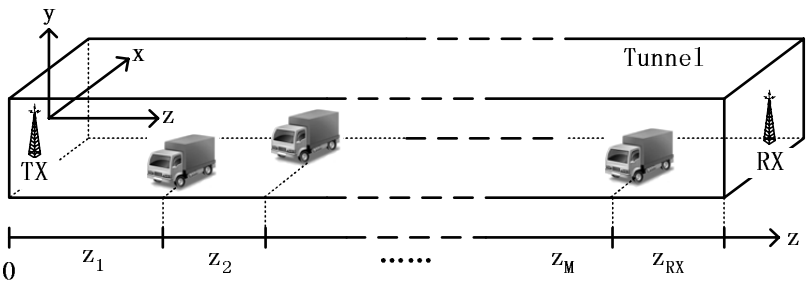


Figure 54: Tunnel Environment

the propagation modes [88, 90]. Since the channel characteristics have straightforward influences on the channel capacity, in this section, we analyze the capacity and outage behavior of the MIMO and cooperative communication systems based on the unique channel model in the underground tunnels.

4.4.1 MIMO Channel in Underground Tunnels

If the underground tunnel is empty, the complex channel gain h_{ij} between receiver i and transmitter j inside the tunnel can be obtained by summing up the gains of all significant modes at receiver's position [86]:

$$h_{ij} = \sqrt{G_t G_r} \cdot \sum_{(m,n)\in \mathbf{N}_{mode}} E^{eign}_{mn,(x_i,y_i)} \cdot C^{TX}_{mn,j} \cdot e^{-\Gamma_{mn} \cdot z_r} , \tag{179}$$

where z_r is the distance between the transmitter and the receiver; (x_i, y_i) is the coordinates on the tunnel cross section of the receiver i; G_t and G_r are the TX and RX antenna gain, respectively; $\Gamma_{mn} = \alpha_{mn} + j\beta_{mn}$ where α_{mn} is the attenuation coefficient of the mode EH_{mn} and β_{mn} is the phase shift coefficient of EH_{mn}; $E^{eign}_{mn,(x,y)}$ is the eigenfunction of mode EH_{mn} at (x, y); $C^{TX}_{mn,j}$ is the intensity of the EH_{mn} mode near the transmitter j. The expressions of $E^{eign}_{mn,(x,y)}$, α_{mn}, β_{mn}, and C^{Tx}_{mn} are given by (110), (112), and (129), respectively.

Actual tunnels are filled with obstructions with different sizes and positions. In Section 4.3, we have theoretically modeled the influence of the traffic flow on the signal propagation in tunnels. Consider that a traffic flow with M vehicles exists between the transceivers in the tunnel, as shown in Fig. 54. The position and size of the k^{th} vehicle are (x^k_v, z^k_v) and (w_k, h_k, l_k), respectively. The transmitter is located at $z_0 = 0$ m and the receiver

is located ate z_r. Then the channel gain h_{ij} between the transmitter j and the receiver i can be expressed as

$$h_{ij} = \sqrt{G_t G_r} \cdot \mathbf{E}^{eign}_{(x_i,y_i)} \cdot \mathbf{A}(z_r - z_M) \cdot \prod_{k=1}^{M} \left[\mathbf{B}_k \cdot \mathbf{A}(z_k - z_{k-1}) \right] \cdot \mathbf{C}^{TX}_{(x_j,y_j)}, \tag{180}$$

where $\mathbf{E}^{eign}_{(x_i y_i)}$ is the eigenfunction vector of the receiver i at (x_i, y_i, z_r); and $\mathbf{C}^{TX}_{(x_j y_j)}$ is the mode intensity vector of the transmitter j at (x_j, y_j, z_0):

$$\mathbf{E}^{eign}_{(x_i y_i)} = \left[E^{eign}_{1,(x_i,y_i)}, E^{eign}_{2,(x_i,y_i)}, \cdots E^{eign}_{N,(x_i,y_i)} \right], \quad \mathbf{C}^{TX}_{(x_j y_j)} = \left[C^{TX}_{1,(x_j,y_j)}, C^{TX}_{2,(x_j,y_j)}, \cdots C^{TX}_{N,(x_j,y_j)} \right]; \tag{181}$$

$\mathbf{A}(z)$ is the $N \times N$ attenuation matrix of transmitting all the N modes for z meters in an empty tunnel; and $\mathbf{B}_k$ is the $N \times N$ influence matrix caused by the k^{th} vehicle in the traffic flow:

$$\mathbf{A}(z) = \begin{pmatrix} e^{-\Gamma_1 \cdot z} & 0 & \cdots & 0 \\ 0 & e^{-\Gamma_2 \cdot z} & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & e^{-\Gamma_N \cdot z} \end{pmatrix}, \quad \mathbf{B}_k = \begin{pmatrix} 1 - L^k_1 & B^{kl}_{2\to 1} & \cdots & B^k_{N\to 1} \\ B^k_{1\to 2} & 1 - L^k_2 & \cdots & B^k_{N\to 2} \\ \vdots & \vdots & \ddots & \vdots \\ B^k_{1\to N} & B^k_{2\to N} & \cdots & 1 - L^k_N \end{pmatrix}, \tag{182}$$

where L^k_{mn} on the diagonal of the matrix is the additional loss of the EH_{mn} mode caused by the k^{th} vehicle in the traffic flow:

$$L^k_{mn} = \frac{1}{4ab} \left[w_k - \frac{2a}{m\pi} (-1)^m \cos\left(\frac{m\pi}{a} x^k_v\right) \sin\left(\frac{m\pi}{a} w_k\right) \right] \cdot \left[h_k - \frac{b}{n\pi} \sin\left(\frac{n\pi}{b} h_k\right) \right]; \tag{183}$$

Except the elements on the diagonal of the matrix $\mathbf{B}_k$, the other elements, i.e. $B^k_{mn\to st}$, are the mode coupling coefficients. According to [88, 90], the influence matrix $\mathbf{B}_k$ can be approximately viewed as a diagonal matrix since the mode coupling coefficients are much smaller than the elements on the matrix diagonal. Consequently, the channel gain h_{ij} in (180) can be simplified as

$$h_{ij} = \sqrt{G_t G_r} \cdot \mathbf{E}^{eign}_{(x_i,y_i)} \cdot \mathbf{D}^M_{z_r} \cdot \mathbf{C}^{TX}_{(x_j,y_j)}, \tag{184}$$

where $\mathbf{D}_{z_r}^{M}$ is the propagation matrix given by

$$\mathbf{D}_{z_r}^{M} = \begin{pmatrix} e^{-\Gamma_1 z_r} \cdot \prod_{k=1}^{M}(1-L_1^k) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & e^{-\Gamma_N z_r} \cdot \prod_{k=1}^{M}(1-L_N^k) \end{pmatrix}. \tag{185}$$

4.4.2 Capacity and Outage Behavior of MIMO Systems in Underground Tunnels

In this section, the capacity and the outage behavior of the MIMO systems are investigated. Explicit formulas of capacity probability density function (PDF), ergodic capacity, and outage capacity are developed. Based on the analysis of MIMO capacity and outage behavior, an outage-optimal MIMO antenna geometry design scheme is developed to maximize the MIMO outage capacity.

4.4.2.1 MIMO Capacity in Underground Tunnels

We consider a narrowband channel with p transmitting (TX) and q receiving antenna elements. The complex channel gain matrix $\mathbf{H}$ is a $q \times p$ matrix $[h_{ij}]_{q\times p}$. The matrix element h_{ij} is the channel gain between RX antenna i and TX antenna j, which is given in (184). We assume that the transmitter does not have the channel state information (CSI). Hence, equal power is allocated to each TX antenna. The MIMO capacity with equal power allocation is given by [96]

$$\mathbb{C}_{MIMO} = E\Big[\log\det\left(\mathbf{I}_{q\times q} + \rho/p \cdot \mathbf{H}\cdot\mathbf{H}^{*}\right)\Big], \tag{186}$$

where ρ is the signal to noise ratio (SNR) at the transmitter, which is defined as $\rho = \frac{P_{tot}}{N_0}$, where P_{tot} is the total transmission power of all TX antenna elements; N_0 is the noise power. The channel gain matrix $\mathbf{H}$ can be derived from (184):

$$\mathbf{H} = \sqrt{G_t G_r} \cdot \mathbf{E}^{RX} \cdot \mathbf{D}_{z_r}^{M} \cdot \mathbf{C}^{TX}, \tag{187}$$

where $\mathbf{D}^M(z_r)$ is the propagation matrix defined in (185); $\mathbf{E}^{RX}$ is the mode eigenfunction matrix at RX side; and $\mathbf{C}^{TX}$ is the mode intensity matrix at TX side:

$$\mathbf{E}^{RX}=\begin{pmatrix} E^{eign}_{1,(x_1,y_1)} & E^{eign}_{2,(x_1,y_1)} & \cdots & E^{eign}_{N,(x_1,y_1)} \\ E^{eign}_{1,(x_2,y_2)} & E^{eign}_{2,(x_2,y_2)} & \cdots & E^{eign}_{N,(x_2,y_2)} \\ \vdots & \vdots & \ddots & \vdots \\ E^{eign}_{1,(x_q,y_q)} & E^{eign}_{2,(x_q,y_q)} & \cdots & E^{eign}_{N,(x_q,y_q)} \end{pmatrix};\quad \mathbf{C}^{TX}=\begin{pmatrix} C^{TX}_{1,(x_1,y_1)} & C^{TX}_{1,(x_2,y_2)} & \cdots & C^{TX}_{1,(x_p,y_p)} \\ C^{TX}_{2,(x_1,y_1)} & C^{TX}_{2,(x_2,y_2)} & \cdots & C^{TX}_{2,(x_p,y_p)} \\ \vdots & \vdots & \ddots & \vdots \\ C^{TX}_{N,(x_1,y_1)} & C^{TX}_{N,(x_2,y_2)} & \cdots & C^{TX}_{N,(x_p,y_p)} \end{pmatrix}.$$

Substitute (187) into (186) and let $G_t = G_r = 1$ for brevity, the MIMO capacity $\mathbb{C}_{MIMO}$ becomes:

$$\mathbb{C}_{MIMO}=\log\det\left(\mathbf{I}_{q\times q}+\frac{\rho}{p}\cdot\mathbf{E}^{RX}\cdot\mathbf{D}^{M}_{z_r}\cdot\mathbf{C}^{TX}\cdot\mathbf{C}^{TX*}\cdot\mathbf{D}^{M*}_{z_r}\cdot\mathbf{E}^{RX*}\right). \tag{188}$$

Since $\det(\mathbf{I}+\mathbf{AB})=\det(\mathbf{I}+\mathbf{BA})$ [81], the MIMO capacity becomes:

$$\mathbb{C}_{MIMO}=\log\det\left[\mathbf{I}_{N\times N}+\frac{\rho}{p}(\mathbf{E}^{RX*}\mathbf{E}^{RX})\cdot\mathbf{D}^{M}_{z_r}(\mathbf{C}^{TX}\mathbf{C}^{TX*})\cdot\mathbf{D}^{M*}_{z_r}\right]. \tag{189}$$

It is difficult to derive the exactly PDF of the MIMO capacity in (189). However, approximate results can be derived if we assume the SNR at the receiver is either high enough or low enough.

High SNR Regime Analysis:

Theorem 3. *In the high SNR regime, the MIMO capacity in underground tunnels follows a normal distribution.*

Proof. As discussed in Section 4.2, only modes that have significant energy need to be considered. Consequently, all the elements on the diagonal of the mode propagation matrix $\mathbf{D}^{M}_{z_r}$ is not trivial. Moreocer, in high SNR regime analysis, we assume that the number of TX antenna elements p and RX antenna elements q are larger than the number of significant modes N (i.e. $p \geq N$ and $q \geq N$). Then $\mathbf{D}^{M}_{z_r}$, $\mathbf{C}^{TX}$ and $\mathbf{E}^{RX}$ are all $N\times N$ full-rank matrix. If the SNR at the RX side is high enough, then the MIMO capacity in (189) can be approximated as

$$\mathbb{C}^{high}_{MIMO}\simeq\sum_{l=1}^{N}\log\left(\rho\cdot\left|e^{-\Gamma_l\cdot z_r}\cdot\prod_{k=1}^{M}(1-L^k_l)\right|^2\right)+\log\det\left(\mathbf{E}^{RX*}\mathbf{E}^{RX}\right)+\log\det\left(\frac{1}{p}\mathbf{C}^{TX}\mathbf{C}^{TX*}\right). \tag{190}$$

The last two terms in (190) are determined by the geometry of the RX and TX antenna elements, respectively. The first term in (190) is the sum of the capacities of N sub-channels. Each sub-channel refers to a propagation mode. The capacity of each sub-channel is governed by the tunnel size and the vehicular traffic flow. Since the number, positions, and sizes of the vehicles between the transceivers are random, the capacities of each sub channels are also random. The first term in (190) can be further developed as

$$\sum_{l=1}^{N} \log\left(\rho \cdot \left| e^{-\Gamma_l \cdot z_r} \cdot \prod_{k=1}^{M} (1 - L_l^k) \right|^2\right) = \sum_{l=1}^{N} \left(\log \rho + 2 \log \left| e^{-\Gamma_l \cdot z_r} \right| \right) + \sum_{l=1}^{N} \sum_{k=1}^{M} 2 \log(1 - L_l^k). \quad (191)$$

The first sum in (191) is determined by the SNR and the mode attenuation coefficients, which are constants after the communication starts. The second sum in (191) is determined by the traffic flow of random vehicles. The additional loss parameters $\{L_l^k, k = 1, 2, ...M; l = 1, 2, ...N\}$ are independent random variables. Since the tunnel is considered to be long, we assume that the number of vehicles M is correspondingly large. Therefore, the second sum in (191) can be viewed as the sum of a sufficiently large number of independent random variables, each with finite mean and variance. According to the central limit theorem and Lindeberg's condition [15], the second sum in (191) are approximately normally distributed. Then the MIMO capacity in (190) is actually the sum of a normal distributed random variable and several constants. Therefore, the MIMO capacity in high SNR regime follows a normal distribution, which completes the proof. □

Since the normal distribution can be completely characterized by its first two moments, we calculate mean and variance of the MIMO capacity in the rest part of this subsection.

Substitute (191) into (190) and calculate the mean value, we derive

$$E[\mathbb{C}_{MIMO}^{high}] = \sum_{l=1}^{N} \left(\log \rho + 2 \log \left| e^{-\Gamma_l \cdot z_r} \right| \right) + \log\det\left(\mathbf{E}^{RX*} \mathbf{E}^{RX}\right) + \log\det\left(\frac{1}{p} \mathbf{C}^{TX} \mathbf{C}^{TX*}\right) + E\left[\sum_{l=1}^{N} \sum_{k=1}^{M} 2 \log(1 - L_l^k)\right]. \quad (192)$$

In underground tunnels, vehicular traffic flow can be modeled as a Poisson flow [37]. If the distance between the transceivers is z_r, the probability that the number of vehicles

$M = m$ is:

$$P(M = m) = \left(\lambda \frac{z_r}{v}\right)^m \cdot \frac{1}{m!} \cdot e^{-\lambda \frac{z_r}{v}}, \tag{193}$$

where λ is the average rate of vehicle arrival (vehicles/sec) in the tunnel; and v is the average velocity of the vehicles. Then the last term in (192) can be further developed as

$$E\Big[\sum_{l=1}^{N}\sum_{k=1}^{M} 2\log(1-L_l^k)\Big] = \sum_{m=0}^{\infty}\Big\{P(M=m)\cdot\sum_{k=1}^{m} E\Big[\sum_{l=1}^{N} 2\log(1-L_l^k)\Big]\Big\}. \tag{194}$$

Every vehicle runs in one of the L lanes in the tunnel. Hence the x-coordinate of the i^{th} vehicles x_v^i belongs to $\{a(\frac{1+2l}{L}-1)|l = 0, 1, \ldots L-1\}$. The x-coordinate of the vehicle obeys uniform distribution with the probability $1/L$. The size of the vehicles are also assumed to have the uniform distribution in the size interval defined in (142). Then the expectations $E[\sum_{l=1}^{N} 2\log(1-L_l^k)]$ in (194) are the same for all $k \in \{1, 2, ...M\}$. Thus, we denote:

$$\mu_{_L} = E\Big[\sum_{l=1}^{N} 2\log(1-L_l^k)\Big], \tag{195}$$

which can be calculated using (183). Substituting (193) and (195) into (194) yields

$$E\Big[\sum_{l=1}^{N}\sum_{k=1}^{M} 2\log(1-L_l^k)\Big] = \sum_{m=0}^{\infty}\Big(\lambda \frac{z_r}{v}\Big)^m \cdot \frac{1}{m!} \cdot e^{-\lambda \frac{z_r}{v}} \cdot m \cdot \mu_{_L} = \lambda \cdot \frac{z_r}{v} \cdot \mu_{_L}. \tag{196}$$

Substituting (196) into (192) yields the mean (ergodic) MIMO capacity in high SNR regime:

$$E[\mathbb{C}_{MIMO}^{high}] = \sum_{l=1}^{N}\Big(\log\rho + 2\log\big|e^{-\Gamma_l \cdot z_r}\big|\Big) + \log\det\Big(\mathbf{E}^{RX*}\mathbf{E}^{RX}\Big) + \log\det\Big(\frac{1}{p}\mathbf{C}^{TX}\mathbf{C}^{TX*}\Big) + \lambda\frac{z_r}{v}\mu_{_L}, \tag{197}$$

The variance of the MIMO capacity in high SNR regime can be calculated in the similar way:

$$Var[\mathbb{C}_{MIMO}^{high}] = \lambda \cdot \frac{z_r}{v} \cdot \sigma_{_L}^2, \quad \text{where} \quad \sigma_{_L}^2 = Var\Big[\sum_{l=1}^{N} 2\log(1-L_l^k)\Big], \tag{198}$$

Note that $\sigma_{_L}^2$ can also be calculated using (183). After the mean and variance of the MIMO capacity are given, the capacity distribution in high SNR regime is completely characterized.

Corollary 4. *In the high SNR regime, the ϵ-outage capacity of the MIMO systems in underground tunnels is given by*

$$Outage_{\epsilon}[\mathbb{C}_{MIMO}^{high}]=E[\mathbb{C}_{MIMO}^{high}]+erf^{-1}(2\epsilon-1)\cdot\sqrt{2Var[\mathbb{C}_{MIMO}^{high}]}, \tag{199}$$

where $erf^{-1}(x)$ is the inverse function of the error function $erf(x)$; $E[\mathbb{C}_{MIMO}^{high}]$ and $Var[\mathbb{C}_{MIMO}^{high}]$ are given by (197) and (198), respectively.

Proof. According to Theorem 1, the MIMO capacity in high SNR regime follows the normal distribution $\mathcal{N}(E[\mathbb{C}_{MIMO}^{high}], Var[\mathbb{C}_{MIMO}^{high}])$. Since the ϵ-outage capacity is the capacity guaranteed for $(1-\epsilon)$ of the channel realizations, we have

$$\epsilon = P_{\mathcal{N}}\Big(\mathbb{C}_{MIMO}^{high} < \text{Outage}_{\epsilon}[\mathbb{C}_{MIMO}^{high}]\Big) = \frac{1}{2}\Bigg[1+\text{erf}\Bigg(\frac{\text{Outage}_{\epsilon}[\mathbb{C}_{MIMO}^{high}]-E[\mathbb{C}_{MIMO}^{high}]}{\sqrt{2Var[\mathbb{C}_{MIMO}^{high}]}}\Bigg)\Bigg], \tag{200}$$

Corollary 1 can be derived from (200), which completes the proof. □

Low SNR Regime Analysis:

Theorem 5. *In the low SNR regime, the MIMO capacity in underground tunnels follows a lognormal distribution.*

Proof. For brevity, we first use a simple matrix to denote the product of matrixes in (189):

$$\mathbf{G} \overset{\text{def}}{=} \frac{\rho}{p}(\mathbf{E}^{RX*}\mathbf{E}^{RX})\cdot\mathbf{D}_{z_r}^{M}\cdot(\mathbf{C}^{TX}\mathbf{C}^{TX*})\cdot\mathbf{D}_{z_r}^{M*} . \tag{201}$$

According to the relationship between matrix determinant and matrix trace [81], (189) can be further developed as

$$\mathbb{C}_{MIMO}=\log\det(\mathbf{I}+\mathbf{G})=\log\Big(\exp\{\text{tr}[\log(\mathbf{I}+\mathbf{G})]\}\Big)=\log\Big\{\sum_{k=0}^{\infty}\frac{1}{k!}\Big[-\sum_{j=1}^{\infty}\frac{(-1)^j}{j!}\text{tr}(\mathbf{G}^j)\Big]^k\Big\}, \tag{202}$$

where $\text{tr}(\mathbf{X})$ is the trace of the matrix $\mathbf{X}$. In low SNR regime,

$$|\text{tr}(\mathbf{G})| \gg |\frac{(-1)^j}{j!}\text{tr}(\mathbf{G}^j)|, \quad j>1; \qquad |\text{tr}(\mathbf{G})| \gg |\frac{1}{k!}[-\text{tr}(\mathbf{G})]^k|, \quad k>1. \tag{203}$$

Substituting (203) into (202) yields

$$\mathbb{C}_{MIMO}^{low} \simeq \log[1+\mathrm{tr}(\mathbf{G})] \simeq \mathrm{tr}(\mathbf{G}) \log e = \frac{\rho}{p} \log e \cdot \sum_{j=1}^{p} \sum_{i=1}^{q} \left| \sum_{l=1}^{N} \left[E_{l,(x_i,y_i)}^{eign} \cdot C_{l,j}^{TX} \cdot e^{-\Gamma_l \cdot z_r} \cdot \prod_{k=1}^{M} (1-L_l^k) \right] \right|^2, \quad (204)$$

For brevity, we denote $f_{l,i,j} \stackrel{\text{def}}{=} E_{l,(x_i,y_i)}^{eign} \cdot C_{l,j}^{TX} \cdot e^{-\Gamma_l \cdot z_r}$. Then

$$\left| \sum_{l=1}^{N} \left[E_{l,(x_i,y_i)}^{eign} \cdot C_{l,j}^{TX} \cdot e^{-\Gamma_l \cdot z_r} \cdot \prod_{k=1}^{M} (1-L_l^k) \right] \right|^2 = \left| \sum_{l=1}^{N} \left\{ \mathrm{Re}(f_{l,i,j}) + j \cdot \mathrm{Im}(f_{l,i,j}) \right] \cdot \prod_{k=1}^{M} (1-L_l^k) \right\} \right|^2 = {X_1}^2 + {X_2}^2, \quad (205)$$

where

$$X_1 = \left| \sum_{l=1}^{N} \left[\mathrm{Re}(f_{l,i,j}) \cdot \prod_{k=1}^{M} (1-L_l^k) \right] \right|; \quad X_2 = \left| \sum_{l=1}^{N} \left[\mathrm{Im}(f_{l,i,j}) \cdot \prod_{k=1}^{M} (1-L_l^k) \right] \right|. \quad (206)$$

where $\mathrm{Re}(x)$ and $\mathrm{Im}(x)$ denote the real and imaginary part of the complex value x, respectively;

Again, we assume that the number of vehicles M is large enough so that $\prod_{k=1}^{M}(1-L_l^k)$ can be viewed as the product of a sufficiently large number of independent and positive random variables. According to the central limit theorem [15], $\prod_{k=1}^{M}(1-L_l^k)$ in (206) approximately follow log-normal distributions. X_1 and X_2 shown in (206) are the absolute values of two weighted sums of the log-normal variables $\prod_{k=1}^{M}(1-L_l^k)$, respectively. According to [63], it has been well recognized that the sum of log-normal random variables can be well approximated by a new lognormal variable. It can be proved that this result also applies to the absolute values of the weighted sums of log-normal variables. Therefore, X_1 and X_2 are also log-normal variables. Moreover, based on the definition of the log-normal distribution, it is easy to prove that the square of a log-normal variable is also a log-normal variable. As a result, ${X_1}^2$ and ${X_2}^2$ follow the log-normal distribution. The MIMO capacity shown in (202) is the sum of the log-normal variables given in (205). Therefore, the MIMO capacity in low SNR regime follows a log-normal distribution. This completes the proof. □

Substitute (205) into (204) and calculate the mean value, we derive

$$\mathbb{C}_{MIMO}^{low} = \frac{\rho}{p} \log e \cdot \sum_{j=1}^{p} \sum_{i=1}^{q} ({X_1}^2 + {X_2}^2). \quad (207)$$

According to the proof of Theorem 2, ${X_1}^2$ and ${X_2}^2$ are log-normal variables. The mean (or variance) of the sum of the log-normal variable can be approximated calculated by the sum of the mean (or variance) of each addend variable [63]. Therefore,

$$E[\mathbb{C}_{MIMO}^{low}] = \frac{\rho}{p}\log e \cdot \sum_{j=1}^{p}\sum_{i=1}^{q}\left(E[{X_1}^2] + E[{X_2}^2]\right), \tag{208}$$

$$Var[\mathbb{C}_{MIMO}^{low}] \simeq (\frac{\rho}{p}\log e)^2 \cdot \sum_{j=1}^{p}\sum_{i=1}^{q}\left(Var[X_1^2] + Var[X_2^2]\right), \tag{209}$$

In the following, we show the detailed calculating process of $E[{X_1}^2]$ and $Var[X_1^2]$, while $E[{X_2}^2]$ and $Var[X_2^2]$ can be easily derived by exchanging $\mathrm{Re}(f_{l,i,j})$ with $\mathrm{Im}(f_{l,i,j})$.

According to [57], the mean and variance of the square of a log-normal variable can be expressed as functions of the mean and variance of the log-normal variable. Then we have:

$$E[{X_1}^2] = E^2[X_1] + Var[X_1], \quad Var[X_1^2] = (E^2[X_1] + Var[X_1])^2[(1+\frac{Var[X_1]}{E^2[X_1]})^4 - 1], \tag{210}$$

Next we calculate the mean and variance of lognormal variable X_1:

$$E[X_1] = \left|\sum_{l=1}^{N}\left\{\mathrm{Re}(f_{l,i,j})\cdot E[\prod_{k=1}^{M}(1-L_l^k)]\right\}\right|, \quad Var[X_1] \simeq \sum_{l=1}^{N}\left\{\mathrm{Re}^2(f_{l,i,j})\cdot Var[\prod_{k=1}^{M}(1-L_l^k)]\right\}, \tag{211}$$

We first calculate $E[X_1]$ in (211), where

$$E[\prod_{k=1}^{M}(1-L_l^k)] = \sum_{m=1}^{\infty}P(M=m)\cdot E[\prod_{k=1}^{m}(1-L_l^k)]. \tag{212}$$

Since the positions and the sizes of the vehicles have independent and identical distributions,

$$E[\prod_{k=1}^{m}(1-L_l^k)] = \prod_{k=1}^{m}E[1-L_l^k] = E^m[1-L_l^k] \stackrel{\text{def}}{=} {\mu_l}^m, \tag{213}$$

where $E[1-L_l^k]$ can be calculate using (183). Here we use μ_l to denote $E[1-L_l^k]$ for brevity. Substituting (193), (212), and (213) into (211) yields:

$$E[X_1] = \left|\sum_{l=1}^{N}\left\{\mathrm{Re}(f_{l,i,j})\cdot e^{\lambda\frac{\tau_r}{v}(\mu_l-1)}\right\}\right|. \tag{214}$$

Similarly, we can derive

$$Var[X_1] \simeq \sum_{l=1}^{N} \left\{ \mathrm{Re}^2(f_{l,i,j}) \cdot \left[e^{\lambda \frac{\bar{z}_r}{v}(\mu_l^2+\sigma_l^2-1)} - e^{\lambda \frac{\bar{z}_r}{v}(\mu_l^2-1)} \right] \right\}; \quad \text{and} \ \ \sigma_l^2 \stackrel{\text{def}}{=} Var[1\text{–}L_l^k]. \tag{215}$$

where μ_l is defined in (213); we use σ_l^2 to denote $Var[1\text{–}L_l^k]$ for brevity.

By substituting (210), (214), (215) into (208) and (209), we can calculate the mean (ergodic capacity) and variance of the MIMO capacity in low SNR regime. Then the outage capacity of the MIMO systems in low SNR regime can be calculated based on the following corollary.

Corollary 6. *In the low SNR regime, the ϵ-outage capacity of the MIMO systems in underground tunnels is given by*

$$Outage_\epsilon[\mathbb{C}_{MIMO}^{low}] = \frac{E[\mathbb{C}_{MIMO}^{low}]}{\sqrt{\kappa}} \cdot e^{erf^{-1}(2\epsilon-1)\cdot\sqrt{2\ln\kappa}}, \ \ where \ \ \kappa = 1 + \frac{Var[\mathbb{C}_{MIMO}^{low}]}{E^2[\mathbb{C}_{MIMO}^{low}]}, \tag{216}$$

where $E[\mathbb{C}_{MIMO}^{low}]$ and $Var[\mathbb{C}_{MIMO}^{low}]$ are calculated in the above analysis.

Proof. According to Theorem 2, the MIMO capacity in low SNR regime follows the log-normal distribution. Then we have

$$\epsilon = P_{\log\text{-}\mathcal{N}}\left(\mathbb{C}_{MIMO}^{low} < \mathrm{Outage}_\epsilon[\mathbb{C}_{MIMO}^{low}]\right) = \frac{1}{2}\left[1 + \mathrm{erf}\left(\frac{\ln \mathrm{Outage}_\epsilon[\mathbb{C}_{MIMO}^{low}] - \ln \frac{E[\mathbb{C}_{MIMO}^{low}]}{\sqrt{\kappa}}}{\sqrt{2\ln\kappa}}\right)\right], \tag{217}$$

Corollary 2 can be derived from (217), which completes the proof. □

4.4.2.2 MIMO Antenna Geometry Design Scheme

So far the MIMO capacity in underground tunnels are thoroughly analyzed. In this subsection, the optimal MIMO antenna geometry is designed to maximize the outage capacity. Since the MIMO capacity in tunnels have different attributes in high SNR and low SNR regime, in this subsection, we first develop the optimal MIMO antenna geometry in high SNR regimes. Then this optimal MIMO antenna geometry is modified to fit the attributes in the low SNR regime.

MIMO Antenna Geometry in High SNR Regime:

In (199), the high SNR ϵ-outage capacity is determined by the mean and variance of the high SNR MIMO capacity. According to (197) and (198), the TX and RX antenna geometries can only affect the mean capacity in (197) but have no effect on the capacity variance in (198). As a result, in high SNR regime, to maximize the ϵ-outage capacity is equal to maximize the mean capacity. Hence, in the next step, we design the TX and RX antenna geometry to maximize the mean capacity given in (197).

The first term and the last term in (197) are determined by the tunnel size and the vehicular traffic flow between the transceivers, which are not affected by the TX and RX antenna geometry. The optimal antenna geometries are expected to maximize the second and third term in (197). It is equal to maximize $\det(\mathbf{E}^{RX*}\mathbf{E}^{RX})$ by selecting optimal RX antenna geometry, and to maximize $\det(\frac{1}{p}\mathbf{C}^{TX}\mathbf{C}^{TX*})$ by selecting optimal TX antenna geometry.

According to (110), the eigenfunctions of different modes are orthogonal to each other:

$$\int_{-a}^{a}\int_{-b}^{b} E^{eign}_{i,(x,y)} \cdot E^{eign*}_{j,(x,y)}\, dx\, dy \simeq \begin{cases} ab, & \text{if } i = j \\ 0, & \text{otherwise} \end{cases} \tag{218}$$

At RX side, the matrix $\mathbf{E}^{RX*}\mathbf{E}^{RX}$ is in fact the covariance matrix of the eigenfunctions of all significant modes at the positions of the RX antenna elements. Due to the mode orthogonality, $\det(\mathbf{E}^{RX*}\mathbf{E}^{RX})$ is maximized if $\mathbf{E}^{RX*}\mathbf{E}^{RX}$ is diagonal. If the RX antenna elements are placed at all the positions where the eigenfunctions of significant modes have extrema values, the matrix $\mathbf{E}^{RX}\cdot\mathbf{E}^{RX*}$ can be approximately diagonalized. The maximum value of $\det(\mathbf{E}^{RX*}\mathbf{E}^{RX})$ is

$$\max\{\det(\mathbf{E}^{RX*}\mathbf{E}^{RX})\} \simeq \prod_{i=1}^{N}\left[\sum_{i=1}^{q}|E^{eign}_{i,(x,y)}|^2\right] \simeq q^N. \tag{219}$$

At TX side, we first check the mode intensity $C^{TX}_{mn,j}$ given in (129). Since $\sqrt{1-(\frac{m\pi}{2ak})^2-(\frac{n\pi}{2bk})^2} \simeq 1$ for low order modes (i.e. the significant modes considered in this section), $C^{TX}_{mn,j}$ approximately equals to the mode eigenfunction multiplied by a constant. Hence, the matrix $\mathbf{C}^{TX}\mathbf{C}^{TX*}$ can also be viewed as the covariance matrix of the eigenfunctions of all significant modes at the positions of the TX antenna elements. Therefore, in high SNR regime, the

optimal TX antenna shares the same geometry as the optimal RX antenna geometry, where antenna elements are placed at all the positions where the eigenfunctions of significant modes have extrema values. And

$$\max\{\det(\frac{1}{p}\mathbf{C}^{TX}\mathbf{C}^{TX*})\} \simeq \prod_{i=1}^{N}\left[\sum_{j=1}^{p}\frac{1}{p}\left|\frac{\pi}{ab}E^{eign}_{j,(x,y)}\right|^2\right] \simeq \left(\frac{\pi}{ab}\right)^{2N} \tag{220}$$

Based on the above discussion, the optimal RX and TX MIMO antenna geometry in high SNR regime is described as follows: the number and geometry of the antenna elements depend on which modes have significant power in the tunnel. For mode EM_{mn}, $p = q = m \times n$ antenna elements are needed. Their positions $\{(x_u, y_v)\}$ should be

$$x_u = -a + (u - \frac{1}{2})\frac{2a}{m}, \;\; u \in [1, m]\;; \qquad y_v = -b + (v - \frac{1}{2})\frac{2b}{n}, \;\; v \in [1, n] \tag{221}$$

Since lower order modes have lower attenuation rates, the probability that the lower order modes have significant power is higher. Therefore, lower order modes are first to be considered when the number of available antenna elements is limited. For example, if we only have one antenna element, the lowest order mode EH_{11} is considered. The position of the antenna element is $(0, 0)$. If we have three antenna elements, both mode EH_{11} and EH_{21} are considered. The positions of the three antenna elements are $(0, 0)$, $(-a/2, 0)$ and $(a/2, 0)$, so on and so forth.

Intuitively, in high SNR regime, each significant mode needs a set of TX and RX antenna elements to be efficiently excited and received. By substituting (219) and (220) into (197) and (199), we find that the high SNR ϵ-outage capacity increases linearly with $\log q$ where q is the number of RX elements. It indicates that more RX elements are always favorable although the capacity increase becomes trivial when q is larger than a threshold. In contrast, the high SNR ϵ-outage capacity remains the same as the number of TX elements increases, which is because that the total TX power P_{tot} is a constant and is equally divided and allocated to each TX elements.

MIMO Antenna Geometry in Low SNR Regime:

In low SNR regime, the optimal RX antenna geometry designed for high SNR regime still works, since the proposed RX antenna geometry can effectively receive all significant propagation modes no matter whether the SNR is high or low. However, the number and geometry of the TX antenna elements in low SNR regime need to be redesigned since it may be not efficient to involve a large number of TX antenna elements to excite all propagation modes in low SNR regime.

Based on the above discussion, our goal in this subsection is to maximize the low SNR ϵ-outage capacity in (216) by designing the number and geometry of the TX antenna elements while keep the RX antenna geometry designed for high SNR regime. In (216), the low SNR ϵ-outage capacity is determined by the mean capacity $E[\mathbb{C}_{MIMO}^{low}]$ and the coefficient κ, where κ is determined by the ratio $\frac{Var[\mathbb{C}_{MIMO}^{low}]}{E^2[\mathbb{C}_{MIMO}^{low}]}$, which can be approximately viewed as a constant if only antenna geometry can be changed, according to (208) - (215). Hence, the coefficient κ becomes a constant. As a result, in low SNR regime, to maximize the ϵ-outage capacity is equal to maximize the mean capacity $E[\mathbb{C}_{MIMO}^{low}]$, in the condition that only the antenna geometry can be designed.

Since the optimal RX antenna geometry for high SNR regime is also utilized in low SNR regime, the correlation matrix $\mathbf{E}^{RX*}\mathbf{E}^{RX}$ becomes a diagonal matrix. Then (201) and (204) becomes:

$$\mathbb{C}_{MIMO}^{low} \simeq \log e \cdot \mathrm{tr}\Big[\frac{\rho}{p}(\mathbf{E}^{RX*}\mathbf{E}^{RX})\cdot\mathbf{D}_{z_r}^{M}(\mathbf{C}^{TX}\mathbf{C}^{TX*})\cdot\mathbf{D}_{z_r}^{M*}\Big] = \log e \cdot \mathrm{tr}\Big[\frac{\rho}{p}\cdot q\cdot\mathbf{I}\cdot\mathbf{D}_{z_r}^{M}(\mathbf{C}^{TX}\mathbf{C}^{TX*})\cdot\mathbf{D}_{z_r}^{M*}\Big]$$
$$= \log e \cdot \mathrm{tr}\Big[\frac{q\rho}{p}(\mathbf{C}^{TX}\mathbf{C}^{TX*})\cdot(\mathbf{D}_{z_r}^{M*}\mathbf{D}_{z_r}^{M})\Big] = \frac{q\rho}{p}\log e\cdot\sum_{l=1}^{N}\sum_{j=1}^{p}\Big|C_{l,j}^{TX}\Big|^2\cdot\Big|e^{-\Gamma_l\cdot z_r}\cdot\prod_{k=1}^{M}(1-L_l^k)\Big|^2. \quad (222)$$

To determine the optimal number and positions of the TX antenna elements in low SNR regime, we first assume that a sufficient large number of TX antenna elements are placed at the whole tunnel cross section. Then we define a set of indicators $\{\delta_j, j = 1, 2, ..., p\}$ where $\delta_j = 1$ if the j^{th} TX antenna element is actually used, otherwise $\delta_j = 0$. Then (222)

becomes:

$$\mathbb{C}_{MIMO}^{low} \simeq \frac{q\rho \log e}{\sum_{j=1}^{p}\delta_j} \cdot \sum_{j=1}^{p} \delta_j \cdot \sum_{l=1}^{N} \left| C_{l,j}^{TX} \cdot e^{-\Gamma_l \cdot z_r} \right|^2 \cdot \prod_{k=1}^{M} (1-L_l^k)^2. \tag{223}$$

Hence,

$$E[\mathbb{C}_{MIMO}^{low}] \tag{224}$$
$$\simeq \frac{q\rho \log e}{\sum_{j=1}^{p}\delta_j} \cdot \sum_{j=1}^{p} \delta_j \cdot \sum_{l=1}^{N} \left| C_{l,j}^{TX} \cdot e^{-\Gamma_l \cdot z_r} \right|^2 \cdot E\Big[\prod_{k=1}^{M} (1-L_l^k)^2 \Big] = \frac{q\rho \log e}{\sum_{j=1}^{p}\delta_j} \cdot \sum_{j=1}^{p} \delta_j \cdot \sum_{l=1}^{N} \left| C_{l,j}^{TX} \cdot e^{-\Gamma_l \cdot z_r} \right|^2 \cdot e^{\lambda \frac{z_r}{v}(\mu_l^2+\sigma_l^2-1)},$$

where μ_l and σ_l^2 is defined in (213) and (215), respectively. Therefore, to find the optimal TX antenna geometry in low SNR regime, it is equal to an optimization problem, which is:

$$Find: \ \{\delta_j, j = 1, 2, ..., p\}$$
$$Maximize: \ \frac{q\rho \log e}{\sum_{j=1}^{p}\delta_j} \cdot \sum_{j=1}^{p} \delta_j \cdot \sum_{l=1}^{N} \left| C_{l,j}^{TX} \cdot e^{-\Gamma_l \cdot z_r} \right|^2 \cdot e^{\lambda \frac{z_r}{v}(\mu_l^2+\sigma_l^2-1)}$$

It is not difficult to find the solution of this optimization problem, which is: $\delta_k = 1$ if $k = \arg\max_j \sum_{l=1}^{N} \left| C_{l,j}^{TX} \cdot e^{-\Gamma_l \cdot z_r} \right|^2 \cdot e^{\lambda \frac{z_r}{v}(\mu_l^2+\sigma_l^2-1)}$; otherwise $\delta_k = 0$. Therefore, the optimal TX antenna geometry in low SNR regime involves only one antenna element. According to the channel model given in Section 4.2, the position of the optimal TX antenna element is located at the center of the tunnel cross section, i.e., $(x, y) = (0, 0)$.

To sum up, the optimal RX antenna geometry in low SNR regime is the same as in the high SNR regime. The optimal TX antenna geometry in low SNR regime involves only one antenna element that is located at the center of the tunnel cross section. This results can be intuitively explained as follows. In low SNR regime, given a fixed total transmitting power P_{tot}, it is more efficient to concentrate all TX power at the antenna element that can achieves smallest path loss.

4.4.3 Capacity and Outage Behavior of Cooperative Communication Systems in Underground Tunnels

In many applications, such as wireless sensor networks, it is impossible to place multiple antenna elements on a single device due to the limited size and cost. In this case, instead of

MIMO, cooperative communication systems can be utilized. In this section, the capacity and outage behavior of the cooperative communication systems in underground tunnels are investigated.

4.4.3.1 Capacity and Outage Probability

In this subsection, we investigate the capacity and the outage performance of the selection Decode-and-Forward (DF) cooperative scheme [51], which has been widely recognized. The selection DF scheme consists of two phases. In the first phase, the source node sends out the information, which is received by both the relay node and the destination node. In the second phase, the relay node check the SNR of the received signal. If the SNR is above a threshold, the relay node decodes and forwards the received data to the destination node. Otherwise, the source nodes just send the information again to the destination node. According to [51], the channel capacity of this cooperative system is given by

$$\mathbb{C}_{coop} = \begin{cases} \frac{1}{2}\log(1+2\rho|h_{sd}|^2), & \text{if } |h_{sr}|^2 < \frac{e^R-1}{\rho} \\ \frac{1}{2}\log(1+\rho|h_{sd}|^2+\rho|h_{rd}|^2), & \text{if } |h_{sr}|^2 \geq \frac{e^R-1}{\rho} \end{cases} \tag{225}$$

where h_{sd}, h_{rd}, and h_{sr} are the channel gains between the source and the destination, the relay and the destination, and the source and the relay, respectively; ρ is the SNR at the transmitter; R is the expected data rate. In underground tunnels, the channel gain h_{sd}, h_{rd}, and h_{sr} can be calculated by (184). Then the instantaneous cooperative capacity in underground tunnels can be derived. However, the distribution of the cooperative capacity in underground tunnels does not fall into any type of classical random distributions. In the following, we calculate the outage probability to characterize this distribution. The outage probability of the DF cooperative scheme can be derived according to (225), which is given by:

$$P[\mathbb{C}_{coop} < R] = P(|h_{sr}|^2 < \frac{e^R-1}{\rho}) P(|h_{sd}|^2 < \frac{e^R-1}{2\rho}) + P(|h_{sr}|^2 \geq \frac{e^R-1}{\rho}) P(|h_{sd}|^2 + |h_{rd}|^2 < \frac{e^R-1}{\rho}). \tag{226}$$

According to (226), we find that the outage probability is determined by the channel gains $|h_{sd}|^2$, $|h_{rd}|^2$, and $|h_{sr}|^2$. In Section 4.1.2, the square of the channel gain norm $|h_{ij}|^2$ is developed as (205), which is proved to follow a log-normal distribution. The mean and variance of this log-normal variable, $E[|h_{ij}|^2]$ and $Var[|h_{ij}|^2]$, can be calculated using (210)-(215). Then,

$$P(|h_{ij}|^2<\frac{e^R-1}{\rho}) = \frac{1}{2}\Big[1+\text{erf}\Big(\frac{\ln\frac{e^R-1}{\rho} - \mu_{|h_{ij}|^2}}{\sqrt{2}\cdot\sigma_{|h_{ij}|^2}}\Big)\Big], \tag{227}$$

where

$$\mu_{|h_{ij}|^2} = \ln(E[|h_{ij}|^2]) - \frac{1}{2}\ln\Big(1+\frac{Var[|h_{ij}|^2]}{E^2[|h_{ij}|^2]}\Big); \qquad \sigma^2_{|h_{ij}|^2} = \ln\Big(1+\frac{Var[|h_{ij}|^2]}{E^2[|h_{ij}|^2]}\Big). \tag{228}$$

The sum of two log-normal variable, $|h_{sd}|^2 + |h_{rd}|^2$, also follows a log-normal distribution:

$$P(|h_{sd}|^2 + |h_{rd}|^2<\frac{e^R-1}{\rho}) = \frac{1}{2}\Big[1+\text{erf}\Big(\frac{\ln\frac{e^R-1}{\rho} - \mu_{(|h_{sd}|^2+|h_{rd}|^2)}}{\sqrt{2}\cdot\sigma_{(|h_{sd}|^2+|h_{rd}|^2)}}\Big)\Big], \tag{229}$$

where

$$\begin{aligned}\mu_{(|h_{sd}|^2+|h_{rd}|^2)} &= \ln(E[|h_{sd}|^2] + E[|h_{rd}|^2]) - \frac{1}{2}\ln\Big(1+\frac{Var[|h_{sd}|^2] + Var[|h_{rd}|^2]}{(E[|h_{sd}|^2] + E[|h_{rd}|^2])^2}\Big);\\ \sigma^2_{(|h_{sd}|^2+|h_{rd}|^2)} &= \ln\Big(1+\frac{Var[|h_{sd}|^2] + Var[|h_{rd}|^2]}{(E[|h_{sd}|^2] + E[|h_{rd}|^2])^2}\Big).\end{aligned} \tag{230}$$

Substituting (227) and (229) into (226) yields the outage probability in underground tunnels:

$$\begin{aligned}P[\mathbb{C}_{coop} < R] = &\frac{1}{4}\Big[1+\text{erf}\Big(\frac{\ln\frac{e^R-1}{\rho} - \mu_{|h_{sr}|^2}}{\sqrt{2}\cdot\sigma_{|h_{sr}|^2}}\Big)\Big]\cdot\Big[1+\text{erf}\Big(\frac{\ln\frac{e^R-1}{2\rho} - \mu_{|h_{sd}|^2}}{\sqrt{2}\cdot\sigma_{|h_{sd}|^2}}\Big)\Big]\\ &+\frac{1}{4}\Big[1-\text{erf}\Big(\frac{\ln\frac{e^R-1}{\rho} - \mu_{|h_{sr}|^2}}{\sqrt{2}\cdot\sigma_{|h_{sr}|^2}}\Big)\Big]\cdot\Big[1+\text{erf}\Big(\frac{\ln\frac{e^R-1}{\rho} - \mu_{(|h_{sd}|^2+|h_{rd}|^2)}}{\sqrt{2}\cdot\sigma_{(|h_{sd}|^2+|h_{rd}|^2)}}\Big)\Big].\end{aligned} \tag{231}$$

Note that the ϵ-outage cooperative capacity $\text{Outage}_\epsilon[\mathbb{C}_{coop}]$ can be also calculated from (231) by letting $R = \text{Outage}_\epsilon[\mathbb{C}_{coop}]$ and $P[\mathbb{C}_{coop} < \text{Outage}_\epsilon[\mathbb{C}_{coop}]] = \epsilon$.

4.4.3.2 Optimal Relay Assignment in Tunnels

So far we assume that the relay position is determined. However, in real applications, especially the mobile wireless networks, the positions of the cooperative users are highly dynamic. Since the relay position has obvious influence on the outage probability, the cooperative relay assignment is of great important. In this subsection, we first use the outage probability given in (231) as the relay assignment metric and propose an outage-optimal and fully distributed cooperative relay assignment protocol. Then, we develop a much simpler relay assignment metric based on the outage probability to reduce the computation burden on each cooperative node.

Outage-Optimal Relay Assignment Protocol:

Assuming that in an underground tunnel, the source node s is to send data to the destination node d. Node s assign one of its neighbors $\{r_i, i = 1, 2, ...\}$ as the relay node and adopts the selection DF scheme. All the nodes have its own position information, which is valid in most applications. The source first sends RTS message to all its neighbors. RTS message contains the position information of the source and the destination. Each neighbor that can correctly receives the RTS message calculates the relay assignment metric. Here we first use the outage probability given in (231) as the relay assignment metric. To calculate this metric, the following information is required: 1) the position information of itself, the source, and the destination, and 2) the statistical vehicular traffic load information in the underground tunnel, which is assume to be available at each node. No information exchanging among neighbor nodes and the destination node is required.

After deriving the relay assignment metric, each neighbor node randomly select a back-off time and start timing. The mean value of the back-off time is proportional to the relay assignment metric, i.e. the outage probability. The neighbor node sends out a CTS message to the source node after the back-off time out if it does not receive any other CTS messages from other neighbors during the back-off time. Then the source node assigns this neighbor as the relay node and start the selection decode-and-forward cooperative communication.

By this relay assignment protocol, the neighbor node that has the least outage probability will be selected as the cooperative relay. Therefore, the protocol is outage-optimal. Moreover, since calculating the relay assignment metric only require the the source and destination position information in the RTS message and the position information of each neighbor node itself, the proposed protocol is also fully distributed.

Simpler Relay Assignment Metric:

Despite the advantages of the proposed relay assignment protocol, the computation burden to calculate the outage probability in (231) may be heavy for low cost and resource limited devices, such as the sensor nodes. Therefore, we propose a simpler relay assignment metric based on the outage probability given in (231).

According to (231), to minimize the outage probability is equal to maximize the parameters $\mu_{|h_{sr}|^2}$ and $\mu_{(|h_{sd}|^2+|h_{rd}|^2)}$ while minimizing the parameters $\sigma_{|h_{sr}|^2}$ and $\sigma_{(|h_{sd}|^2+|h_{rd}|^2)}$. Similar to the analysis in Section 4.2.2, it can be proved that the position of the relay node does not significantly affect the ratio $\frac{Var[|h_{sd}|^2]+Var[|h_{rd}|^2]}{(E[|h_{sd}|^2]+E[|h_{rd}|^2])^2}$ and the ratio $\frac{Var[|h_{sr}|^2]}{E^2[|h_{sr}|^2]}$. Hence, according to (228) and (230), to minimize the outage probability in (231) is equal to maximize $\ln(E[|h_{sr}|^2])$ and $\ln(E[|h_{sd}|^2] + E[|h_{rd}|^2])$. Since $E[|h_{sd}|^2]$ is determined, ony $E[|h_{sr}|^2]$ and $E[|h_{rd}|^2]$ need to be considered. However, there exists tradeoff between $|h_{sr}|$ and $|h_{rd}|$. The optimal relay node should be placed at the position where $|h_{sr}|$ and $|h_{rd}|$ are both sufficient large. Motivated by the above discussion, to reduce the metric computation complexity, we use $\min\{E[|h_{sr}|], E[|h_{rd}|]\}$ as a simpler metric instead of the outage probability.

This simpler metric metric can be calculated based on (205), (206), and (214), which can be further simplified by the following approximation. Since the lowest order mode EH_{11} has the lowest attenuation rate in tunnels, if the distance between source and destination is large enough, we can approximately assume that only the lowest order mode exists. This assumption is valid since the cooperative communication is not necessary in short range distance communication where the impact of multipath fading is negligible. Then

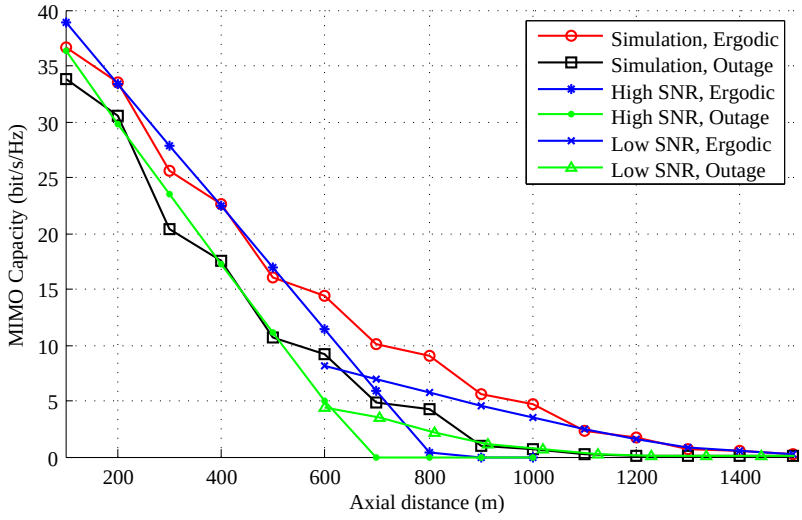


Figure 55: Ergodic and 10%-outage MIMO capacity as functions of the axial distance between transceivers.

the simpler metric can be calculated as

$$\min\{E[|h_{sr}|], E[|h_{rd}|]\} \tag{232}$$
$$\simeq \min\left\{\left|\frac{\pi}{ab}E^{eign}_{1,(x_r,y_r)}\cdot E^{eign}_{1,(x_s,y_s)}\cdot e^{z_{sr}[-\Gamma_1+\frac{\lambda}{v}(\mu_1-1)]}\right|, \left|\frac{\pi}{ab}E^{eign}_{1,(x_r,y_r)}\cdot E^{eign}_{1,(x_d,y_d)}\cdot e^{z_{rd}[-\Gamma_1+\frac{\lambda}{v}(\mu_1-1)]}\right|\right\},$$

where z_{sr} is the axial distance between the source and the relay; z_{rd} is the axial distance between the destination and the relay; Γ_1 is the attenuation coefficient of mode EH_{11}; $E^{eign}_{1,(x,y)}$ is given in (110); μ_1 is defined in (213).

The simpler relay assignment metric given in (232) can be adopted in resource limited devices instead of the outage probability given in (231). Note that the back-off time is inversely proportional to this simpler relay assignment metric.

4.4.4 Numerical Analysis

In this section, the capacity and outage behavior of MIMO and cooperative communication systems in underground tunnels are numerically analyzed. Then, the proposed MIMO antenna geometry as well as the cooperative relay assignment protocol are compared with existing geometry and protocols, respectively. Except studying the effect of certain parameters, the default simulation parameters are set as follows: The one-lane tunnel has a

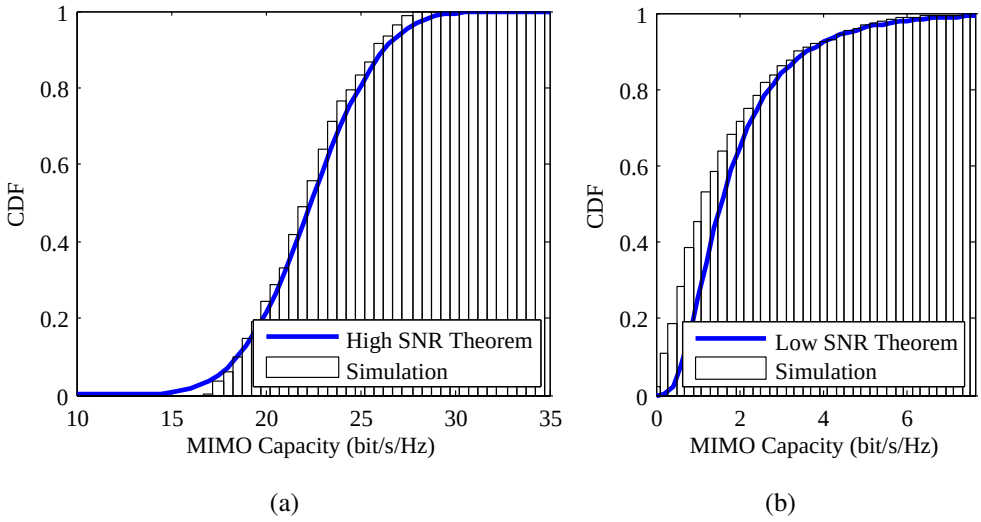


Figure 56: CDF of MIMO capacity: (a) high SNR, (b) low SNR.

rectangle cross section with a height of 4 m and a width of 6 m. The operating frequency is 1 GHz. The SNR at the transmitter $\rho = \frac{P_{tot}}{N_0}$ is 100 dB. The traffic load is described using the average rate of vehicle arrival λ and average vehicle velocity v. The default values are set as $\lambda = 0.5\ s^{-1}$ and $v = 72$ km/hour.

4.4.4.1 MIMO Systems in Underground Tunnels

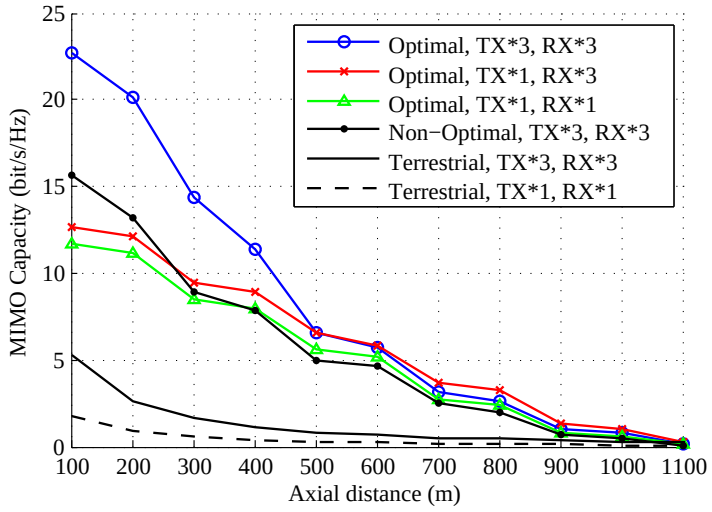
Fig. 55 shows the ergodic and the 10%-outage MIMO capacity as functions of of the axial distance between the transceivers, where the 5×5 MIMO antenna with optimal geometry is adopted. It shows that the theoretical results in high SNR regime have a good match with the simulations when the axial distance is smaller than 500 m, while the theoretical results in low SNR regime have a good match with the simulations when the axial distance is larger than 1000 m. Therefore, the theoretical ergodic and outage MIMO capacities have a good match with the simulations in both high SNR regime and low SNR regime.

In Fig. 56(a), the cumulative distribution function (CDF) of the normal distribution with parameters derived in high SNR regime is compared with the simulated capacity distribution when axial distance is 400 m. In Fig. 56(b), the CDF of the log-normal distribution

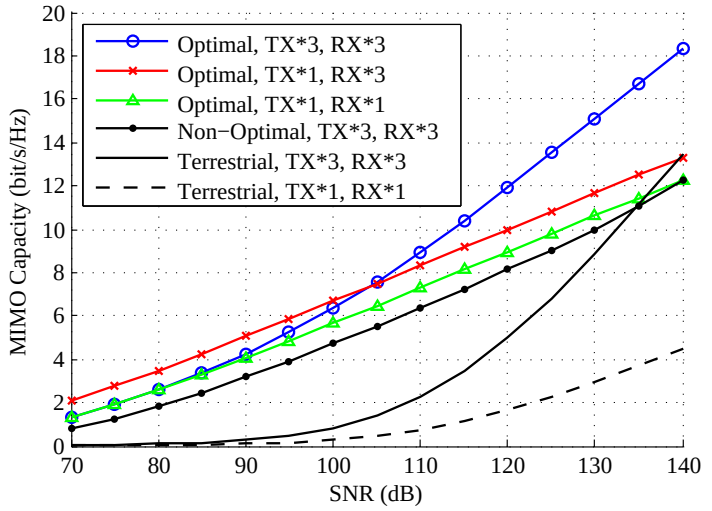
with parameters derived in low SNR regime is compared with the simulated capacity distribution when axial distance is 1200 m. A good match is shown between the theoretical and simulation results.

Fig. 57(a) and Fig. 57(b) show the 10%-Outage capacity of the optimal MIMO antenna geometries with different element number as a function of the axial distance and the SNR at the transmitter, respectively. The capacity of a traditional linear antenna array with 3 antenna elements is also provided for comparison. For fair comparison, one terminal of the traditional linear antenna array is placed at the center of the tunnel. The array is placed parallel to the floor with interval of one wavelength (0.33 m). Fig. 57(a) and Fig. 57(b) indicate that, the MIMO capacity with optimal geometry is significantly higher than that of the undesigned geometry. In low SNR regime, the undesigned linear array has even smaller capacity than that of the SISO (1×1) system placed at the optimal position. In high SNR regime (axial distance smaller than 500 m), the optimal 3×3 MIMO antenna achieves much higher capacity than the optimal 1×3 SIMO antenna. In low SNR regime (axial distance larger than 600 m), the optimal 1×3 SIMO antenna achieves higher outage capacity than the optimal 3×3 MIMO antenna. This phenomenon is consistent with our analysis on the MIMO antenna geometry design in Section 4.4.2.2, i.e. the spatial diversity at the TX side is more efficient in the high SNR regime that in the low SNR regime.

Moreover, Fig. 57(a) and Fig. 57(b) also show the capacity of a 3×3 MIMO system and a SISO system in terrestrial environments. A widely used terrestrial channel model is utilized, where the Friis transmission equation with a exponent of 3 is used to model the path loss and a Rayleigh random variable is used to model the multipath fading. The MIMO antenna elements are placed far enough to guarantee they are not correlated. Fig. 57(a) and Fig. 57(b) show that the capacity of MIMO and SISO system is much smaller than the capacity in tunnels since the signal energy in tunnels does not spread as much as in terrestrial environments. Meanwhile, the gain of MIMO compared with SISO in terrestrial environments is larger than in tunnels since the MIMO antenna elements in tunnels are



(a)



(b)

Figure 57: 10%-Outage MIMO capacity with different antenna geometries as a function of (a) axial distance (SNR at the transmitter is 100 dB) and (b) SNR at the transmitter (axial distance is 500 m).

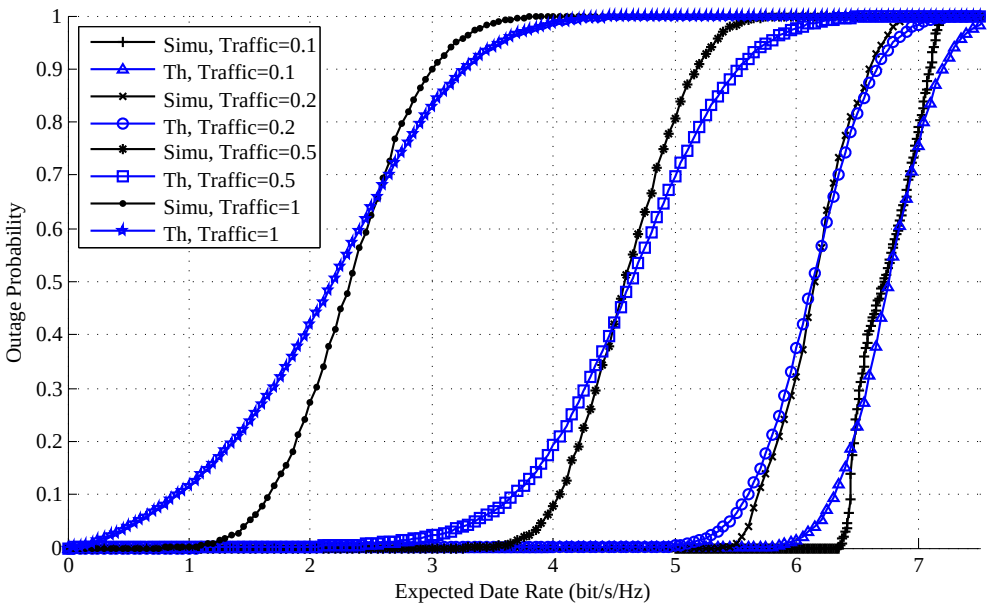


Figure 58: Outage probabilities of cooperative communication systems with different traffic loads.

more correlated.

4.4.4.2 Cooperative Communication Systems in Underground Tunnels

In the following analysis on cooperative communication systems in tunnels, we assume the source node and the destination node are 400 m apart. Both nodes are placed at the center of the tunnel cross section. First, we assume that the relay node is fixed, which is placed in the middle point of the distance between the source and the destination, i.e., at the center of the tunnel cross section and 200 m apart from both the source and the destination. In Fig. 58, it is shown that the theoretical outage probability is consistent with the simulation results with different traffic load (average arrival rate λ). As the traffic load become heavier, the mean cooperative capacity decreases and the variance of the cooperative capacity increases, which is due to the reason that heavier traffic load causes higher additional path loss and more randomness.

Then, the performance of the relay assignment protocols is analyzed. In the following simulations, we let 50 cooperative relay nodes uniformly distributed between the source

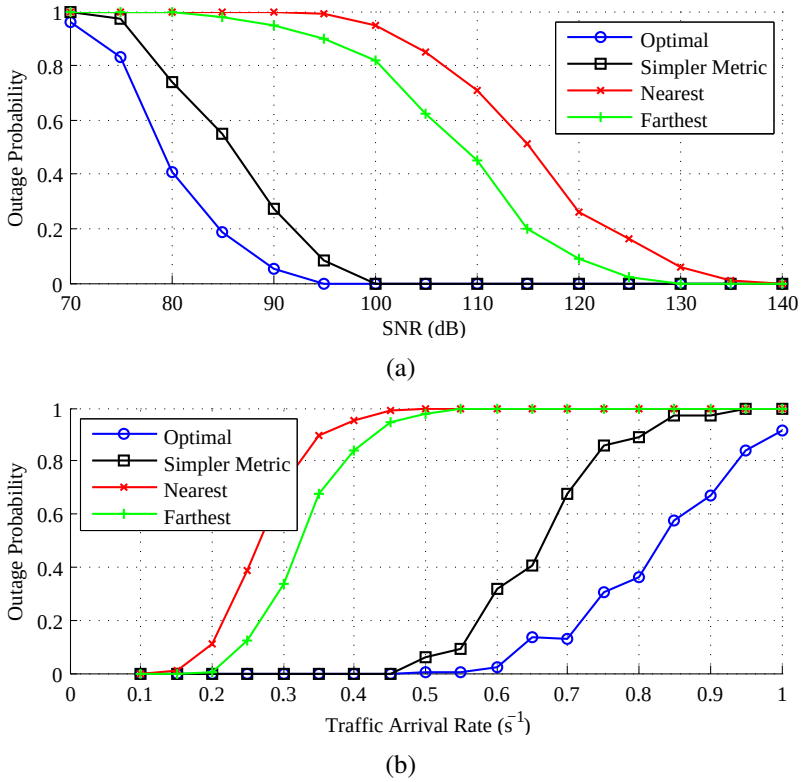


Figure 59: Outage probability with different cooperative relay assignment strategy as functions of different (a) SNR at the transmitter and (b) traffic load.

and the destination. The position of each relay node on the tunnel cross section is also uniformly distributed. Four relay assignment strategies are adopted: 1) optimal strategy that use the outage probability given in (231) as the metric; 2) simpler metric strategy that use the metric given in (232); 3) nearest neighbor strategy given in [73] that select the nearest neighbor as the relay node; and 4) farthest neighbor strategy that that select the farthest neighbor as the relay node. In Fig. 59, the outage probability of the four strategies are provided as functions of the TX SNR and the traffic load (λ). As expected, the optimal strategy achieves the lowest outage probability in all conditions. The simpler metric strategy has higher outage probability than the optimal strategy but much lower than the nearest and farthest neighbor strategy.

CHAPTER V

CONCLUSION

WUSNs enable a wide variety of novel applications, such as intelligent irrigation, smart underground power grid, border patrol and intruder detection, assisted navigation, sports field maintenance, and leakage detection of underground pipelines and tanks. Despite the potential advantages, the harsh underground environments are challenging for wireless communications. Hence, existing communication and networking solutions and protocols for terrestrial WSNs have to be reexamined and redesigned.

In this thesis, we analyze the unique characteristics of the WUSNs in two underground environments, i.e. soil medium and underground mines/tunnels. For WUSNs in soil medium, we develop the WUSNs based two types of signal propagation techniques based on either EM wave-based or MI to overcome the challenges brought by the soil medium. Corresponding networking solutions are also developed for each signal propagation technique. For WUSNs in underground mines and tunnels, we optimize the MIMO system and cooperative communication system based on a comprehensive channel model to establish reliable and efficient communications.

In Chapter 2, the WUSNs in soil medium based on EM wave technique are analyzed.

- First, a complete channel model that can characterize all the three types of EM wave channels in soil medium are developed. The transmission ranges of the three types of channels are analytically expressed. The analysis shows that the range of UG-UG channel is the smallest ($\leq$ 5 m) among the three channels, while the range of UG-AG and AG-UG channel are in the range of 10 m to 50 m, depending on the soil water content and the sensor burial depth. The range of UG-AG channel is larger than the AG-UG channel due to the reflection and refraction on the air-ground interface.

The soil water content has significant influences on all the three types of channels in WUSNs. The sensor burial depth only affects the UG-AG channel and AG-UG channel while does not dramatically influence the UG-UG channel. Beside the soil water content and the sensor burial depth, the antenna height of the AG sinks also has obvious effect on the AG-UG channel.

- Second, the dynamic network connectivity in WUSNs is theoretically investigated based on the channel model of EM waves in soil medium. The soil transmission medium and the heterogeneous network architecture make the WUSN connectivity analysis much more complex than the terrestrial wireless sensor networks. The upper and lower bounds of the connectivity probability are analytically developed to provide the necessary and sufficient conditions to achieve the full-connected network, which give the guidelines to design the system parameters of the WUSNs according to the environmental conditions. The theoretical bounds are validated by simulations in various scenarios. Our analysis quantitatively captures the effects of multiple system and environmental parameters on the WUSN connectivity, including the UG sensor density, the AG fixed sink density, the number of AG mobile sinks, the UG sensor burial depth, the AG sink antenna height, the soil moisture, the tolerable latency, the mobility of the AG mobile sinks, and the AG sink antenna height.

- Third, a spatio-temporal correlation-based data collection scheme is proposed to release the unfeasible sensor density requirement of the full connectivity in WUSNs. More importantly, an explicit solution for the sensor density optimization in WUSNs under the proposed data collection scheme, which can guarantee the overall monitoring accuracy with minimum deployment/maintenance cost. To formalize the optimization solution, the dynamic underground channel conditions, the spatio-temporal correlation, the network connectivity, and the random or controlled mobility of mobile sinks are jointly analyzed. Our optimization solution clearly reveals the tradeoff

between the optimal sensor density and the number and mobility of the mobile sinks. Moreover, the effects of multiple environmental parameters, including the sensor burial depth, the soil water content, and the spatio-temporal model of the monitored physical quantity, are explicitly captured. The developed optimization solution provides a flexible tool to determine the optimal sensor density for different application requirements and environmental conditions in WUSNs.

In Chapter 3, the WUSNs in soil medium based on MI technique are investigated.

- First, an analytical model is developed to characterize the underground MI communication channel. The MI technique has constant channel condition because its path loss only depends on the permeability of the propagation medium, which remains the same if the medium is air, water and most kinds of soil and rock. However, the path loss of a simple MI communication system is still high since magnetic field strength falls off much faster than the EM waves in most environments. In soil medium, although it is known that the soil absorption causes high signal attenuation in the EM waves systems but does not affect the MI systems, the total path loss of the MI system is still similar to that of the EM waves system.

- Second, based on the MI channel analysis, we develop the MI waveguide technique to significantly enlarge the transmission range. In underground environments, the path loss of the MI system is slightly smaller than the EM wave system in normal and wet soil medium; while in very dry soil, the EM wave system has smaller path loss. However, due to the high path loss, both the systems can not provide a transmission range that is more than 10 m, which prevent them from practical applications. To this end, the MI waveguide is developed. The MI waveguide technique can greatly reduce the path loss, which is attributed to the relay coils deployed between the transceivers. It should be noted that the relay coils do not consume any energy and the cost is very low. The relay distance is also larger than the maximum transmission range of

the EM wave system. With the help of MI relay coils, the transmission range of the MI waveguide system is increased dramatically compared with that of the ordinary MI system and the EM wave system. Although the bandwidth of the MI system and the MI waveguide system is only 1 to 2 KHz, which is much smaller than the EM wave system, it is enough for the low data rate monitoring applications of WUSNs. Another advantage of the MI and MI waveguide system is that, as the transmission range increases, the transmission power decreases simultaneously with the received power, which is favorable for the energy-constrained WUSNs.

- Third, the deployment of those MI waveguides to connect the underground sensors in the WUSNs is investigated. In 1D WUSNs, the optimal number of relay coils between two adjacent sensors is derived. Based on the results in 1D WUSNs, two solutions to deploy the MI waveguides in 2D WUSNs is developed. To minimize the relay coil number, the MST algorithm is provided. The MST algorithm use the minimum spanning tree to connect all the sensors with the optimal relay coil number. However, the WUSN constructed by the MST algorithm is 1-connected hence is not robust to sensor failure. To enhance the network robustness while not increasing the relay coil number too much, the TC algorithm is proposed. The TC algorithm first use the Voronoi diagram to partition the whole network into non-overlapping triangle cells. Then the MI waveguides with the shape of the three-pointed star is deployed in every other triangle cells. The network constructed by the TC algorithm is k-connected ($k \geq 3$). Hence, the TC algorithm is robust to sensor failures.

In Chapter 4, the WUSNs deployed in underground mines and tunnels are explored.

- First, the propagation of EM waves in empty underground mines and tunnels is modeled. The typical structures of current underground mines and road tunnels is analyzed. The channels in those environments are divided into two cases: tunnel channel and room-and-pillar channel. The multimode model is developed to characterize

both tunnel channel and room-and-pillar channel. Based on the multimode model, our analysis shows that:

  - For tunnel environment:

    1) Due to the combination of multiple modes, high signal attenuation and intense fluctuation occur in the near region. The fall in the received power is gradual in the far region because the higher order modes attenuate very fast as the distance increases. The division of near and far region depends on the operation frequency, the tunnel size and the transmitter positions, which is quantitatively analyzed for the first time.

    2) The attenuation is mostly determined by the tunnel size and operating frequency, while the power distribution among modes is governed by the position of the transmitter's antenna.

    3) The delay spread of the wideband signals is determined by how many significant modes exist and how long the transmission distance is.

    4) In their normal value range, the humidity, pressure and temperature of the tunnel air, as well as the material of tunnel walls have little influence on the signal propagation in tunnels.

  - For room-and-pillar environment:

    Signal propagation has similar characteristics as in the tunnel case. The difference is that signals experience significant fluctuations in both near and far regions. Additionally, the operating frequency, the room height, the antenna position/polarization and the electrical parameters have much smaller influence on the channel characteristics in the room-and-pillar environment than in the tunnels.

- Second, an analytical channel model for the tunnel with a vehicular traffic flow is developed. Specifically, the propagation modes experience in-mode loss and mode

coupling around a vehicle. The UTD method, Poisson sum formula and saddle point method are utilized to calculate the in-mode loss and mode coupling coefficients. Then based on the deterministic model, a statistical model is developed to characterize the signal propagation in a tunnel with random vehicular traffic flow. Instead of using the deterministic vehicular traffic information, the traffic flow theory and the vehicle size distribution model is utilized. The entries in the propagation matrix are theoretically proved to be normally distributed. In addition, the means and variables of the entries are analytically calculated. Based on the proposed channel model, our analysis shows that: the vehicles in tunnels induce additional path losses and signal fluctuations. In the tunnel with deterministic vehicular traffic flow, the number and size of the vehicle, as well as the size of the tunnel determine the intensity of the vehicles' influence. In the tunnel with random vehicular traffic flow, the average vehicle arriving rate, the average vehicle velocity, and the size of the tunnel determine the additional loss and signal fluctuation caused by the vehicular traffic flow.

- Third, the capacity and outage behavior of the MIMO and cooperative communication systems is investigated based on the underground tunnel channel model. For MIMO system, it is proved that the MIMO capacity in underground tunnels follows either a normal distribution in high SNR regime or a log-normal distribution in low SNR regime. The ergodic and outage capacity of MIMO systems in tunnels are explicitly expressed as functions of tunnel size, transmission power, vehicular traffic load, and MIMO antenna geometry. Then the optimal MIMO antenna geometry design scheme is proposed, which obviously increases the outage capacity. For cooperative communication systems, the outage probability of such systems in underground tunnels is calculated. Based on the formula of the outage probability, we proposed an outage-optimal cooperative relay assignment protocol, which significantly outperforms the existing relay assignment protocols in term of outage probability in underground tunnel environments.

In the future, for WUSNs in soil medium, we plan to extend our work in several directions, including 1) an information theoretical study of the channel and network capacity, especially the MI waveguide-based WUSNs; 2) the protocol development that involves close coupling between the unique soil medium conditions and the system design, leading to a cross-layer communication and coordination module; 3) the evaluation testbed development to verify the theoretical result through field experiments; and 4) the development of a hybrid WUSN system using both EM Waves and MI Waveguides that takes advantages of both techniques.

For WUSNs in underground mines and tunnels, so far we have developed the communication strategies for normally operating mines and tunnels. However, no existing techniques support communications and localization after mine disasters or tunnel collapse. To this end, we plan to investigate the MI-based WUSNs in underground mines and tunnels to achieve robust wireless communications and localization for survivor rescue after mine disasters.

REFERENCES

[1] "Coping with water scarcity - challenge of the twenty first century," tech. rep., United Nations - WorldWater Day 2007 Report, 2007.

[2] Akyildiz, I. F. and Stuntebeck, E. P., "Wireless underground sensor networks: Research challenges," *Ad Hoc Networks Journal (Elsevier)*, vol. 4, pp. 669–686, July 2006.

[3] Akyildiz, I. F., Su, W., Sankarasubramaniam, Y., and Cayirci, E., "Wireless sensor networks: A survey," *Computer Networks (Elsevier) Journal*, vol. 38, pp. 393–422, March 2002.

[4] Akyildiz, I. F., Sun, Z., and Vuran, M. C., "Signal propagation techniques for wireless underground communication networks," *Physical Communication Journal (Elsevier)*, vol. 2, pp. 167–183, September 2009.

[5] Arshad, K., Katsriku, F., and Lasebae, A., "Modelling obstructions in straight and curved rectangular tunnels by finite element approach," *Journal of Electrical Engineering*, vol. 59, no. 1, pp. 9–13, 2008.

[6] ASTM International, ASTM Standard B 258-02, *Standard specification for standard nominal diameters and cross-sectional areas of AWG sizes of solid round wires used as electrical conductors*, 2002.

[7] Avestimehr, A. S. and Tse, D. N. C., "Outage capacity of the fading relay channel in the low-snr regime," *IEEE Transactions on Information Theory*, vol. 53, pp. 1401–1415, April 2007.

[8] AWE Communications, Technical Report, *Entwicklung eines Programmpakets fur die Berechnung der Ausbreitung elektromagnetischer Wellen in Tunnels*, http://www.awe-communications.com/Propagation/Tunnel/pdf/tunnel.pdf.

[9] Bansal, R., "Near-field magnetic communication," *IEEE Antennas and Propagation Magazine*, Apirl 2004.

[10] Beres, E. and Adve, R., "Selection cooperation in multi-source cooperative networks," *IEEE Transactions on Wireless Communications*, vol. 7, pp. 118–127, January 2008.

[11] Berg, M. D., Cheong, O., Kreveld, M., and Overmars, M., *Computational Geometry: Algorithms and Applications*. Berlin, New York: Springer, 3rd ed., 2000.

[12] Berkeley Varionics Systems, Inc., www.bvsystems.com, *YellowJacket wireless spectrum analyzer*.

[13] Bettstetter, C., "On the minimum node degree and connectivity of a wireless multihop network," in *Proc. ACM MOBIHOC '02*, (Lausanne, Switzerland), June 2002.

[14] Bettstetter, C. and Hartmann, C., "Connectivity of wireless multihop networks in a shadow fading environment," *ACM-Springer Journal of Wireless Networks (WINET)*, vol. 11, no. 5, pp. 571–579, 2005.

[15] Billingsley, P., *Probability and Measure*. New York: John Wiley and Sons, 3rd ed., 1995.

[16] Bogena, H. R., Huisman, J. A., Meier, H., Rosenbaum, U., and Weuthen, A., "Hybrid wireless underground sensor networks: Quantification of signal attenuation in soil," *Vadose Zone J.*, vol. 8, pp. 755–761, August 2009.

[17] Borovikov, V. and Kinber, B. Y., *Geometrical theory of diffraction*. London: Institution of Electrical Engineers, 1994.

[18] Bunszel, C., "Magnetic induction: a low-power wireless alternative," *RF Design*, vol. 24, pp. 78–80, November 2001.

[19] Cai, H. and Eun, D. Y., "Crossing over the bounded domain: From exponential to power-law inter-meeting time in mobile ad-hoc networks," *IEEE/ACM Transactions on Networking*, vol. 17, pp. 1578–1591, October 2009.

[20] Camp, T., Boleng, J., and Davies, V., "A survey of mobility models for ad hoc network research," *Wireless Communications and Mobile Computing*, vol. 2, pp. 483–502, September 2002.

[21] Cerasoli, C., "Rf propagation in tunnel environments," in *Proc. IEEE Military Communications Conference MILCOM'04*, vol. 1, pp. 363–369, November 2004.

[22] Charles, L. A. and Kenneth, W. A., *Electronic engineering*. New York: Wiley, 3rd ed., 1973.

[23] Chehri, A., Fortier, P., and Tardif, P. M., "Security monitoring using wireless sensor networks," in *Proc. Communication Networks and Services Research CNSR '07*, pp. 13–17, May 2007.

[24] Chen, C. H., Chiu, C., Hung, S., and Lin, C., "Ber performance of wireless bpsk communication system in tunnels with and without traffic," *Wireless Personal Communications*, vol. 30, pp. 1–12, July 2004.

[25] Chen, S. and Jeng, S., "Br image approach for radio wave propagation in tunnels with andwithout traffic," *IEEE Trans. on Vehicular Technology*, vol. 45, pp. 570–578, August 1996.

[26] Cocheril, Y., Langlais, C., Berbineau, M., and Moniak, G., "Advantages of simple mimo schemes for robust or high data rate transmission systems in underground tunnels," in *Proc. IEEE VTC 2008-Fall*, (Calgary, Alberta), September 2008.

[27] Collin, R., *Field theory of guided waves*. New York: IEEE Press, 1991.

[28] Deift, P. and Zhou, X., "A steepest descent method for oscillatory riemann-hilbert problems. asymptotics for the mkdv equation," *Ann. of Math.*, vol. 137, no. 2, pp. 295–368, 1993.

[29] Delogne, P., *Leaky Feeders and Subsurface Radio Communications*. Stevenage Herts, New York: P. Peregrinus, August 1982.

[30] Dong, X. and Vuran, M. C., "Spatio-temporal soil moisture measurement with wireless underground sensor networks," in *Proc. IFIP Med-Hoc-Net '10*, (Juan-les-pins, France), June 2010.

[31] Dousse, O., Mannersalo, P., and Thiran, P., "Latency of wireless sensor networks with uncoordinated power saving mechanisms," in *Proc. ACM MOBIHOC*, (Tokyo, Japan), May 2004.

[32] Dousse, O., Thiran, P., and Hasler, M., "Connectivity in ad-hoc and hybrid networks," in *Proc. IEEE INFOCOM '02*, (New York, USA), June 2002.

[33] Dudley, D. G., Lienard, M., Mahmoud, S. F., and Degauque, P., "Wireless propagation in tunnels," *IEEE Antenna and Propagation Magazine*, vol. 49, pp. 11–26, Apirl 2007.

[34] Emslie, A. G., Lagace, R. L., and Strong, P. F., "Theory of the propagation of uhf radio waves in coal mine tunnels," *IEEE Trans. on Antenna and Propagation*, vol. AP-23, pp. 192–205, March 1975.

[35] Fabbri, F. and Verdone, R., "A statistical model for the connectivity of nodes in a multi-sink wireless sensor network over a bounded region," in *Proc. European Wireless Conference (EW'08)*, (Prague, Czech Republic), June 2008.

[36] Frankl, D. R., *Electromagnetic Theory*. Prentice-Hall, 1986.

[37] Gerlough, D. L. and Huber, M. J., *Traffic Flow Theory: A Monograph*. Washington: Transportation Research Board, National Research Council, 1975.

[38] Gertsch, R. E. and Bullock, R. L., *Techniques in underground mining: selections from Underground mining methods handbook*. Littleton, CO : Society for Mining, Metallurgy, and Exploration, 1998.

[39] Goovaerts, P., *Geostatistics for Natural Resources Evaluation*. New York: Oxford University Press, 1997.

[40] Gupta, P. and Kumar, P. R., "Critical power for asymptotic connectivity," in *IEEE CDC*, (Tampa, USA), December 1998.

[41] Hwang, Y., Zhang, Y. P., and Kouyoumjian, R. G., "Ray-optical prediction of radio-wave propagation characteristics in tunnel environments: Theory, analysis and measurements," *IEEE Trans. on Antenna and Propagation*, vol. 46, pp. 1328–1345, September 1998.

[42] Hyytia, E., Lassila, P., and Virtamo, J., "Spatial node distribution of the random waypoint mobility model with applications," *IEEE Trans. on Mobile Computing*, vol. 5, pp. 680–694, June 2006.

[43] Isham, V., Cox, D., Rodríguez-Iturbe, I., Porporato, A., and Manfreda, S., "Representation of space-time variability of soil moisture," *Proc. Royal Society*, vol. 461, no. 2064, pp. 4035–4055, 2005.

[44] Jack, N. and Shenai, K., "Magnetic induction ic for wireless communication in rf-impenetrable media," in *Proc. IEEE Workshop on Microelectronics and Electron Devices (WMED'07)*, April 2007.

[45] Kalinin, V. A., Ringhofer, K., and Solymar, L., "Journal of applied physics," *Journal of Applied Physics*, vol. 92, no. 10, pp. 6252–6261, 2002.

[46] Klemenschits, T., Scholtz, A., and Bonek, E., "Microwave measurements in scaled road tunnels modeling 900 mhz propagationin," in *Proc. IEEE VTC '93*, (Secaucus, NJ, USA), May 1993.

[47] Kong, Z. and Edmund, M. Y., "Connectivity and latency in large-scale wireless networks with unreliable links," in *Proc. IEEE INFOCOM '08*, (Phoenix, USA,), April 2008.

[48] Kouyoumjian, R. and Pathak, P., "A uniform geometrical theory of diffraction for an edge in a perfectly conducting surface," *Proceedings of the IEEE*, vol. 62, pp. 1448–1461, November 1974.

[49] Kraus, J. D. and Fleisch, D. A., *Electromagnetics*. New York: McGraw-Hill, 5th ed., 1999.

[50] Laakmann, K. D. and Steier, W. H., "Waveguides: Characteristic modes of hollow rectangular dielectric waveguides," *Applied Optics*, vol. 15, pp. 1334–1340, May 1976.

[51] Laneman, J. N., Tse, D. N. C., and Wornell, G. W., "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Transactions on Information Theory*, vol. 50, pp. 3062–3080, December 2004.

[52] Lee, J. and Bertoni, H. L., "Coupling at cross, t, and l junctions in tunnels and urban street canyons," *IEEE Transactions on Antenna and Propagation*, vol. 51, pp. 1403–1412, May 2003.

[53] Li, L., Vuran, M. C., and Akyildiz, I. F., "Characteristics of underground channel for wireless underground sensor networks," in *Proc. Med-Hoc-Net '07*, (Corfu, Greece), June 2007.

[54] Lienard, M., Betrencourt, S., and Degauque, P., "Propagation in road tunnels: a statistical analysis of the field distribution and impact of the traffic," *Annals of Telecommunications*, vol. 55, pp. 623–631, November 2000.

[55] Lienard, M. and Degauque, P., "Mobile telecommunication in mine: Characterization of the radio channel," in *Proc. IEEE Electrotechnical Conference MELECON apos '96*, vol. 3, pp. 1663–1665, May 1996.

[56] Lienard, M. and Degauque, P., "Natural wave propagation in mine environments," *IEEE Transactions on Antenna and Propagation*, vol. 48, pp. 1326–1339, September 2000.

[57] Limpert, E., Stahel, W., and Abbt, M., "Log-normal distributions across the sciences: Keys and clues," *BioScience*, vol. 51, pp. 341–352, May 2001.

[58] Loyka, S., "Multiantenna capacities of waveguide and cavity channels," *IEEE Transactions on Vehicular Technology*, vol. 54, pp. 863–872, May 2005.

[59] Mahmoud, S. F. and Wait, J. R., "Geometrical optical approach for electromagnetic wave propagation in rectangular mine tunnels," *Radio Science*, vol. 9, pp. 1147–1158, December 1974.

[60] Marzetta, T. L. and Hochwald, B. M., "Capacity of a mobile multiple-antenna communication link in rayleigh flat fading," *IEEE Transactions on Information Theory*, vol. 45, pp. 139–157, January 1999.

[61] Medhi, J., *Stochastic processes*. New York: J. Wiley, 2nd ed., 1994.

[62] Meester, R. and Roy, R., *Continuum percolation*. Cambridge University Press, 1996.

[63] Mehta, N., Wu, J., Molisch, A., and Zhang, J., "Approximating a sum of random variables with a lognormal," *IEEE Transactions on Wireless Communications*, vol. 6, pp. 2690–2699, July 2007.

[64] Molina-Garcia-Pardo, J. M., Lienard, M., Degauque, P., Dudley, D. G., and Juan-Llacer, L., "Interpretation of mimo channel characteristics in rectangular tunnels from modal theory," *IEEE Transactions on Vehicular Technology*, vol. 57, pp. 1974–1979, May 2008.

[65] Moustakas, A. L. and Simon, S. H., "On the outage capacity of correlated multiple-path mimo channels," *IEEE Transactions on Information Theory*, vol. 53, pp. 3887–3903, November 2007.

[66] Murray, R. M., "Recent research in cooperative control of multivehicle systems," *Journal of Dynamic Systems, Measurement, and Control*, vol. 129, pp. 571–583, September 2007.

[67] Palermo, V., "Near-field magnetic comms emerges," *Electronic Engineering Times*, November 2003.

[68] Peplinski, N., Ulaby, F., and Dobson, M., "Dielectric properties of soils in the 0.3-1.3-ghz range," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 33, pp. 803–807, May 1995.

[69] Popov, A. V. and Zhu, N. Y., "Modeling radio wave propagation in tunnels with a vectorial parabolic equation," *IEEE Transactions on Antenna and Propagation*, vol. 48, pp. 1403–1412, September 2000.

[70] Porrat, D., *Radio Propagation in Hallways and Streets for UHF Communications*. PhD thesis, Stanford University, 2002.

[71] Proakis, J. G., *Digital Communications*. New York: McGraw-Hill, 4th ed., 1995.

[72] Rajagopalan, R. and Varshney, P. K., "Connectivity analysis of wireless sensor networks with regular topologies in the presence of channel fading," *IEEE Trans. on Wireless Communications*, vol. 8, no. 7, pp. 3475–3483, 2009.

[73] sadek, A. K., Han, Z., and Liu, K. J. R., "Distributed relay-assignment protocols for coverage expansion in cooperative wireless networks," *IEEE Transactions on Mobile Computing*, vol. 9, pp. 505–515, April 2010.

[74] Santi, P. and Blough, D. M., "The critical transmitting range for connectivity in sparse wireless ad hoc networks," *IEEE Trans. Mobile Computing*, vol. 2, no. 1, pp. 25–39, 2003.

[75] Scaglione, A., Goeckel, D. L., and Laneman, J. N., "Cooperative communications in mobile ad hoc networks," *IEEE Signal Processing Magazine*, vol. 23, pp. 18–29, September 2006.

[76] Shi, Y., Sharma, S., Hou, Y. T., and Kompella, S., "Optimal relay assignment for cooperative communications," in *Proc. ACM MobiHoc '08*, (Hong Kong SAR, China), May 2008.

[77] Shiu, D., Foschini, G. J., Gans, M. J., and Kahn, J. M., "Fading correlation and its effect on the capacity of multielement antenna systems," *IEEE Transactions on Communications*, vol. 48, pp. 502–513, March 2000.

[78] Silva, A. R. and Vuran, M. C., "Communication with aboveground devices in wireless underground sensor networks: An empirical study," in *Proc. IEEE ICC '10*, (Cape Town, South Africa), May 2010.

[79] Silva, A. R. and Vuran, M. C., "Development of a testbed for wireless underground sensor networks," *EURASIP Journal on Wireless Communications and Networking (JWCN)*, 2010.

[80] Sojdehei, J., Wrathall, P., and Dinn, D., "Magneto-inductive (mi) communications," in *Proc. MTS/IEEE Conference and Exhibition (OCEANS'01)*, November 2001.

[81] Strang, G., *Linear algebra and its applications*. Belmont, CA: Thomson, Brooks/Cole, 4th ed., 2006.

[82] Stuber, G. L., *Principles of mobile communication*. Boston: Kluwer Academic, 2001.

[83] Sun, M., Hackworth, S., Tang, Z., Gilbert, G., Cardin, S., and Sclabassi, R. J., "How to pass information and deliver energy to a network of implantable devices within the human body," in *Proc. IEEE Conference on Engineering in Medicine and Biology Society (EMBS'07)*, pp. 5286–5289, August 2007.

[84] Sun, Z. and Akyildiz, I. F., "Channel modeling of wireless networks in tunnels," in *Proc. IEEE GLOBECOM 2008*, (New Orleans, USA), November 2008.

[85] Sun, Z. and Akyildiz, I. F., "Optimal antenna geometry analysis of mimo communication systems in underground tunnels," in *Proc. IEEE GLOBECOM 2009*, (Honolulu, USA), November 2009.

[86] Sun, Z. and Akyildiz, I. F., "Channel modeling and analysis for wireless networks in underground mines and road tunnels," *IEEE Transactions on Communications*, vol. 58, pp. 1758–1768, June 2010.

[87] Sun, Z. and Akyildiz, I. F., "Connectivity in wireless underground sensor networks," in *Proc. of IEEE SECON 2010*, (Boston, MA, USA), June 2010.

[88] Sun, Z. and Akyildiz, I. F., "Influences of vehicles on signal propagation in road tunnels," in *Proc. of IEEE ICC 2010*, (Cape Town, South Africa), May 2010.

[89] Sun, Z. and Akyildiz, I. F., "Magnetic induction communications for wireless underground sensor networks," *IEEE Transactions on Antenna and Propagation*, vol. 58, pp. 2426–2435, July 2010.

[90] Sun, Z. and Akyildiz, I. F., "A mode-based approach for channel modeling in underground tunnels under the impact of vehicular traffic flow," *submitted for journal publication*, April 2010.

[91] Sun, Z., Wang, P., Vuran, M. C., Al-Rodhaan, M. A., Al-Dhelaan, A. M., and Akyildiz, I. F., "Bordersense: Border patrol through advanced wireless sensor networks," *Ad Hoc Networks Journal (Elsevier)*, vol. 9, pp. 468–477, May 2011.

[92] Sun, Z., Wang, P., Vuran, M. C., Al-Rodhaan, M. A., Al-Dhelaan, A. M., and Akyildiz, I. F., "Mise-pipe: Magnetic induction-based wireless sensor networks for underground pipeline monitoring," *Ad Hoc Networks Journal (Elsevier)*, vol. 9, pp. 218–227, May 2011.

[93] Syms, R. R. A., Young, I. R., and Solymar, L., "Low-loss magneto-inductive waveguides," *Journal of Physics D: Applied Physics*, vol. 39, pp. 3945–3951, 2006.

[94] Syms, R., Shamonina, E., and Solymar, L., "Magneto-inductive waveguide devices," *Proc. IEE Microwaves, Antennas and Propagation*, vol. 153, no. 2, pp. 111–121, 2006.

[95] Taflove, A. and Hagness, S. C., *Computational Electrodynamics: The Finite-Difference Time-Domain Method*. Norwood, MA: Artech House, 2005.

[96] Telatar, I. E., "Capacity of multi-antenna gaussian channels," *European Transactions on Telecommunications*, vol. 10, pp. 585–595, 1999.

[97] Telford, W. M., Geldart, L. P., and Sheriff, R. E., *Applied Geophysics*. Cambridge University Press, 2nd ed., 1990.

[98] Tiusanen, J., "Wireless soil scout prototype radio signal reception compared to the attenuation model," *Precision Agriculture*, November 2008.

[99] Vuran, M. C., Akan, O. B., and Akyildiz, I., "Spatio-temporal correlation: Theory and applications for wireless sensor networks," *Computer Networks Journal (Elsevier)*, vol. 45, pp. 245–259, June 2004.

[100] Vuran, M. C., Akyildiz, I. F., and Al-Dhelaan, A. M., "Channel modeling and analysis for wireless underground sensor networks in soil medium," *to appear in Physical Communication Journal (Elsevier)*, 2010.

[101] Wang, G., Cao, G., and LaPorta, T. F., "Movement-assisted sensor deployment," *IEEE Trans. on Mobile Computing*, vol. 5, pp. 640–652, June 2006.

[102] Wiltshire, M. C. K., Pendry, J. B., Young, I. R., Larkman, D. J., Gilderdale, D. J., and Hajnal, J. V., "Microstructured magnetic materials for rf flux guides in magnetic resonance imaging," *Science*, vol. 291, pp. 849–851, February 2001.

[103] Wiltshire, M., Shamonina, E., Young, I., and Solymar, L., "Dispersion characteristics of magneto-inductive waves: comparison between theory and experiment," *Electronics Letters*, vol. 39, no. 2, pp. 215–217, 2003.

[104] Wisniewski, S., *Wireless and cellular networks*. Upper Saddle River, N.J.: Pearson/Prentice Hall, 2005.

[105] Wu, B. Y. and Chao, K.-M., *Spanning Trees and Optimization Problems*. Boca Raton, FL: Chapman and Hall / CRC, 2004.

[106] X. Xu, J. L. and Zhang, Q., "Delay tolerant event collection in sensor networks with mobile sink," in *Proc. IEEE INFOCOM '10*, (San Diego, USA), Apirl 2010.

[107] Yamaguchi, Y., Abe, T., and Sekiguchi, T., "Radio wave propagation loss in the vhf to microwave region due tovehicles in tunnels," *IEEE Trans. on Electromagnetic Compatibility*, vol. 31, pp. 87–91, February 1989.

[108] Yee, H. Y., Felsen, L. B., and Keller, J. B., "Ray theory of reflection from the open end of a waveguide," *SIAM J. Appl. Math.*, vol. 16, pp. 268–300, March 1968.

[109] Yee, H. Y., Felsen, L. B., and Keller, J. B., "Ray theory of reflection from the open end of awaveguide," *SIAM J. Appl. Math.*, vol. 16, pp. 268–300, March 1968.

[110] Zhang, S., "The multipath propagation model of rectangular tunnel channel," in *Proc. IEEE TENCON 2002*, vol. 2, pp. 1016–1019, October 2002.

[111] Zhang, Y. P. and Hong, H. J., "Ray-optical modeling of simulcast radio propagation channels in tunnels," *IEEE Trans. on Vehicular Technology*, vol. 53, pp. 1800–1808, November 2004.

[112] Zhang, Y. P., Zheng, G. X., and Sheng, J. H., "Excitation of uhf radio waves in tunnels," *Microwave and Optical Technology Letters*, vol. 22, pp. 408–410, Augest 1999.

[113] Zou, Y. and Chakrabarty, K., "Sensor deployment and target localization based on virtual forces," in *Proc. IEEE INFOCOM '03*, (San Francisco, USA), March 2003.

VITA

Zhi Sun was born on November 9, 1981, in Beijing, China. He received the B.S. degree in Communication Engineering from Beijing University of Posts and Telecommunications (BUPT), and the M.S. degree in Electronic Engineering from Tsinghua University, Beijing, China, in 2004 and 2007, respectively. He received Ph.D. degree in Electrical and Computer Engineering from Georgia Institute of Technology, Atlanta, GA. in 2011, under the supervision of Prof. Ian F. Akyildiz. His current research focuses on wireless sensor networks and cyber physical systems, specifically physical layer and cross-layer design and optimization, performance analysis and modeling, and systems-level implementation of wireless sensor and actuator networks in challenged environments, such as in underground soil medium, in underground mines and tunnels, in sandstorms, among others. Zhi Sun received the Best Paper Award for "Deployment Algorithms for Wireless Underground Sensor Networks using Magnetic Induction" in IEEE GLOBECOM 2010. He also received the BWN Researcher of the Year Award at the Georgia Institute of Technology in 2009.