4th International Conference on Mathematical Modeling in Physical Sciences (IC-MSquare2015) IOP Publishing
Journal of Physics: Conference Series **633** (2015) 012001 doi:10.1088/1742-6596/633/1/012001

Forecasting systemic transitions in high dimensional stochastic complex systems

Duccio Piovani[1], Jelena Grujic[2] and Henrik Jeldtoft Jensen[1]

[1]Centre for Complexity Science and Department of Mathematics, Imperial College London, South Kensington Campus, SW7 2AZ, UK

E-mail: duccio.piovani@gmail.com,jelena.grujic@gmail.com,h.jensen@imperial.ac.uk

Abstract. We briefly describe a new procedure to monitor and forecast transitions occurring in high dimensional complex systems. The method is illustrated by applications to two model systems: firstly to the Tangled Nature Model of evolutionary ecology and secondly to a stochastic replicator system. The quasi-stable configurations of the stochastic dynamics are taken as input for a stability analysis of the deterministic mean field approximation of the dynamics. We demonstrate that the largest overlap between the observed configuration and the unstable eigendirections serves as a precursor that allows us to forecast transitions with an efficiency of about 80% even if we only know the couplings matrix describing the dynamics to with 10% accuracy.

1. Introduction

Many types of complex systems are prone to sudden transitions that are challenging to forecast[1]. High dimensional complex systems both physical and biological exhibit intermittent dynamical evolution consisting of stretches of relatively little change interrupted by often sudden and dramatic transitions to a new meta-stable configurations[1, 2].

These transitions may change the state of the affairs in dramatic ways, think for instance of ecological catastrophes or crashes in financial markets. Forecasting could help to alleviate the consequences. Different types of complex systems will have different types of dynamics and each type may very well need specific types of analysis. Scheffer and collaborators have developed a method pertinent to systems in which the transition takes the form as a bifurcation captured by a robust macroscopic variable[2].

We describe an alternative scenario, suggested recently in [3], in which the transitions are induced by intrinsic fluctuations at the microscopic level of the individual components. Our approach is relevant to systems in which the dynamics creates new hitherto components that can perturb the existing order. One may think of a new and more virulent virus being created through a mutation of an existing strain (e.g. the SARS virus in 2003), or a new economic agent arriving in the market (e.g. the dot-com bubble in 1997-2000).

We exemplify the methodology by applications to to two models: (1) The Tangled Nature (TaNa) model of evolutionary ecology[4]. This model has had considerable success in reproducing both macro-evolutionary aspects such as the intermittent mode of extinctions[5] and ecological aspects such as species abundance distributions[6] and species area laws[7].(2) A high dimensional replicator with a stochastic element of mutation[8, 9].


Content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.
Published under licence by IOP Publishing Ltd

We recall that the stochastic replicator system is an example of the generic Langevin approach is a generic way to including a stochastic element to otherwise deterministic dynamics, see e.g. [10] and that the deterministic version of the replicator dynamics used routinely in a large variety of applications, not least because of its relation to game theory[11, 12]. This makes us hopeful that successful application of our method to the stochastic replicator indicates a broad relevance of the way we identify and analyse precursors of endogenous transitions.

2. Methodology

The procedure consists in a stability analysis of a deterministic approximation of the dynamics. If the time evolution consists of a set of stochastic processes (as they do for the Tangled Nature model, see below) a mean filed dynamics will have to be established. In the case of Langevin dynamics one simply needs to eliminate the stochastic part of the dynamics (this is the situation for the stochastic replicator system).

We define a state vector $\vec{n}(t) = (n_1(t), \ldots, n_d(t))$ and consider deterministic dynamics of the form

$$\vec{n}(t+1) - \vec{n}(t) \simeq \mathbb{T}(\vec{n}(t))\vec{n}(t) \tag{1}$$

where the matrix $\mathbb{T}$ is given by the mean field evolution matrix of derived from the stochastic dynamics.

For intermittent dynamics Eq. (10) will support periods during which little change occur from time step to time step. I.e. $\vec{n}(t+1) \simeq \vec{n}(t) \simeq \vec{n}^*$. This corresponds to configurations for which the right hand side of eq.(10) will essentially be zero. These are to a good approximation the fixed point configurations of our dynamical system. Since we are interested in very high dimensional systems, it is typically impossible to solve $\mathbb{T}(\vec{n}(t))\vec{n}(t) = 0$. Instead we use as an approximation the observed metastable state during which $\vec{n}(t+1) \simeq \vec{n}(t)$. We do a local time average of the observed $\vec{n}(t)$ and denote this average by $\vec{n}^*$. We then linearise the dynamics around this observed state by expressing the configuration vector as $\vec{n}(t) = \vec{n}^* + \delta\vec{n}(t)$

$$\begin{aligned} \delta\vec{n}(t+1) - \delta\vec{n}(t) &= \left(\mathbb{T}(\vec{n}^*) + \partial_n \mathbb{T}(\vec{n}^*)\vec{n}^*\right) \delta\vec{n}(t) \\ &= \mathbb{M}(\vec{n}^*) \cdot \delta\vec{n} \end{aligned} \tag{2}$$

and extract the stability matrix $\mathbb{M}(\vec{n}^*)$

$$\mathbb{M}(\vec{n}^*) = \left(\mathbb{T}(\vec{n}^*) + \partial_n \mathbb{T}(\vec{n}^*)\vec{n}^*\right) \tag{3}$$

As usual the stability of the dynamics near $\vec{n}^*$ is determined to linear order by the nature of the eigenvalues of $\mathbb{M}(\vec{n}^*)$. We know that the eigenvectors or *generalised* eigenvectors (in case of a non diagonalizable Jacobian) $\vec{e}_+$ associated with λ with $\mathrm{Re}(\lambda) > 0$ indicate unstable directions. These can be identified with *toxic* components n_{t} of the configuration vector.

A sudden growth of these components would indicate the approach of a transition. We now use as a quantitative precursor of transitions

$$Q(t) = \max_{\lambda \in Sp^+(\mathbb{M}[\vec{n}^*])} \left| \langle \delta n(t), \vec{e}_\lambda \rangle \right| . \tag{4}$$

Here $\langle \delta\vec{n}(t)), \vec{e}_\lambda \rangle$ estimates the time evolution (over one time unit) of the overlap between the deviation away from the metastable state and the eigendirection $\vec{e}_\lambda$. $Q(t)$ is then the maximal overlap between this forward propagated deviation and the unstable eigendirections.

3. Applications to two models

3.1. The Tangled Nature Model

The dynamical entities of the TaNa model consist of agents represented by a sequence of binary variables with fixed length L [13], denoted as, where. We denote by $n(\mathbf{S}^a, t)$ the number of agents of type $\mathbf{S}^a$ at time t and the total population is $N(t) = \sum_{a=1}^{2^L} n(\mathbf{S^a}, \mathbf{t})$. A time step is defined as a succession of one annihilation and of one reproduction attempt. Annihilation consists of choosing an agent at random with uniform probability and remove the agent with probability p_{kill}, taken to be constant in time and independent on the type. Reproduction: choose with uniform probability an agent, $\mathbf{S^a}$, at random and duplicate the agent (and remove the mother) with probability: $p_{off}(\mathbf{S^a}, \mathbf{t}) = \frac{\exp(\mathbf{H(S^a,t)})}{\mathbf{1}+\exp(\mathbf{H(S^a,t)})}$ which depends on the occupancy distribution of all the types at time t through the weight function:

$$H(\mathbf{S^a}, \mathbf{t}) = \frac{\mathbf{k}}{\mathbf{N(t)}} \sum_{\mathbf{b}} \mathbf{J}(\mathbf{S^a}, \mathbf{S^b}) \mathbf{n_b(t)} - \mu \mathbf{N(t)} \tag{5}$$

In Eq. (5), the first term couples the agent $\mathbf{S^a}$ to one of type $\mathbf{S^b}$ by introducing the interaction strength $\mathbf{J}(\mathbf{S^a}, \mathbf{S^b})$, whose values are randomly distributed in the interval $[-1, +1]$. For simplification and to emphasize interactions we here assume: $\mathbf{J}(\mathbf{S^a}, \mathbf{S^a}) = \mathbf{0}$. The parameter k scales the interactions strength and μ can be thought of as the carrying capacity of the environment. An increase (decrease) in μ corresponds to harsher (more favourable) external conditions.

Mutations occur in the following way: For each of the two copies $\mathbf{S^{a_1}}$ and $\mathbf{S^{a_2}}$, a single mutation changes the sign of one of the genes: $S_i^{a_1} \to -S_i^{a_1}$, $S_i^{a_2} \to -S_i^{a_2}$ with probability p^{mut}. We define a generation to consist of $N(t)/p^{kill}$ time steps, i.e. the average time needed to kill all the individuals at time t. These microscopic rules generate intermittent macro dynamics[5] as showen in Fig. 1.

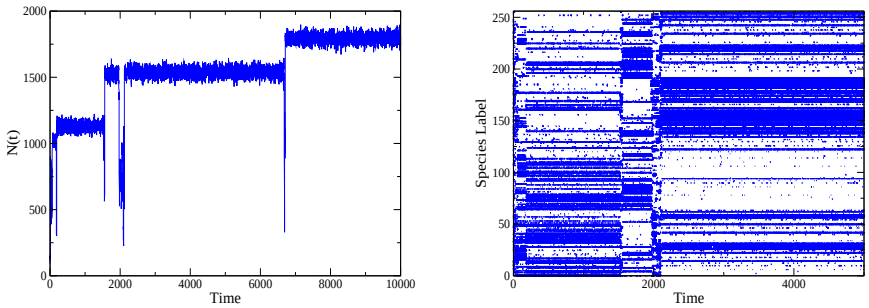


Figure 1. Left Panel: total population as a function of time (in generations) for a single realization of the TaNa. The punctuated dynamics is clearly visible: quasi-stable periods alternate with periods of hectic transitions, during which $N(t)$ exhibits large amplitude fluctuations. Right panel: occupancy distribution of the types. The genotypes are labelled arbitrarily and a blue (whilte) dot indicates a type which is (non) occupied at the time t. The lines therefore represent fit core species that fom the qESS, while the dots are the cloud of unfit species that are populated by mutations but rapidly go extinct.

3.1.1. Mean Field Description In this model the state vector $\vec{n}(t) = (n_1(t), \cdots, n_d(t))$ tells us how many individuals of each species are present at time t, therefore $n_i(t) \in \mathbb{N}$. The microscopic

dynamics of the TaNa model is fully stochastic: reproduction, mutations and deaths. The average of this dynamics leads [3, 14] to a mean field evolution matrix given by

$$T_{ij} = \left(\alpha p_j^{\text{off}}(t) - p^{kill}\right) \cdot \delta_{ij} + p_j^{\text{off}} \cdot p_{j\to i}^{\text{mut}} \cdot (1 - \delta_{ij})\,. \tag{6}$$

where $\alpha = (2p_o - 1)$, the corresponding stability matrix of the linearized dynamics is

$$\mathbb{M}_{ij} = (\alpha p_j^{\text{off}} - p^{kill})\delta_{ij} + 2(1 - \delta_{ij}) p_j^{\text{off}} p_{j\to i}^{\text{mut}} \tag{7}$$

$$+ \sum_k \left[\alpha\delta_{ik} + (1 - \delta_{ik}) \cdot p_{k\to i}^{mut}\right] \frac{\partial p_k^{\text{off}}}{\partial n_j} n_k^*,$$

from which we can determine the stability spectrum and compute the indicator $Q(t)$ by use of Eq. (4).

3.2. The Stochastic Replicator System

Describes a large number of agents involved in games and applying different strategies [12]. We consider a version in which strategies are removed from the system if the frequency with which appear are too low. We also include a stochastic element, which describes players that change their strategy[9].

The instantaneous state of the system is given by a configuration vector $\vec{n}$, which contains the frequency (hence $n_i \in [0, 1]$) of each possible strategy. There a total of d possible strategies. The outcome of the games are described by a dxd payoff matrix J. This matrix is constant. Each strategy is related to the other strategies by a set of payoffs, or interactions.

The deterministic part of the time evolution is done by first computing the *fitness*, $h_i(t) = \sum_j J_{ij} n_j(t)$ of each active strategy and compare it with the average fitness $\bar{h}(t) = \sum_{ij} J_{ij} n_i(t) n_j(t)$. Each frequency is updated according to

$$n_i(t+1) = n_i(t) + \left(\sum_j J_{ij} n_j(t) - \sum_{ij} J_{ij} n_i(t) n_j(t)\right) \cdot n_i(t) \tag{8}$$

The stochastic part of the dynamics is implemented in the following way. With probability p^{mut} a strategy number i mutates into another one. In this process a fraction α^{mut} of the frequency of that strategy is moved to a strategy with a label chosen from the normal distribution $N(n_i(t), \sigma)$ centred on label i with variance σ. The closer the labels of two strategies the more likely it is for one to mutate into the other.

Should a frequency become smaller than an extinction threshold $n_i(t) < n^{\text{ext}}$ it is remove $n_i(t+1) = 0$. After an extinction event the system is renormalized in order to respect the condition $\sum_i n_i(t) = 1$.

We show in Fig.(2) and example of the complex macroscopic dynamics generated by these dynamical rules.

3.2.1. Mean Field Description The mean field evolution equation obtained by averaging over the stochastic mutations and keeping in mind that strategies may go extinct takes the form

$$n_i(t+1) \simeq n_i(t) + \left(\sum_j J_{ij} n_j(t) - \sum_{ij} J_{ij} n_i(t) n_j(t)\right) \cdot n_i(t)$$

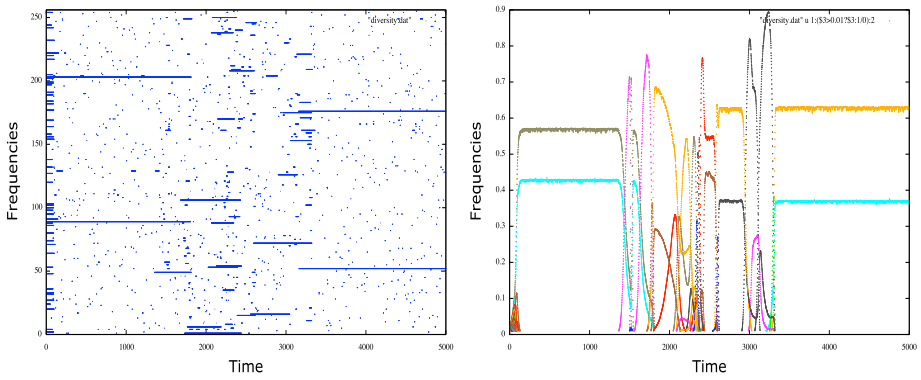


Figure 2. Left panel: We present once again the occupation plot. This is done exactly like in the Tangled Nature's case. The blue lines represent strategies, typically very fit, that are consistently in the system. The blue dots are strategies populated by mutations, that quickly go extinct. Populating a new fit strategy usually indicates the arrival of a transition (label 50 around $t = 1500$) Right panel: Each colour is associated with a different strategy, and indicates the frequency of players that are using it. We can clearly see how the system jumps from one configuration to the other.

$$-\alpha^{\rm mut}p^{\rm mut}\cdot n_i(t)+\frac{p^{\rm mut}}{N_a(t)}\sum_j \alpha^{\rm mut}n_j(t)p_{j\to i} \tag{9}$$

where $p_{j\to i}=\frac{\exp(-\frac{\|i-j\|}{2\sigma^2})}{\sqrt{2\pi\sigma^2}}$ follows the normal distribution. We can express Eq.(9), in compact form as

$$n_i(t+1)-n_i(t)\simeq\sum_j \mathbb{T}_{\rm ij}n_j(t) \tag{10}$$

where

$$\mathbb{T}_{ij}=\left(\sum_j J_{ij}n_j(t)-\sum_{ij}J_{ij}n_i(t)n_j(t)-\alpha_{\rm mut}\right)\cdot\delta_{ij} \tag{11}$$
$$-\frac{\alpha_{\rm mut}}{2^L-1}\cdot(1-\delta_{ij})\,.$$

The corresponding linear stability matrix takes the form

$$\mathbb{M}_{ij}=\mathbb{T}_{ij}(\vec{n})+\left[J_{ij}-\sum_k(J_{ik}+J_{ki})\vec{n}^*_k\right]\vec{n}^*_i \tag{12}$$

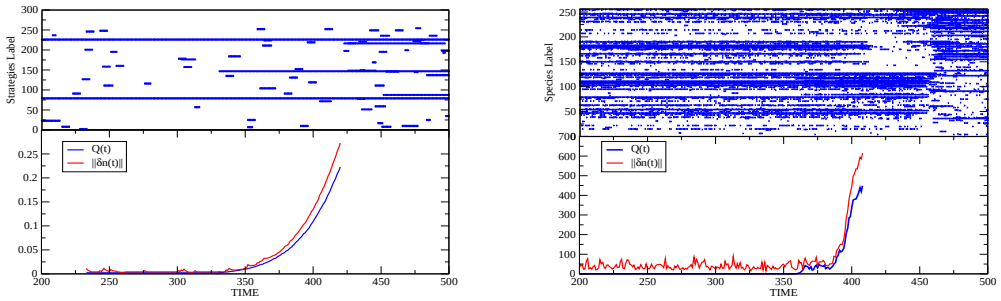


Figure 3. We show the behaviour of our stability indicator $Q(t)$, in both the Replicator Model (Left Panel) and the Tangled Nature (Right Panel). We compare Q's behaviour (blue curve) both to the displacement's $\delta n(t)$ (red curve) behaviour and the occupation plot. We can see how in both models the $Q(t)$ peaks only before the transition, while it doesn't feel the previous fluctuations.

4. Results

We now present how the indicator defined in Eq. (4) behaves as function of time for the two models. A comparison between the temporal evolution of $Q(t)$ and the deviation between the meta-stable configuration and the instantaneous state

$$\|\delta\vec{n}(t)\| = \|\vec{n}(t) - \bar{\vec{n}}_{\rm stoc}\| \tag{13}$$

is given in Fig.(3). We see that $Q(t)$ is not sensible to the stochastic fluctuations of that happen during the qESS, but peaks only just before the transition actually occurs. This suggests that the $Q(t)$ correctly distinguishes between harmless fluctuations, that happen in the stable subspace, and the dangerous ones that are capable to push the system out of its configuration. As one can see from Fig.(3) the $Q(t)$ starts peaking while the $\delta n(t)$ is stil in its characteristic values. This happens several steps before the transition occurs, proving the forecasting power of the variable.

A more systematic analysis is shown in Fig.(4) where we investigate the behaviour of the the indicator $Q(t)$ and $\delta n(t)$ indicator just before the transition, on 10000 transitions for each model. As we can see far from the transition (t=20) most of the Q(t) are close to zero, but approaching the transition (t = 0) the average values of $Q(t)$ start increasing while the $\delta n(t)$ stays more or less constant.

As one can clearly see in Fig.(4) some of the transitions are not predicted by our method. This means, in these *missed transitions*, that the system is driven out of its configuration by a perturbation parallel to what we consider the the stable subspace; for details see [3].

5. Conclusion

We have presented a procedure for forecasting the approach of systemic transitions in which stochastic dynamics is combined with mean field stability analysis about the "observed" configuration at a given time. The crux of the method is to include in the analysis new mutants or new strategies which originates from the existing agents but are distinct from them. We believe that our method can be applied to the dynamics of many real complex systems such as ecosystems and the financial market. The crucial point will be to carefully monitoring what happens at the perimeter of the extant agents.

6. Acknowledgments

This work is supported by the European project CONGAS (Grant FP7-ICT-2011-8-317672)

7. References

[1] P. Sibani and H.J. Jensen. *Stochastic Dynamics of Complex Systems*. Imperial College Press, 2013.
[2] Marten Scheffer, Jordi Bascompte, William A Brock, Victor Brovkin, Stephen R Carpenter, Vasilis Dakos, Hermann Held, Egbert H Van Nes, Max Rietkerk, and George Sugihara. Early-warning signals for critical transitions. *Nature*, 461(7260):53–59, 2009.
[3] A. Caroli, D. Piovani, and H.J. Jensen. Forecasting transitions in systems with high dimensional stochastic complex dynamics: A linear stability analysis of the tangled nature model. *Phys. Rev. Lett.*, 113:264102, 2014.
[4] K. Christensen, S. A. di Collobiano, M. Hall, and H. J. Jensen. Tangled nature model: A model of evolutionary ecology. *J. Theor. Biol.*, 216:73–84, 2002.
[5] M. Hall, K. Christensen, S. A. di Collobiano, and H. J. Jensen. Time-dependent extinction rate and species abundance in a tangled-nature model of biological evolution. *Phys. Rev. E*, 66, 2002.
[6] E. P. Anderson and H. J. Jensen. Network properties, species abundance and evolution in a model of evolutionary ecology. *J. Theor. Biol.*, 232:551–558, 2005.
[7] D. Lawson and H. J. Jensen. The species-area relationship and evolution. *J. Theor. Biol.*, 241:590–600, 2006.
[8] Diederich Sigurd and Manfred Opper. Replicators with random interactions: A solvable model. *Physical Review A*, 39:4333, 1989.
[9] K.and Yasutomi A. Tokita. Emergence of a complex and stable network in a model ecosystem with extinction and mutation. *Theoretical population biology*, 63:131–146, 2003.
[10] G. A. Pavliotis. *Stochastic Processes and Applications Diffusion Processes, the Fokker-Planck and Langevin Equations*. Springer, 2014.
[11] R. Veltri A. Kianercy and K.J. Pienta. Critical transitions in a game theoretic model of tumour metabolism. *Interface focus*, 4(4):20140014, 2014.
[12] Josef Hofbauer and Karl Sigmund. *Evolutionary Games and Population Dynamics*. Cambridge University Press, 1998.
[13] P.G. Higgs and B. Derrida. Genetic distance and species formation in evolving populations. *J. Mol. Evolution*, 35(5):454–465, 1992.
[14] Simone Avogadro di Collobiano, Kim Christensen, and Henrik Jeldtoft Jensen. The tangled nature model as an evolving quasi-species model. *J. Phys A*, 36:883–891, 2003.

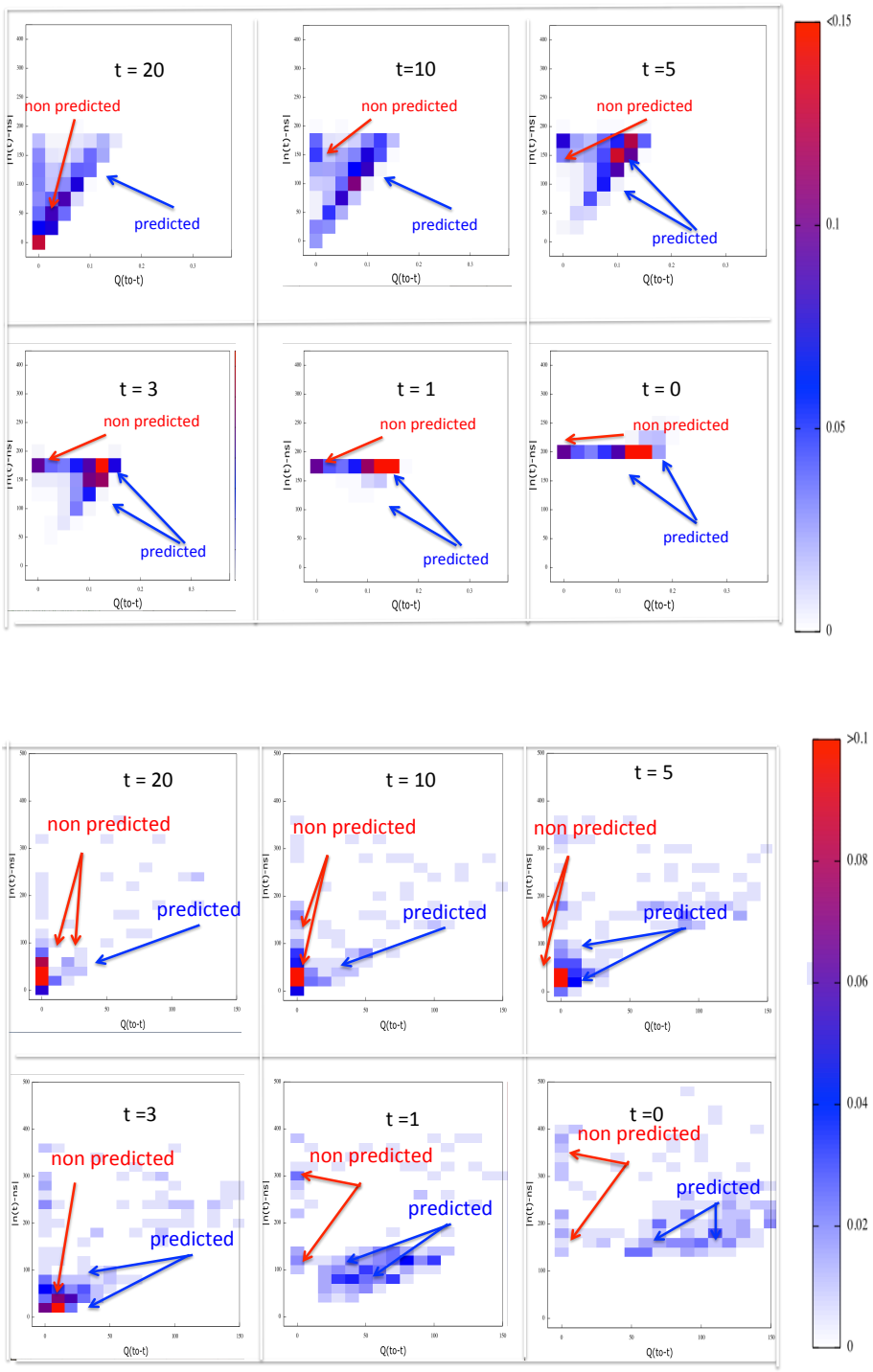


Figure 4. Top panel Replicator Model and bottom panel Tangled Nature model,. The figures compare the indicator $Q(t^*-t)$ (x-axis) and $\delta n(t^*-t)$ (y-axis) where t^* is the time the transitions begins. As we can see for $t = 20$, meaning 20 steps before the transitions most of the transitions are not predicted ($Q(t^* - t) \simeq 0$), but as the system approaches the transition the vast majority of transitions (80/90%) are predicted by an increase of $Q(t)$.