An information theory perspective on tipping points in dynamical networks

Casper van Elteren, Rick Quax, Peter Sloot

Abrupt, system-wide transitions can be endogenously generated by seemingly stable networks of interacting dynamical units, such as mode switching in neuronal networks or public opinion changes in social systems. However, it remains poorly understood how such 'noise-induced transitions' emerge from the interplay of network structure and dynamics on the network. We identify two key roles that nodes can play in the progression towards a tipping point can emerge and illustrate it in dynamical networks governed by the Boltzmann-Gibbs distribution. In the initial phase, initiator nodes absorb and transmit short-lived fluctuations to neighboring nodes, causing a domino-effect by making neighboring nodes more dynamic. Conversely, towards the tipping point we identify stabilizer nodes whose state information becomes part of the long-term memory of the system. We validate these roles by targeted interventions that make tipping points more (less) likely to begin or lead to systemic change. This opens up possibilities for understanding and controlling endogenously generated metastable behavior.

1 Introduction

Multistability is an important characteristic in many real-world complex systems [1, 2]. It entails the phenomenon whereby a system under the influence of noise explores its state space on different spatio-temporal scales. For example, the brain is able to operate in one cognitive mode for an extended time before rapidly switching to a different mode [3]. Similarly, the life cycle of a cell is tightly regulated between two bistable states of mitosis and interphase [4]. Other examples exist on larger scales such as the emergence of multistability in opinion dynamics [5], or the multistability of ecosystems or climate systems [6, 7].

There is increasing evidence that for some complex systems noise plays a fundamental role in the transition between the metastable states [8–11]. Noise is traditionally viewed as an unwanted signal that reduces the effectiveness of a system or reduces the quality of a measured signal. Indeed it is well known that for linear systems, the signal to noise ratio is maximized in the absence of noise. For non-linear systems, however, noise may allow the system to explore larger parts of its state space by allowing it to escape local minima [12, 13]. In networked systems, this noise is embodied as state fluctuations of individual nodes. These fluctuations can be amplified which allows the system to transition along different paths in the network of interactions resulting in metastable transitions (fig. 1).

The mechanisms underlying such transitions are often not well understood. It is of vital importance to understand the patterns of "fluctuations enhancing fluctuations" that can cause metastable behavior.

Here, we consider dynamical systems consisting of a static network where the states of the nodes are governed by a Boltzmann-Gibbs distribution (fig. 1). This type of model has been used to describe a wide range of behaviors such as neural dynamics [14], opinion dynamics, ferromagnetic spins [15], and organized criminal gang interactions [16].

In this class of systems, each node chooses its state in local equilibrium according to the potential induced by its neighbor states. In physical applications this potential is the classic energy potential, but in other applications it can be interpreted, for instance, as frustration level, assortativity(homophily), or more broadly speaking, a fitness score of the state of the node given its neighbors. The second ingredient in this model is a global 'temperature' which is essentially a noise level; at zero noise a node always picks the absolute minimum energy state, whereas the higher this noise level, the more likely it is that high energy states are chosen. Without loss of generality, this study considers no external field forces for simplicity.

For low noise levels, it is common for systems governed by the Boltzmann-Gibbs distribution to exhibit metastable behavior because of the existence of multiple (local) minima in the system's potential (fig. 1a-c). In finite systems and nonzero temperature, there is a finite probability that the system moves (eventually) from one local minimum to another. In this paper we present a generic method to identify how these metastable transitions are generated by paths of reinforcing fluctuations, and demonstrate it on the well-known kinetic Ising spin model without external forces. Here, nodes have only two possible states: 0 and $+1$. At system level, there are two global minima: all nodes in state 0 or all nodes in state $+1$ (fig. 1c). Between these two system states lies a 'potential barrier' (fig. 1b): many possible paths of system states connect the two systemic minima, all of which having a growing potential, making these paths less likely than paths of similar length that remain close to one of the minima. The peak of this potential along each crossing path lies, informally speaking, at a 'checkerboard pattern': each node being maximally different from (the majority of) its neighbors. We refer to this peak as the 'tipping point'.

The crucial point here is that the network structure by itself can make systemic transitions much more likely [7, 17–19]. Without network structure, each node has an independent (low) probability of choosing a state closer to the tipping point (higher potential energy). The probability that all nodes in the system happen to do this simultaneously, (thereby transitioning the system state to another potential minimum) decreases to zero rapidly (as $\mathcal{O}(e^{-N^2})$ for dense networks with N representing the number of nodes). This means that transitions become unlikely for all but the smallest systems. When adding a network structure, however, transitions can potentially occur along a path of nodes that form a "domino effect": a first node choosing may sometimes choose a higher energy state under the influence of noise and in doing so making that same transition more likely for all its neighbors. For some of these neighbors this new situation may suffice to make their own transition t o a higher energy state with (almost) equal likelihood as the first node, and so on, until the tipping point has been reached. The likelihood of such a transition is much higher than without network effects (which is up to $\mathcal{O}(e^{-N})$). This is still an exponentially decaying function of system size, highlighting the fact that such noise-induced transitions are expected only to occur for finite-sized systems, but exceedingly more likely with network structure than without it.

Here, we present a method to uncover the network percolation process that facilitates endogenously generated, noise-induced transitions. The approach differs differs from the traditional approaches that focus on how the system as a whole approaches a tipping point. The computational method is broadly applicable. It makes no assumptions about the network structure nor the type of dynamics governing the nodes. In principle it requires access only to the system state probabilities over time. These could be obtained, for instance, through cross-sections of time-series.

2 Results

Fluctuations and their correlations at time τ are captured using Shannon's mutual information [20] shared between a node and the entire future system state $I(s_i^\tau : S^{\tau+t})$. The time lag t is used to analyze two key features of information flows of a system: the area under the curve (AUC) of short-term information, and sustained level of long term information.

Depending on the network connectivity of a node (fig. A.9), its contribution to the dynamics of the system will differ [21, 22]. The total amount of fluctuations shared between the node's current state and the system's short-term future trajectory is computed as the integrated mutual information

$$\mu(s_i) = \sum_{t=0}^{\infty} (I(s_i^\tau : S^{\tau+t}) - \omega_{s_i})\Delta t. \qquad (1)$$

Intuitively, $\mu(s_i)$ represents a combination of the intensity and duration of the short-term fluc-

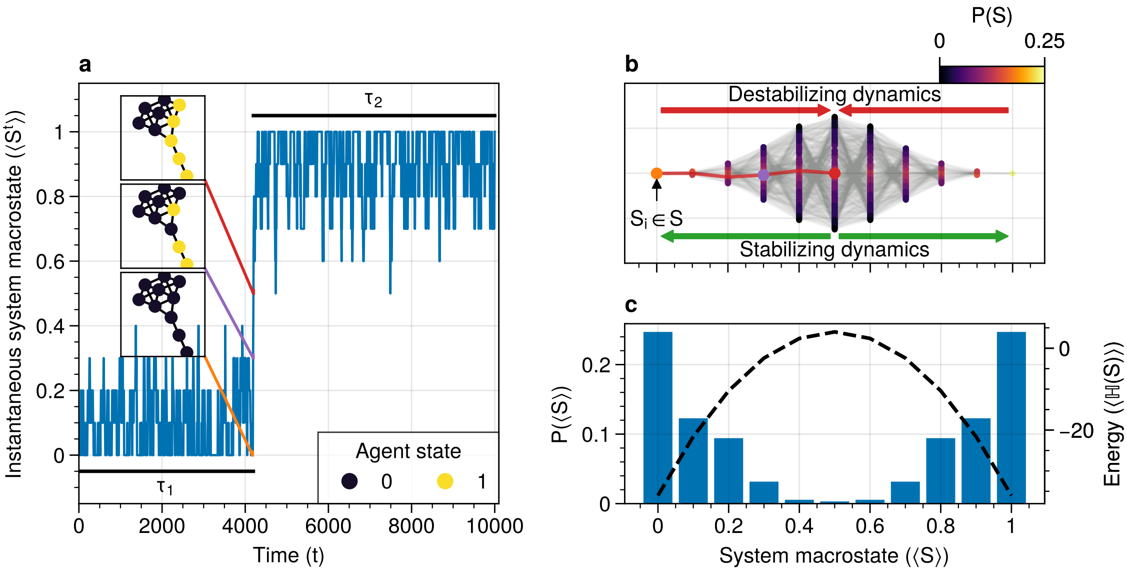

Figure 1: *A dynamical network governed by kinetic Ising dynamics produces multistable behavior. (a) A typical trajectory is shown for a kite network for which each node is governed by the Ising dynamics with $\beta \approx 0.534$. The panels show system configurations $S_i \in S$ as the system approaches the tipping point (orange to purple to red). For the system to transition between metastable points, it has to cross an energy barrier (c). (b) The dynamics of the system can be represented as a graph. Each node represents a system configuration $S_i \in S$ such as depicted in (a). The probability for a particular system configuration $p(S)$ is indicated with a color; some states are more likely than others. The trajectory from (a) is visualized. Dynamics that move towards the tipping point (midline) destabilize the system, whereas moving away from the tipping point are stabilizing dynamics. (c) The stationary distribution of the system is bistable. Transitions between the metastable states are infrequent and rare. For more information on the numerical simulations see A.2.*

tuations on the (transient) system dynamics [21]. It reflects how much of the node state is in the "working memory" of the system.

The term $\omega(s_i)$ quantifies the long term memory of the system. As the dynamics of the system evolve, the short-lived fluctuations will cancel and the long term behavior of the system will dominate more and more. These slow fluctuations are correlated with the metastable state the system is in. Around a low energy state, the system produces short-lived fluctuations. However, as the system approaches a tipping point a new low energy state will be chosen. Correspondingly, the correlations of a node with this new future system state will produce long(er) time scale correlations. The next tipping point will be reached on a much longer timescale. Consequently, ω quantifies the system returning to a stable system regime. For nodes with fast dynamics, $\mu(s_i)$ is generally high and ω_{s_i} would be generally low.

In fig. 2a-e the information flows are shown at different stages in the metastable transition. The results are shown for the kite graph to illustrate the information features and to enhance clarity of the meaning of the information flows. Results for synthetic networks are shown in fig. 6. The metastable transition was decomposed by considering the local information flows from a given system partition $S_\gamma = \{S' \subseteq S | \langle S' \rangle = \gamma\}$ where $\gamma \in [0, 1]$ is the fraction of nodes having state +1. This yields the conditional integrated mutual information as

$$\mu(s_i|\langle S \rangle) = \sum_{t=0}^{\infty} (I(s_i^\tau : S^{\tau+t}|\langle S^\tau \rangle) - \omega_{s_i})\Delta t. \quad (2)$$

Details about the estimation procedure can be found in appendix: A.5.

Two things are observed from fig. 2. First, the tipping point is reached by a domino effect where low degree nodes play an "initiator" role early in the process. In this model, low degree nodes are most susceptible to noise fig. A.9 and therefore are more likely to pass on fluctuations to neighbors. Far away from the tipping point (fig. 2a), nodes with lower degree have higher shared information (higher $\mu(s_i|\langle S \rangle)$) than higher degree nodes. Lower degree nodes can initiate a metastable tran-

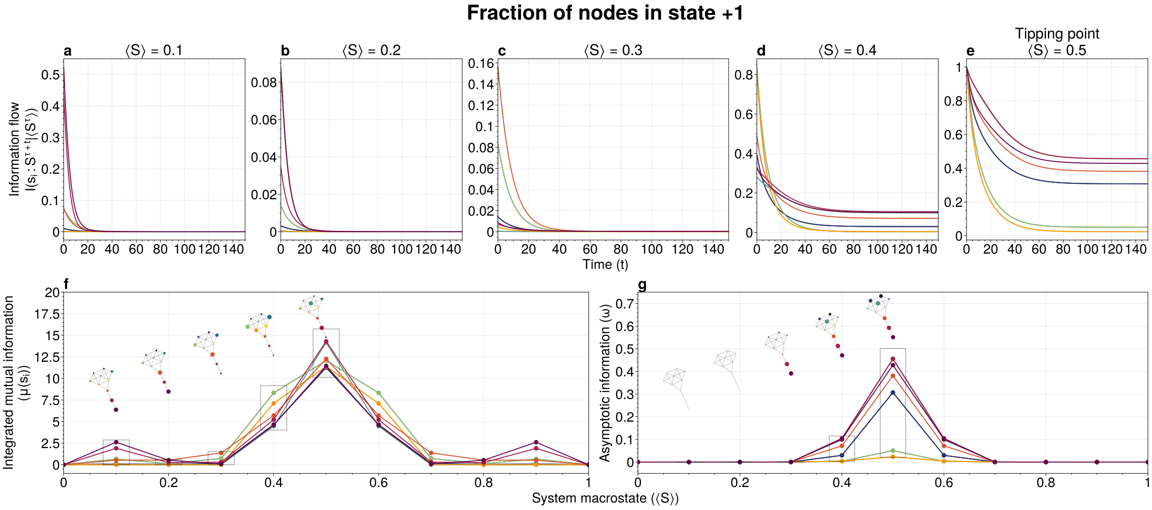

Figure 2: *(a-e) Information flows as distance to tipping point. Far away from the tipping point most information processing occurs in low degree nodes (f,g). As the system moves towards the tipping point, the information flows increase and the information flows move towards higher degrees. (f) Integrated mutual information as function of distance to tipping point. The graphical inset plots show how noise in introduced far away from the tipping point in the tail of the kite graph. As the system approaches the tipping point, the local information dynamics move from the tail to the core of the kite. (g) A rise in asymptotic information indicates the system is close to a tipping point. At the tipping point, the decay maximizes as trajectories stabilize into one of the two metastable states.*

sition by injecting noise into the system. Without this injected noise, it would be less likely for a metastable transition to occur. In other words, in a system that is slightly destabilized by low degree nodes with high energy (fluctuating states), the transition towards a tipping point becomes more likely as neighboring higher degree nodes are more likely to become "initiator" nodes. This cascade progresses whereby new initiator nodes are formed through local fluctuations.

Second, an increase in asymptotic behavior correlates with the system transitioning from one attractor state to another. The asymptotic information remains low far away from the tipping point, and monotonically increases as the system approaches the tipping point fig. 2{b, c}). A node's asymptotic information encodes how much predictive information a node has about the future system state. After a tipping point, the system either relaxes to the closest attractor state or transitions across the tipping point into the next attractor state. After such a transition, the dynamics of the nodes slow down. That is, all but the nodes with the lowest degrees are locally frozen as the system dynamics restabilizes after a noise-induced perturbation. A node with high asymptotic information will have more information regarding which side of the tipping point the system ends up being.

To illustrate what is encoded in the information flows trajectories were computed from the attractor state $S = \{0, \dots, 0\}$ and simulated for $t = 5$ steps. In fig. 3 a trajectory is shown that maximizes

$$\log p(S^{t+1}|S^t, S^0 = \{0, \dots, 0\}, \langle S^5 \rangle = 0.5).$$

These trajectories reveal how the information flows measured in fig. 2c are caused by the sequence of flips generated from the "tail" in the kite graph. These tail nodes are uniquely positioned due to their higher potential to pass on fluctuations to their neighbors eventually causing a cascade of flips that reach the tipping point.

The domino effect is not completely correlated with degree. As the system approaches the tipping point, destabilizing fluctuations tend to be caused by lower degree nodes, but as the system approaches the tipping point network effects play a profound role. For example, consider node 8 and node 3. Node 8 has degree 2 and has the highest integrated mutual information when 2 bits are flipped in the system (fig. 2b). The dynamics for node 8 for all states where $\langle S \rangle = 0.2$ (or 0.8 by symmetry) indicate that 8 is essential in propagating the fluctuations generated by 9. At the tipping point, node 8 shares the highest information with the system. In contrast, node 3 which has degree 6 has low shared information prior to the tipping, indicating that 3 is less involved with initializing the tipping point. At the tipping point, however, node 3 has high amounts of shared information with the future system states, similar to that of node 8. This makes it hard to generate a strict rule based on network connectivity alone what role a node has in the tipping behavior. Both the network structure and the dynamics fundamentally interact in generating the tipping points. Furthermore, the role of a node may change as the system approaches a tipping point.

To further illustrate the intricacies encoded in the information flows, the most likely trajectory to and from the tipping point were computed. The path analysis reveal that at the tipping point the system is most likely to either (a) move from one attractor state to another, or (b) relax back to the attractor state it evolved from (fig. 3). The most likely paths reaching the tipping point from one of the ground state results in a configuration in which a high degree cluster set of nodes has to flip (e.g. 1,0,3,4,6 in fig. 3 at $\langle S \rangle = 0.5$). This trajectory is less likely than essentially reversing the path shown in fig. 3. Hence, most of the tipping points "fail" and relax back to the attractor state from which it evolved (fig. 4b). If, however, it does make the metastable transition to the other side, the "tail" in the graph remains stable for these transitions, yielding relative high correlation for node 8, 9. The information flows reflect how certain a given node is about the future system state, e.g. $H(S^{t+\tau}|s_i^t)$, revealing how much uncertainty it has on how quickly $p(S^{t+\tau})$ converges to some stable trajectory around a future attractor state.

The increased information of node 8 around the tipping point can now be understood by considering the what kind of information 8 has about the future of the system. The path analysis revealed that the network structure plays a fundamental role whereby a domino effect from the "bottom-up" is the most likely path to and from a tipping point. This implies that the information that node 8 and node 3 store about the future of the system differs but ends up providing the same amount of shared information. In fig. 4a the conditional probabilities are shown of each node relative to the tipping point. Both node 3 and node 8 have the lowest uncertainty about the future system state. However, the nature of this uncertainty differs. Relative to the tipping point, the node 3 has more certainty that the average of the system state will be equal to its state at the tipping point. This reflects the node's ability to "choose" the next stable point. This is most likely caused for the kite graph by a failure of the system to transition between attractor states (fig. 4b): most transitions are more likely to transition back to the metastable state it transition from towards the tipping point. Node 8, however, shares different information about the future system state. Figure 4 shows that node 8 has higher certainty that the future system state will most likely have opposite sign to its state at the tipping point. As most tipping points fail to transition between metastable points, node 8 will have the opposite state to what it was at the tipping point. This gives node 8 a non-intuitive high predictive power of the system's future.

The information flows reflect the most probable trajectories around the partition $\langle S \rangle = c$ and give unique insights into the mechanism driving the tipping behavior. Over time, local clusters will stabilize. Some nodes will experience more "frustration" than others. In other words, the node will tend to change state more as the effect of a node flip percolates through the system. For example, nodes 5 (yellow) and 6 (orange) have the lowest asymptotic information while still having a relatively high degree. These nodes experience more frustration as they attempt to reconcile with the states of the nearest neighbors.

2.1 Simulated interventions

The cascade of flips is further studied using simulated interventions (fig. 5, fig. 6). By pinning each node state to 0 in separate simulations, the effect on the occurrence of tipping points is studied. The interventions highlight two distinct roles for the metastable transitions. Intervention on low degree nodes remove fluctuations in the system macrostate 0 but increases the fluctuations when the system reaches the macrostate 1. The effect is most prominent for node 9 which has degree 1 (fig. 5b); interventions on node 9 yields the lowest time spent in the 0 metastable state (fig. 5a), and the highest time spent in the 1 macrostate relative to interventions on other nodes(fig. 5b). Notable, the number of tipping transitions is the least affected by lower degree nodes. In contrast, high degree nodes seem to be essential for the tipping behavior to endure; lower degree nodes are necessary to destabilize the system, but the higher degree nodes have to flip in order for the new metastable state to endure. This can be seen by

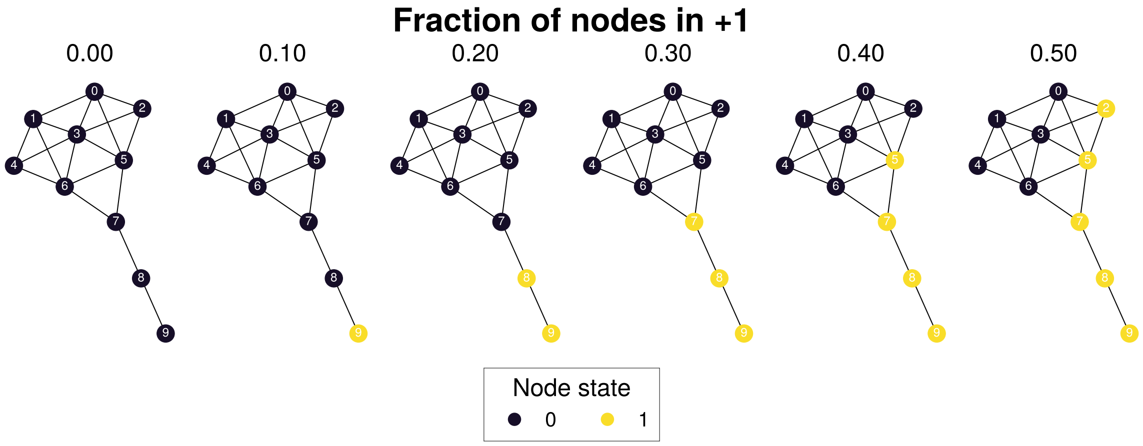

Figure 3: *The tipping point is initiated from the bottom up. Each node is colored according to state 0 (black) and state 1 (yellow) Shown is a trajectory towards the the tipping point that maximizes* $\sum_{t=1}^{5} \log p(S^{t+1}|S^t, S^0 = \{0\}, \langle S^5 \rangle) = 0.5)$*. As the system approaches the tipping point, low degree nodes flip first, and recruit "higher" degree nodes to further destabilize the system and push it towards a tipping point. In total 30240 trajectories that reach the tipping point in 5 steps, and there are 10 trajectories that have the same maximized values as the trajectory shown in this figure.*

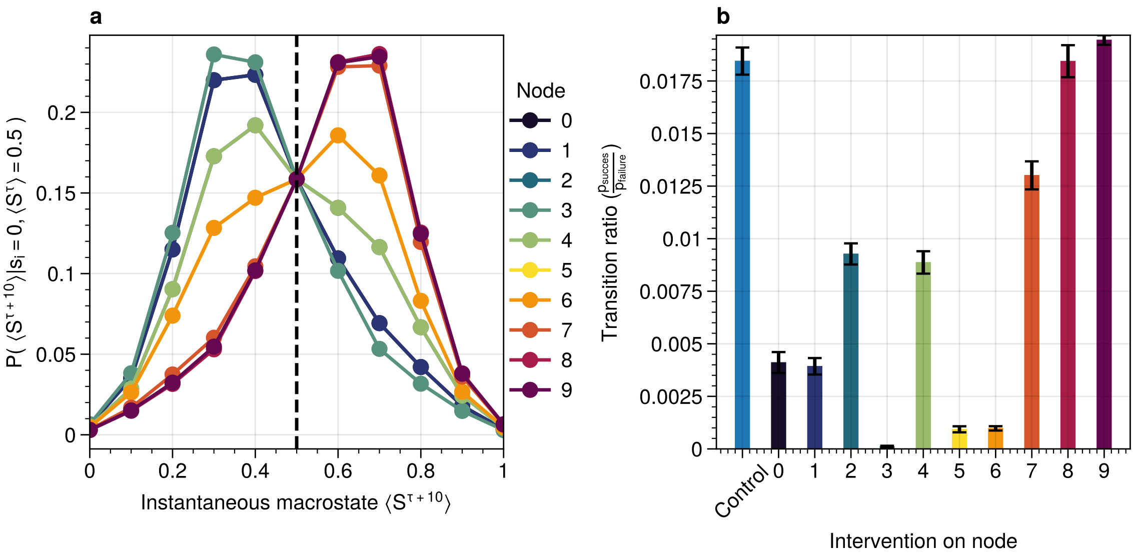

Figure 4: *(a) Shown are the conditional probability at time* $t = 10$ *relative to the tipping point. The shared information between the hub node 3 and the tail node 8 is shared is similar but importantly caused through different sources. The hub (node 3) has high certainty on that the system macrostate will be the same sign as its state. In contrast, node 8 has high certainty that the system macrostate will be opposite to its state at the tipping point. This is caused by the interaction between the network structure and the system dynamics whereby the most likely trajectories to the tipping point from the stable regime is mediated by the noise-induced dynamics from the tail to the core in the kite graph (see main text).(b) Successful metastable transitions are affected by network structure. Successful metastable transitions are those for which the sign of the macrostate is not the same prior and after the tipping point, e.g. the system going from the 0 macrostate side to the +1 macrostate side or vice versa. Shown here are the number of successful metastable transitions for fig. 5 under control and pinning interventions on the nodes in the kite graph.*

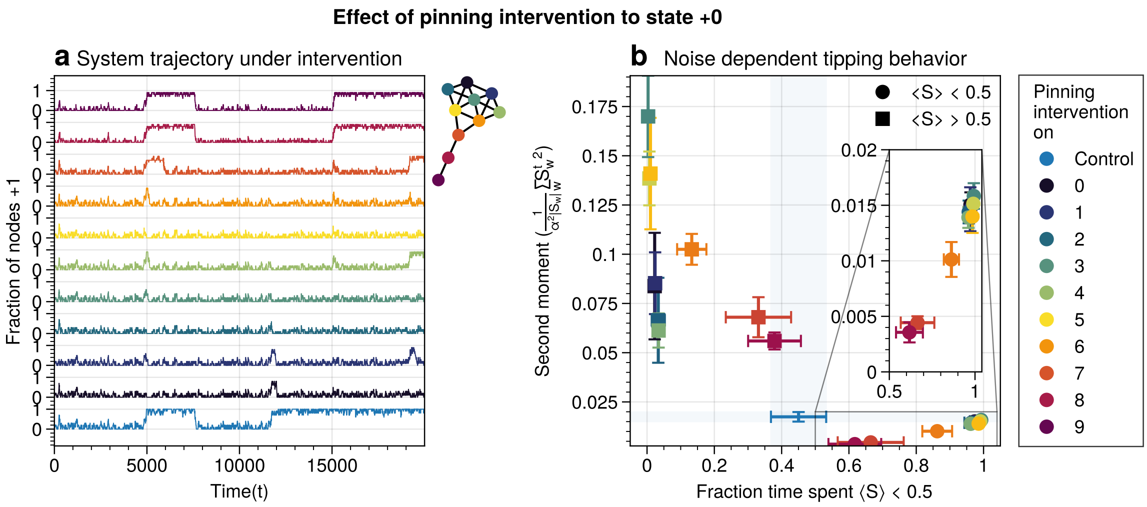

Figure 5: *For a system to cross a tipping point two different types of nodes are identified. High degree nodes are essential for system to move from one metastable point to another. Low degree nodes are essential to propagate noise into the system. In (a) typical system trajectories are shown under pinning intervention on a node. Each color indicates a targeted intervention on the colors matching in (a). (b) The effect of intervention has a different effect depending on which node is targeted; Targeting a high degree node to the 0 state (e.g. node 3) prevents the system into tipping the opposite side of the pinning effect. In contrast, targeting a low degree node (e.g. 9) the system is still able to explore the full state space. Intermediate connected nodes (e.g. node 7, 8) removed merely nudges the system macrostate to one side, and increases the probability to remain in the 0 macrostate. In (b) ± 2 standard error of the mean are shown.*

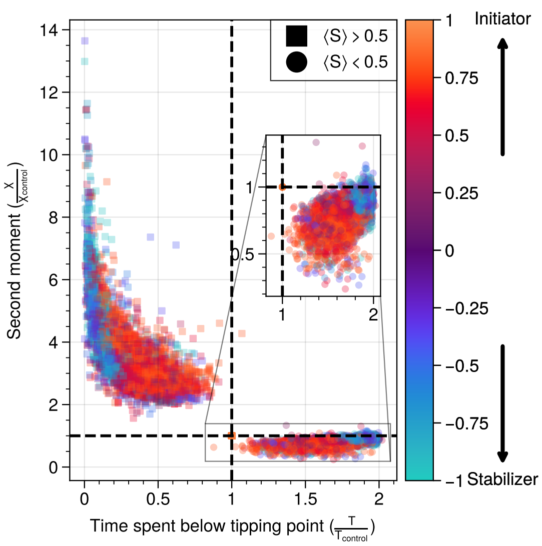

Figure 6: *The effect of pinning intervention per node on 0 state in Erdos-Reniy graphs(N = 100 graphs with 10 nodes each and p = 0.2, 6 different seeds). Shown are the second moment (noise) and time spent below the tipping relative to the control per network. Pinning intervention on initiator nodes increases the occurrence of tipping points. In contrast, interventions on stabilizers prevents tipping points and increases noise above the tipping point. These results extend fig. 5, for more details on the role approximation please see appendix: 2.2.*

the time spent in the 1 macrostate: interventions on a hub node has increased white noise compared to control conditions in the 0 macrostate (fig. 5a). This indicates that noise is propagated and nodes are flipped towards the tipping point, but are less likely to cross the tipping point. This is further strengthened by the reduced time spent in the 1 macrostate as a function of degree fig. 5b.

2.2 Role division in random networks

The contribution of each node to the tipping point be decomposed as a continuum between two roles. Initiator node are essential to kickstart the domino effect. These nodes are influential early on to inject and transfer fluctuations to neighboring nodes. At the tipping point, stabilizing nodes contribute to the system returning to a stable attractor state. A role for node i can be approximated by the difference between integrated mutual information and asymptotic information

$$r_i = \max_{\langle S\rangle} \mu^Z(s_i|\langle S\rangle) - \max_{\langle S\rangle} \omega^Z(s_i) \in [-1, 1], \quad (3)$$

with

$$\mu^Z(s_i|\langle S\rangle) = \frac{\mu(s_i|\langle S\rangle)}{\max_j \mu(s_j)}$$
$$\omega^Z(s_i|\langle S\rangle) = \frac{\omega(s_i|\langle S\rangle)}{\max_j \omega(s_j)}. \tag{4}$$

For role values close to 1, the node is classified as a (pure) initiator. These nodes have high integrated mutual information indicating high predictive information about short-lived system trajectories. However, these nodes lack long-term predictive information about future system states. Conversely, a node classified as -1, has is a pure stabilizer. Roles having a value $r_i \sim 0$ are more difficult to interpret as the zero value could be caused by an equally large integrated mutual information and asymptotic information or a generally lacking high score in both.

In fig. 6 roles were computed for Erdos-Renyi networks under simulated interventions (N = 10, all graphs are non-isomorphic, see appendix: A.10). The interventions performed were similar as in fig. 5. System fluctuations are quantified using the second moment and normalized per system (see appendix: A.10). Three observations can be made.

First, intervention on initiator nodes increases tipping behavior of the system. As the role r_i approaches 1, the noise fluctuations and tipping behavior approaches control (dashed lines). As the role approaches -1, however, the tipping behavior decreases compared to the control condition.

Second, intervention on initiator reduces the noise fluctuations below the tipping point. This is similar reduced discussed in fig. 5: higher frequency fluctuations are more likely removed when intervening on initiators.

Lastly, system fluctuations are higher when interventions are performed on stabilizers. Fluctuations are lower when interventions are performed on roles $r_i \to 1$. As r_i decreases, the fluctuations in the macro states increases both below as above the tipping point. The effect is stronger above tipping point as the intervention is signed to the 0 state, which generates a drive towards system states $\langle S\rangle < 0.5$.

3 Discussion

Understanding how metastable transitions occur may help in understanding how, for example, a pandemic occurs, or a system undergoes critical failure. In this paper, dynamical networks governed by the Boltzmann-Gibbs distribution were used to study how endogenously generated metastable transitions occur. The external noise parameter (temperature) was fixed such that the statistical complexity of the system behavior was maximized (see appendix A.2).

The results show that in the network two distinct node types could be identified: *initiator* and *stabilizer* nodes. Initiator nodes are essential early in the metastable transition. Due to their high degree of freedom, these nodes are more effected by external noise. They are instigators and inject noise into the system, destabilizing more stable nodes. In contrast, stabilizer nodes, have high degree of freedom and require more energy to change state. These nodes are essential for the metastable behavior as they stabilize the system macrostate. During the metastable transition a domino sequence of node state changes are propagated in an ordered sequence towards the tipping point.

This domino effect was revealed through two information features unvealing an *information cascade* underpinning the trajectories towards the tipping point.

Integrated mutual information captured how short-lived correlations are passed on from the initator nodes. In the stable regime (close to the ground state) low degree nodes drive the system dynamics. Low degree nodes destabilize the system, pushing the system closer to the tipping point. In most cases, the initiator nodes will fail in propagating the noise to their neighbors. On rare occasions, however, the cascade is propagated progressively from low degree, to higher and higher degree. A similar domino mechanism was recently found in climate science [7, 19]. Wunderling and colleagues provided a simplified model of the climate system, analyzing how various components contribute to the stability of the climate. They found that interactions generally stabilize the system dynamics. If, however, a metastable transitions was initialized, noise was propagated through a similar mechanism as found here. That is, an "initializer" node propagated noise through the system which created a domino effect that percolated through the system. The results from this study mirrors these conclusions and provides a model-free language to express these domino effects.

An increase in asymptotic information forms an indicator of how close the system is to a tipping point. Close to the ground state, the asymptotic information is low, reflecting how transient noise perturbations are not amplified and the system macrostate relaxes back to the ground state. As the system approaches the tipping point, the asymptotic information increases. As the distance to the ground state increases, the system is more likely to transition between metastable states. After the transition, there remains a longer term correlation. Asymptotic information reflects the long(er) timescale dynamics of the system. This “rest” information peaks at the tipping point, as the system chooses its next state.

The information viewpoint uniquely reveals a complex mechanism of interaction underlying the system macrostate. It allows for compressing the high dimension probability distribution in a away to understand what elements are fundamental for a tipping point ot be reached. It revealed how some nodes may have high predictive information, which is hard to infer from their interaction structure alone fig. 4. Integrated information and asymptotic information jointly readout the separation of fast-time scale dynamics that tend to stabilize noise-induced dynamics, and slow timescale dynamics indicating a metastable transition. Importantly, these measures can be directly computed on data.

4 Conclusions

Our information theoretic approach offers an alternative view to understand *how* metastable transitions are generated by dynamical networks. Two information features were introduced that decompose the metastable transition in sources of high information processing (integrated mutual information) and distance of the system to the tipping point (asymptotic information). A domino effect was revealed, whereby low degree nodes initiate the tipping point, making it more likely for higher degree nodes to tip. On the tipping point, long-term correlations stabilizes the system inside the new metastable state. Importantly, the information perspective allows for estimating integrated mutual information directly from data without knowing the mechanisms that drive the tipping behavior. The results highlight how short-lived correlations are essential to initiate the information cascade for crossing a tipping point.

5 Limitations

Integrated mutual information was computed based on exact information flows. This means that for binary systems it requires to compute a transfer matrix on the order of $2^{|S|} \times 2^{|S|}$. This reduced the present analysis to smaller graphs. It would be possible to use Monte-Carlo methods to estimate the information flows. However, $I(s_i^\tau : S^{\tau+t})$ remains expensive to compute. When using computational models, it requires to compute the conditional and marginal distributions which are on order $\mathbb{O}(2^{|S|})$ and $\mathbb{O}(2^{t|S|})$ respectively.

In addition, the decomposition of the metastable transition depends on the partition of the state space. Information flows are in essence statistical dependencies among random variables. Here, the effect of how the tipping point was reached was studied by partition the average system state in terms of number of bits flipped. This partitioning assumes that the majority of states prior to the tipping point are reached by having fraction $c \in [0, 1]$ bits flipped. The contribution of each system state over time, however, reflects a distribution of different states; reaching the tipping point from the ground state 0, can be done at $t-2$ prior to tipping by either remaining in 0.4 bits, or transitioning from 0.3 bits flipped to 0.4 and eventually to 0.5 in 2 time steps. The effect of these additional paths showed marginal effects on the integrated mutual information and asymptotic information.

Information flows conditioned on a partition is a form of conditional mutual information [23]. Prior results showed that conditional information produces synergy, i.e. information that is only present in the joint of all variables but cannot be found in any of the subset of each variable. Unfortunately, there is no generally agreed upon definition on how to measure synergy [24, 25] and different estimates exist that may over or underestimate the synergetic effects. By partitioning one can create synergy as for a given partition each spin has some additional information about the other spins. For example, by taking the states such that $\langle S \rangle = 0.1$, each spin “knows” that the average of the system equals 0.1. This creates shared information among the spins. Analyses were performed to estimate synergy using the redundancy estimation I_{min}[26]. Using this approach, no synergy was measured that affected the outcome of this study. However, it should be emphasized that synergetic effects may influence the causal interpretation of the approach presented here.

A general class of systems was studied governed by the Boltzmann-Gibbs distribution. For practical purposes the kinetic Ising model was only tested, but we speculate that the results should hold (in principle) for other systems dictated by the Boltzmann-Gibbs distribution. We leave the extension to other system Hamiltonians for future work.

6 Acknowledgments

CvE would like to thank Fiona Lippert, and Jair Lenssen for providing insights and feedback in various ideas present in this paper. This research is supported by grant Hyperion 2454972 of the Dutch National Police.

7 References

1. Ladyman, J., Lambert, J. & Wiesner, K. What Is a Complex System? *European Journal for Philosophy of Science* **3,** 33–67. ISSN: 1879-4912 (Jan. 19, 2013).
2. Van Nes, E. H. *et al.* What Do You Mean, 'Tipping Point'? *Trends in Ecology & Evolution* **31,** 902–904. ISSN: 01695347 (Dec. 2016).
3. Fries, P. Rhythms for Cognition: Communication through Coherence. *Neuron* **88,** 220–235. ISSN: 08966273 (Oct. 2015).
4. Kandel, E. R., Schwartz, J. H. & Jessell, T. M. *Principles of Neural Science* 4th ed. ISBN: 0-07-112000-9. `http://www.amazon.com / exec / obidos / redirect ? tag = citeulike07 - 20 % 5C & path = ASIN / 0071120009` (McGraw-Hill Medical, July 2000).
5. Galam, S. & Cheon, T. Tipping Points in Opinion Dynamics: A Universal Formula in Five Dimensions. *Frontiers in Physics* **8,** 566580. ISSN: 2296-424X. arXiv: `1901 . 09622 [cond-mat, physics:nlin, physics:physics]` (Nov. 10, 2020).
6. © 1986 Nature Publishing Group (1986).
7. Wunderling, N., Donges, J. F., Kurths, J. & Winkelmann, R. Interacting Tipping Elements Increase Risk of Climate Domino Effects under Global Warming. *Earth System Dynamics* **12,** 601–619. ISSN: 2190-4987 (June 3, 2021).
8. Beggs, J. M. & Timme, N. Being Critical of Criticality in the Brain. *Frontiers in Physiology* **3.** ISSN: 1664-042X (2012).
9. Mitchell, M., Hraber, P. & Crutchfield, J. P. *Revisiting the Edge of Chaos: Evolving Cellular Automata to Perform Computations* 1993. arXiv: `adap-org/9303003`. `http://arxiv.org/abs/adap-org/9303003`.
10. Mitchell, M., Crutchfield, J. P. & Hraber, P. Dynamics, Computation, and the "Edge of Chaos": A Re- Examination, 17.
11. Forgoston, E. & Moore, R. O. A Primer on Noise-Induced Transitions in Applied Dynamical Systems. *SIAM Review* **60,** 969–1009. ISSN: 0036-1445, 1095-7200 (Jan. 2018).
12. Czaplicka, A., Holyst, J. A. & Sloot, P. M. Noise Enhances Information Transfer in Hierarchical Networks. *Scientific Reports* **3.** ISSN: 20452322 (2013).
13. Nicolis, G. & Nicolis, C. Stochastic Resonance, Self-Organization and Information Dynamics in Multistable Systems. *Entropy* **18,** 172. ISSN: 1099-4300 (May 4, 2016).
14. Hopfield, J. J. Neural Networks and Physical Systems with Emergent Collective Computational Abilities. *Proceedings of the National Academy of Sciences of the United States of America* **79,** 2554–8. ISSN: 0027-8424. pmid: `6953413` (May 1982).
15. Glauber, R. J. Time-Dependent Statistics of the Ising Model. *Journal of Mathematical Physics* **4,** 294–307. ISSN: 00222488 (1963).
16. D'Orsogna, M. R. & Perc, M. Statistical Physics of Crime: A Review. *Physics of Life Reviews* **12,** 1–21. ISSN: 15710645 (Mar. 2015).
17. Harush, U. & Barzel, B. Dynamic Patterns of Information Flow in Complex Networks. *Nature Communications* **8,** 1–11. ISSN: 20411723 (2017).

18. Gao, J., Barzel, B. & Barabási, A.-L. Universal Resilience Patterns in Complex Networks. *Nature* **536,** 238–238. ISSN: 0028-0836. pmid: 11507039 (2016).

19. Wunderling, N. *et al.* How Motifs Condition Critical Thresholds for Tipping Cascades in Complex Networks: Linking Micro- to Macro-Scales. *Chaos: An Interdisciplinary Journal of Nonlinear Science* **30,** 043129. ISSN: 1054-1500, 1089-7682 (Apr. 2020).

20. Cover, T. M. & Thomas, J. A. *Elements of Information Theory* 1–748. ISBN: 978-0-471-24195-9. pmid: 20660925 (2005).

21. Van Elteren, C., Quax, R. & Sloot, P. Dynamic Importance of Network Nodes Is Poorly Predicted by Static Structural Features. *Physica A: Statistical Mechanics and its Applications,* 126889. ISSN: 03784371 (Jan. 2022).

22. Quax, R., Apolloni, A. & a Sloot, P. M. The Diminishing Role of Hubs in Dynamical Processes on Complex Networks. *Journal of the Royal Society, Interface / the Royal Society* **10Q,** 20130568. ISSN: 1742-5662. pmid: 24004558 (2013).

23. James, R. G., Barnett, N. & Crutchfield, J. P. Information Flows? A Critique of Transfer Entropies. *Physical Review Letters* **116,** 1–6. ISSN: 10797114. pmid: 27341264 (2016).

24. Beer, R. D. & Williams, P. L. Information Processing and Dynamics in Minimally Cognitive Agents. *Cognitive Science* **39,** 1–38. ISSN: 1551-6709 (2015).

25. Kolchinsky, A. A Novel Approach to the Partial Information Decomposition. *Entropy* **24,** 403. ISSN: 1099-4300 (Mar. 13, 2022).

26. Williams, P. L. & Beer, R. D. *Nonnegative Decomposition of Multivariate Information* 2010. arXiv: 1004.2515. http://arxiv.org/abs/1004.2515.

27. Forgoston, E. & Moore, R. O. A Primer on Noise-Induced Transitions in Applied Dynamical Systems. *SIAM Review* **60,** 969–1009. ISSN: 0036-1445, 1095-7200. arXiv: 1712.03785 (Jan. 2018).

28. Calim, A., Palabas, T. & Uzuntarla, M. Stochastic and Vibrational Resonance in Complex Networks of Neurons. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* **379,** rsta.2020.0236, 20200236. ISSN: 1364-503X, 1471-2962 (May 31, 2021).

29. Czaplicka, A., Holyst, J. A. & Sloot, P. M. Noise Enhances Information Transfer in Hierarchical Networks. *Scientific Reports* **3.** ISSN: 20452322 (2013).

30. Lizier, J. T., Prokopenko, M. & Zomaya, A. Y. The Information Dynamics of Phase Transitions in Random Boolean Networks, 9 (2008).

31. Lizier, J. T., Flecker, B. & Williams, P. L. Towards a Synergy-Based Approach to Measuring Information Modification. *IEEE Symposium on Artificial Life (ALIFE)* **2013-Janua,** 43–51. ISSN: 21606382. arXiv: 1303.3440 (January 2013).

32. Lizier, J. T., Bertschinger, N., Jost, J. & Wibral, M. Information Decomposition of Target Effects from Multi-Source Interactions: Perspectives on Previous, Current and Future Work. *Entropy* **20,** 307 (4 Apr. 2018).

33. Quax, R., Har-Shemesh, O. & Sloot, P. M. Quantifying Synergistic Information Using Intermediate Stochastic Variables. *Entropy* **19,** 7–10. ISSN: 10994300. arXiv: 1602.01265 (2017).

34. Lizier, J. T., Prokopenko, M. & Zomaya, A. Y. Information Modification and Particle Collisions in Distributed Computation. *Chaos: An Interdisciplinary Journal of Nonlinear Science* **20,** 037109. ISSN: 1054-1500, 1089-7682 (Sept. 2010).

35. Scheffer, M. *et al.* Early-Warning Signals for Critical Transitions. *Nature* **461,** 53–9. ISSN: 1476-4687. pmid: 19727193 (2009).

36. Prokopenko, M., Lizier, J. T., Obst, O. & Wang, X. R. Relating Fisher Information to Order Parameters. *Physical Review E* **84,** 041116. ISSN: 1539-3755, 1550-2376 (Oct. 13, 2011).

37. Scheffer, M., Carpenter, S., Foley, J. A., Folke, C. & Walker, B. Catastrophic Shifts in Ecosystems. *Nature* **413,** 591–596. ISSN: 0028-0836, 1476-4687 (Oct. 2001).

38. Eason, T., Garmestani, A. S. & Cabezas, H. Managing for Resilience: Early Detection of Regime Shifts in Complex Systems. *Clean Technologies and Environmental Policy* **16,**

773–783. ISSN: 1618-954X, 1618-9558 (Apr. 2014).

39. Schreiber, M. *Volume 1 Edited by K. Krickeberg· R. C. Lewontin . J. Neyman M. Schreiber* ISBN: 978-3-642-46246-7 ().
40. Ay, N. & Polani, D. Information Flows in Causal Networks. *Advances in Complex Systems* **11,** 17–41. ISSN: 0219-5259. pmid: 15823657 (2008).
41. Runge, J. *et al.* Inferring Causation from Time Series in Earth System Sciences. *Nature Communications* **10,** 1–13. ISSN: 20411723. pmid: 31201306 (2019).
42. Li, C. Functions of Neuronal Network Motifs. *Physical Review E* **78,** 037101. ISSN: 1539-3755, 1550-2376 (Sept. 3, 2008).
43. James, R. G., Barnett, N. & Crutchfield, J. P. Information Flows? A Critique of Transfer Entropies. *Physical Review Letters* **116,** 1–6. ISSN: 10797114. pmid: 27341264 (2016).
44. Bialek, W. & Tishby, N. *Predictive Information* Feb. 25, 1999. arXiv: `cond-mat/9902341`. `http://arxiv.org/abs/cond-mat/9902341` (2022).
45. López-Ruiz, R., Mancini, H. L. & Calbet, X. A Statistical Measure of Complexity. *Physics Letters A* **209,** 321–326. ISSN: 03759601 (1995).
46. Virtanen, P. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods* **17,** 15 (2020).

8 Author contribution

Casper van Elteren: first draft, (code) implementation, visualization. **Rick Quax**: feedback, supervision, conceptualization. **Peter Sloot**: feedback, conceptualization.

A Appendix

A.1 Background, scope & innovation

Noise induced transitions produces produces metastable behavior that is fundamental for the functioning of complex dynamical systems. For example, in neural systems, the presence of noise increases information processing. Similarly, the relation between glacial ice ages and earth eccentricity has been shown to have a strong correlation. Metastability manifests itself by means of noise that can be of two kinds [27]. External noise originates from events outside the internal system dynamics [28, 29]. Examples include the influence of climate effects, population growth or a random noise source on a transmission line. External noise is commonly modeled by replacing an external control or order parameter by a stochastic process. Internal noise, in contrast, is inherent to the system itself and is caused by random interactions of elements of the system, e.g. individuals in a population, or molecules in chemical processes. Both types of noise can generate transitions between one metastable state and another. In this paper, the metastable behavior is studied of internal noise in complex dynamical networks governed by the kinetic Ising dynamics.

The ubiquity of multistability in complex systems calls for a general framework to understand *how* metastable transitions occur. The diversity of complex systems can be captured by an interaction networks that dynamically evolves over time. These dynamics can be seen as a distributive network of computational units, where each unit or element of the interaction network changes it state based on the input it gets from its local neighborhood. Lizier proposed that these proposed that the dynamic interaction of complex systems can be understood by their local information processing [30–32]. Instead of describing the dynamics of the system in terms of their domain knowledge such as voltage over distance, disease spreading rate, or climate conditions, one can understand the dynamics in terms of the *information dynamics*. In particular, the field of information dynamics is concerned with describing the system behavior along its capacity to store information, transmit information, and modify information. By abstracting away the domain details of a system and recasting the dynamics in terms of *how* the system computes its next state, one can capture the intrinsic computation a system performs. The system behavior is encoded in terms of probability, and the relationship among these variables are explored using the language of information theory [33].

Information theory offers profound benefits over traditional methods used in meta-stability analysis as the methods developed are model-free, can capture non-linear relationships, can be used for both discrete and continuous variables, and can be estimated directly from data [20]. Shannon information measures such as mutual information and as well as Fisher information can be used to study how much information the system dynamics shares with the control parameter [13, 34].

Past research on information flows and metastable transitions focuses on methods to detect the onset of a tipping point [35–37]. It often centers around an observation that the system's ability to absorb noise reduces prior to the system going through a critical point. This critical slowing down, can be captured as a statistical signature where the Fisher information peaks [38]. However, these methods traditionally use some form of control parameter driving the system towards or away from a critical point. Most real-world systems lack such an explicit control parameter and require different methods. Furthermore, detecting a tipping point does not necessarily lead to further understanding how the tipping point was created. For example, for a finite size Ising model, the system produces bistable behavior. As one increases the noise parameter, the bistable behavior disappears. The increase in noise effectively changes the energy landscape, but little information is gained as to how initially the metastable behavior emerged.

In this work, a novel approach using information theory is explored to study metastable behavior. The statistical coherence between parts of the system are quantified by the the capability of individual nodes to predict the future behavior of the system [31]. Two information features are introduced. *Integrated mutual information* measure predictive information of a node on the future of the system. *Asymptotic information measures* the long timescale memory capacity of a node. These measures differ from previous information methods such as transfer entropy [39], conditional mutual information under causal intervention [40], causation entropy [41], and time-delayed variants [42] in that these methods are used to infer the transfer of information between sets of nodes by possible correcting for a third variable. Here, instead, we aim to understand how the elements in the system contribute to the macroscopic properties of the system. It is important to emphasize that information flows are not directly comparable to causal flows [43]. A rule of thumb is that causal flows focus on micro-level dynamics (X causes Y), whereas information flows focus on the predictive aspects, a holistic view of emergent structures [31]. In this sense, this work is similar to predictive information [44] where predictive information of some system S is projected onto its consistent elements $s_i \in S$ and computed as a function of time t.

A.2 Methods and definitions

A.2.1 Model

To study metastable behavior, we consider a system as a collection of random variables $S = \{s_1, \ldots, s_n\}$ governed by the Boltzmann-Gibbs distribution

$$p(S) = \frac{1}{Z}\exp(-\beta \mathcal{H}(S)),$$

where is the inverse temperature $\beta = \frac{1}{T}$ which control the noise in the system, $\mathcal{H}(S)$ is the system Hamiltonian which encodes the node-node dynamics. The choice of the energy function dictates what kind of system behavior we observe. Here, we focus on arguable the simplest models that shows metastable behavior: the kinetic Ising model, and the Susceptible-Infected-Susceptible model.

Temporal dynamics are simulated using Glauber dynamics sampling. In each discrete time step a spin is randomly chosen and a new state $X' \in S$ is accepted with probability

$$p(\text{accept}X') = \frac{1}{1 + \exp(-\beta \Delta E)}, \tag{5}$$

where $\Delta E = \mathcal{H}(X') - \mathcal{H}(X)$ is the energy difference between the current state X and the proposed state X'.

A.2.2 Kinetic Ising model

The traditional Ising model was originally developed to study ferromagnetism, and is considered one of the simplest models that generate complex behavior. It consists of a set of binary distributed spins $S = \{s_1, \dots s_n\}$. Each spin contains energy given by the Hamiltonian

$$\mathcal{H}(S) = -\sum_{i,j} J_{ij} s_i s_j - h_i s_i. \tag{6}$$

where J_{ij} is the interaction energy of the spins s_i, s_j.

The interaction energy effectively encodes the underlying network structure of the system. Different network structures are used in this study to provide a comprehensive numerical overview of the relation between network structure and information flows (see A.2). The interaction energy J_{ij} is set to 1 if a connection exists in the network.

For sufficiently low noise (temperature), the Ising model shows metastable behavior (fig. 1c). Here, we aim to study *how* the system goes through a tipping point by tracking the information flow per node with the entire system state.

A.3 Information flow on complex networks

Informally, the information flows measures the statistical coherence between two random variables X and Y over time such that the present information in Y cannot be explained by the past of Y but rather by the past of X. Estimating information flow is inherently difficult due to the presence of confounding which potential traps the interpretation in the "correlation does not equal causation". Under some context, however, information flow can be interpreted as causal [21]. Let $S = \{s_1, \dots, s_n\}$ be a random process, and S^t represent the state of the random process at some time t. The information present in S is given as the Shannon entropy

$$H(S) = -\sum_{x \in S} p(x) \log p(x) \tag{7}$$

where $\log$ is base 2 unless otherwise stated, and $p(x)$ is used as a short-hand for $p(S = x)$. Shannon entropy captures the uncertainty of a random variable; it can be understood as the number of yes/no questions needed to determine the state of S. This measure of uncertainty naturally extends to two variables with Shannon mutual information. Let s_i be an element of the state of S, then the Shannon mutual information $I(S; s_i)$ is given as

$$\begin{aligned} I(S; s_i) &= \sum_{S_i \in S, s' \in s_i} p(S_i, s') \log \frac{p(S_i, s')}{p(S_i)p(s')} \\ &= H(S) - H(S|s_i) \end{aligned} \tag{8}$$

Shannon mutual information can be interpreted as the uncertainty reduction of S after knowing the state of s_i. Consequently, it encodes how much statistical coherence s_i and S share. Shannon mutual information can be measured over time to encode how much *information* (in bits) flows from state s_i^τ to $S^{\tau+t}$

$$I(S^{\tau+t}; s_i^\tau) = H(S^{\tau+t}) - H(S^{\tau+t}|s_i^\tau). \tag{9}$$

Prior results showed that the nodes with the highest causal importance are those nodes that have the highest information flow (i.e. maximize 9) [21]. Intuitively, the nodes for which the future system "remembers" information from a node in the past, is the one that "drives" the system dynamics. Formally, these driver nodes can be identified by computing the total information flow between S^t and s_i can be captured with the integrated mutual information [21]

$$\mu(s_i) = \sum_{\tau=0}^{\infty} I(s_i^{t-\tau}; S^t). \tag{10}$$

In some context, the nodes that maximizes the (10) are those nodes that have the highest causal influence in the system [21]. However in general information flows are difficult to equate to causal flows [31, 43]. Here, the local information flows are computed by considering the integrated mutual information conditioned on part of the entire state space. This allows for mapping the local information flows between nodes and the system over time, but does not guarantee that the measured information flows are directly causal. The main reason being that having predictive power about the future, could be completely caused by the partitioning. In [21] the correlation measured considered all possible states, and the measures were directly related to a causal effect.

In addition, in [21] the shared information between the system with a node shifted over time ($I(S^\tau : s_i^{\tau+t})$) was considered. Applying this approach under a state partition $I(S^\tau : s_i^{\tau+t}|\langle S\rangle)$causes a violation of the data processing as information may flow from a node at a particular $t = t_1$ and then flow back to the node at $t = t2, t_2 > t_1$. In order to simplify the interpretation of the information flows and keep the data processing inequality, the reverse $I(S^{t+\tau} : s_i^\tau|\langle S\rangle)$ was computed in the present study.

A.4 Noise matching procedure

The Boltzmann-Gibbs distribution is parameterized by noise factor $\beta = \frac{1}{kT}$ where T is the temperature and k is the Boltzmann constant. For high β values metastable behavior occurs in the kinetic Ising model. The temperature was chosen such that the statistical complexity [45] was maximized. The statistical complexity C is computed as

$$C = \bar{H}(S)D(S),$$

where $\bar{H}(S) = \frac{H(s)}{-\log_2(|S|)}$ is the system entropy, and $D(S)$ measures the distance to disequilibrium

$$D(S) = \sum_i (p(S_i) - \frac{1}{|S|})^2.$$

A typical statistical complexity curve is seen in fig. A.7. The noise parameter β is set such that it maximizes the statistical complexity using numerical optimization (COBYLA method in scipy's `optimize.minimize` module) [46].

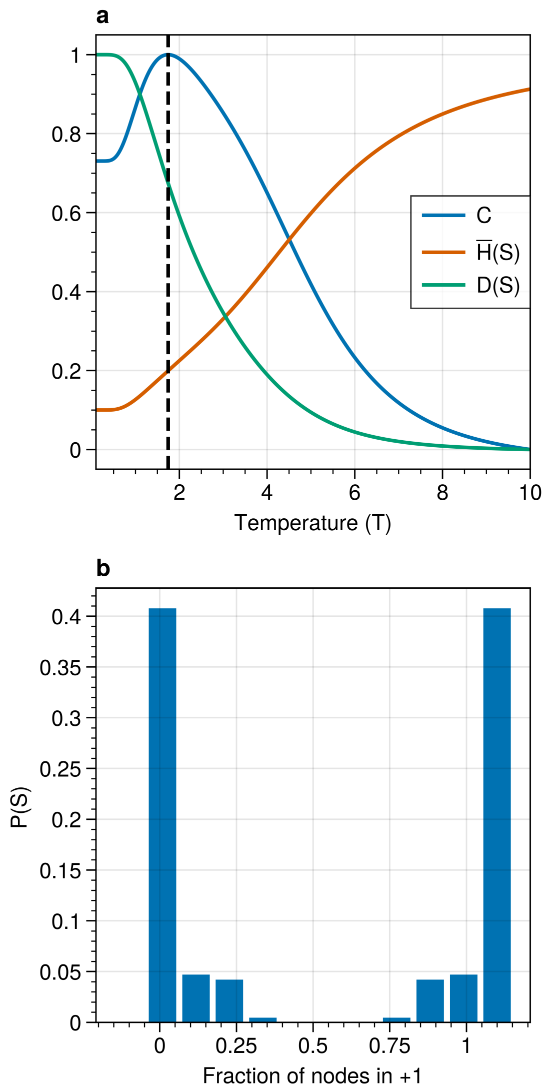

Figure A.7: *(a) Statistical complexity (C), normalized system entropy ($H(S)$) and disequilibrium ($D(S)$) as a function of the temperature ($T = \frac{1}{\beta}$) for Krackhardt kite graph. The noise parameter was set such that it maximizes the statistical complexity (vertical black line). The values are normalized between [0,1] for aesthetic purposes. (b) State distribution $p(S)$ for temperature that maximizes the statistical complexity in (a) as a function of nodes in state 1.*

A.5 Exact information flows $I(s_i^\tau; S^{\tau+t})$

In order to compute $I(s_i^\tau : S^{\tau+t})$, the conditional distribution $p(S^{\tau+t}|s_i^\tau)$ and $p(S^{\tau+t})$ needs to be computed. For Glauber dynamics, the system S transitions into S' by considering to flips by randomly choosing node s_i. The transition matrix $p(S^t|s_i) = \mathbf{P}$ can be constructed by computing each entry p_{ij} as

$$p_{ij,i\neq j} = \frac{1}{|S|}\frac{1}{1+\exp(-\Delta E)}$$
$$p_{ii} = 1 - \sum_{j,j\neq i} p_{ij},$$

where $\Delta E = \mathcal{H}(S_j) - \mathcal{H}(S_j)$ encodes the energy difference of moving from S_i to S_j. The state to state transition $\mathbf{P}$ matrix will be of size $2^{|S|} \times 2^{|S|} \times |\mathcal{A}_{s_i}|$, where $|\mathcal{A}_{s_i}|$ is the size of the alphabet of s_i, which becomes computationally intractable due to its exponential growth with the system size $|S|$. The exact information flows can then be computed by evaluating $p(S^t|s_i)$ out of equilibrium by evaluating all S^t for all possible node states s_i where $p(S^t)$ is computed as

$$p(S^{\tau+t}) = \sum_{s_i} p(S^{\tau+t}|s_i^\tau)p(s_i^\tau).$$

A.5.1 Extrapolation with regressions

Exact information flows were computed per graph for $t = 500$ times steps. Using ordinary least squares a double exponential was fit to estimate the information flows for longer t and estimate the integrated mutual information and asymptotic information.

A.6 Noise estimation procedure

Tipping point behavior under intervention was quantified by evaluating the level of noise on both side of the tipping point. Let $T1$ represent the ground state where all spins are 0, $T2$ where all spins, and the tipping point TP is where the instantaneous macrostate $M(S^t) = 0.5$. Fluctuations of the system macrostate was evaluated by analyzing the second moment above and below the tipping point. This was achieved by numerically simulating the system trajectories under 6 different seeds for $t = 10e6$ time-steps. The data was split between two sets (above and below the tipping point) and the noise η was computed as

$$\eta = \frac{1}{\alpha^2|S_w|} \sum_w {S_w^t}^2,$$

where $w \in \{\langle S \rangle < 0.5, \langle S \rangle > 0.5\}$, and

$$S_w^t = \begin{cases} S^t \text{ if } S^t < 0.5 \\ 1 - S^t \text{ if } S^t > 0.5 \end{cases} \tag{11}$$

is the instantaneous system trajectory for the system macrostate above or below the tipping point value. The factor α corrects for the reduced range the system macrostate has under interventions. For example pinning a node s_i to state 0, reduces the maximum possible macrostate to $1 - \frac{1}{n}$ where n is the size of the system. The correction factor α is set such that for an intervention on 0 for a particular node, the range $S_{\langle S \rangle > 0.5}$ alpha is set to $\frac{n}{2} - \frac{1}{n}$.

A.7 Switch susceptibility as a function of degree

First, we investigate the susceptibility of a spin as a function of its degree. The susceptibility of a spin switching its state is a function both of the system temperature T and the system dynamics. The system dynamics would contribute to the susceptibility through the underlying network structure either directly or indirectly. The network structure produces local correlations which affects the switch probability for a given spin.

As an initial approximation, we consider the susceptibility of a target spin s_i to flip from a majority state to a minority state given the state of its neighbors where the neighbors are not connected among themselves. Further, the assumption is that for the instantaneous update of s_i the configuration of the neighborhood of s_i can be considered as the outcome of a binomial trial. Let, N be a random variable with state space $\{0, 1\}^{|N|}$, and let $n_j \in N$ represent a neighbor of s_i. We assume that all neighbors of s_i are i.i.d. distributed given the instantaneous system magnetization

$$M(S^t) = \frac{1}{|S^t|} \sum_i s_i^t.$$

Let the minority state be 1 and the majority state be 0, the expectation of s_i flipping from the majority state to the minority state is given as:

$$\begin{aligned} E[p(s_i = 1|N)]_{p(N)} &= \sum_{N_i \in N} p(N_i)p(s_i = 1|N_i) \\ &= \sum_{N_i \in N} \prod_j^{|N_i|} p(n_j)p(s_i = 1|N_i) \\ &= \sum_{N_i \in N} \binom{n}{k} f^k(1-f)^{n-k} p(s_i = 1|f), \end{aligned} \tag{12}$$

where f is the fraction of nodes in the majority states, n is the number of neighbors, k is the number of nodes in state 0. In fig. A.9. This is computed as a function of the degree of spin s_i. As the degree increases, the susceptibility for a spin decreases relatively to the same spin with a lower degree. This implies that the susceptibility of change to random fluctuations are more likely to occur in nodes with less external constraints as measured by degree.

A.8 Additional networks

The kite graph was chosen as it allowed for computing exact information flows while retaining a high variety of degree distribution given the small size. Other networks were also tested. In fig. A.8) different network structure were used. Each node is governed by kinetic Ising spin dynamics.

A.9 Flip probability per degree

In fig. A.9 the tendency for a node to flip from the majority to the minority state is computed as function of fraction of nodes possessing the majority states 1 in the system, denoted as N. Two things are observed. First, nodes with lower degree are more susceptible to noise than nodes with higher degree. For a given system stability, nodes with lower degree tend to have a higher tendency to flip. This is true for all distances of the system to the tipping point. In contrast, the higher the degree of the node, the closer the system has to be to a tipping point for the node to change its state. This can be explained by the fact that lower degree nodes, have fewer constraints compared to nodes with higher degree nodes. For Ising spin kinetics, the nodes with higher degree tend to be more "frozen" in their node dynamics than nodes with lower degree. Second, in order for a node to flip with probability with similar mass, i.e. ($E[p(s_i)|N] = 0.2$) a node with higher degree needs to be closer to the tipping point than nodes with lower degree. In fact, the order of susceptibility is correlated with the degree; the susceptibility decreases with increasing degree and fixed fraction of nodes in state 1.

A.10 Synthetic networks

For the synthetic graphs, 100 non-isomorphic connected Erdos-Renyi networks were generated with a p = 0.2. Graphs were generated randomly and rejected if the graph did not contain a giant component, or was isomorphic with already generated graphs. For each of the graphs, information curves were computed as function of the macrostate $\langle S \rangle$.

A.10.1 Noise and time spent

Various network structures are generated in the synthetic networks. The variety of network structure has non-linear effects on the information flows. The effect of intervention in fig. 6 is made relative to the control values for the graph and seed. The second moment (appendix: A.6) and the time spent below the tipping point are normalized with respect to the graph (fig. A.11) and the seed. In total 6 seeds are used (0, 12, 123, 1234, 123456, 1234567).

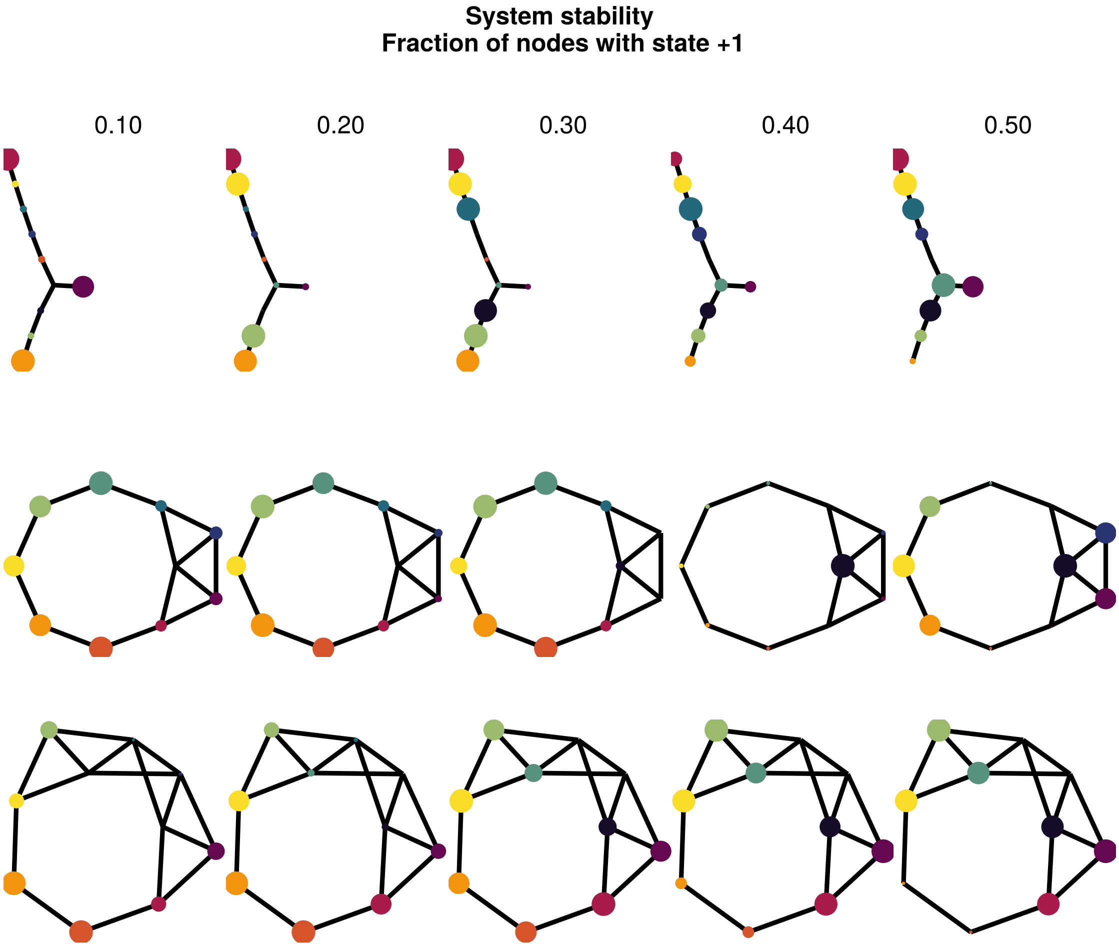

Figure A.8: *Adjusted mutual information for a random tree (top), and Leder-Coxeter Fruchte graphs (middle, bottom). Each node is goverened by kinetic Ising spin dyanmics. Far away from the tipping point (fraction nodes 1 = 0.5) most information flows are concentrated on non-hub nodes. As the system approaches the tipping point (fraction = 0.5), the information flows move inwards, generating higher adjusted integrated mutual information for nodes with higher degree.*

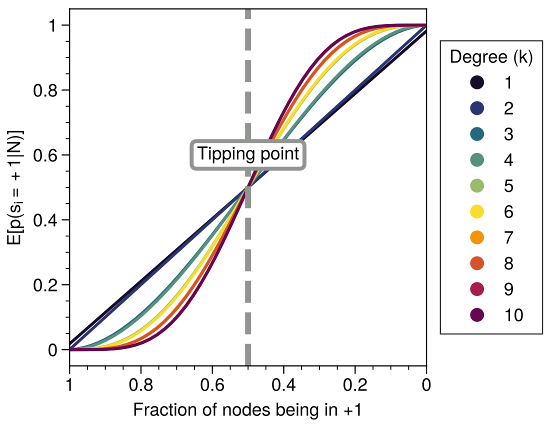

Figure A.9: *Susceptibility of a node with degree k switching from the minority state 0 to the majority state 1 as a function of the neighborhood entropy for $\beta = 0.5$. The neighborhood entropy encodes how stable the environment of a spin is. As the system approaches the tipping point, the propensity of a node to flip from to the minority state increases faster for low degree nodes than for high degree nodes. Higher degree nodes require more change in their local environment to flip to the majority state. See for details A.7.*

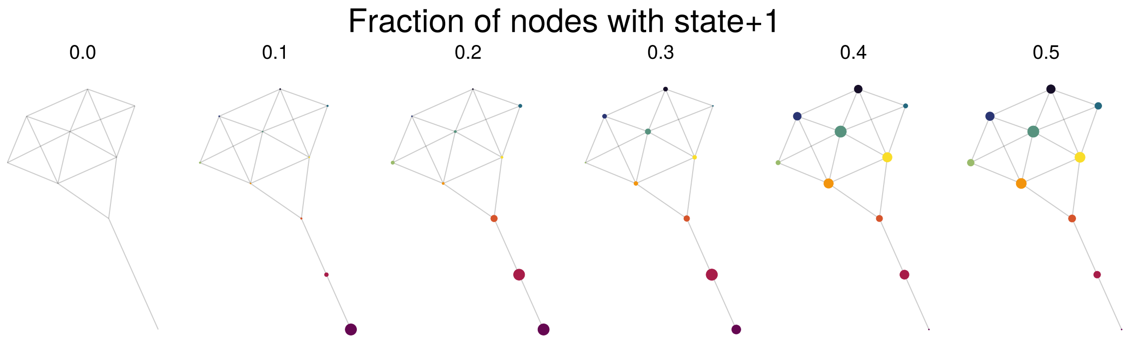

Figure A.10: *Shortest path analysis of the system ending up in the tipping point from the state where all nodes have state 0. The node size is proportional to the expectation value of a node having state 1 ($E[s_i = 1]_{S^t, M(S^5)}$ as a function of the fraction of nodes having state 1. The expectation values are computed based on 30240 trajectories, an example trajectory can be seen in fig. 3.*

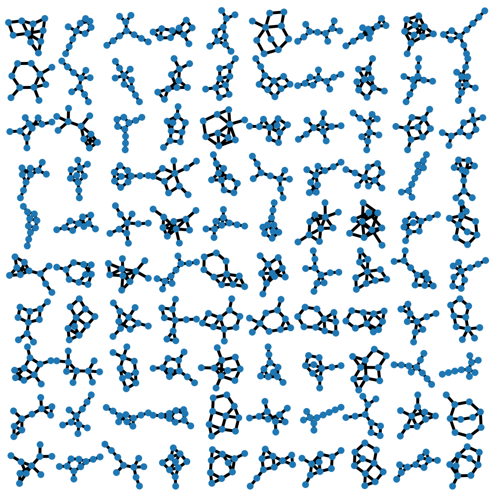

Figure A.11: *Erdos-Renyi graphs generated from seed = 0 to produce non-isomorphic connected graphs.*