

**Harnessing Stochastic Forces:
Exploring Noise-Induced Phenomena
in Physarum-Inspired Network Design
Models**

DISSERTATION

ZUR ERLANGUNG DES AKADEMISCHEN GRADES
DES DOKTORS DER NATURWISSENSCHAFTEN
DER NATURWISSENSCHAFTLICH-TECHNISCHEN FAKULTÄT
DER UNIVERSITÄT DES SAARLANDES

von
Frederic Christopher Folz

Saarbrücken
2025

Tag des Kolloquiums: 20.08.2025

Dekan: Prof. Dr.-Ing. Dirk Bähre

Berichterstatter: Prof. Dr. Giovanna Morigi

Prof. Dr. Adish Singla

Prof. Dr. Dr. Karsten Kruse

Akad. Mitarbeiter: Priv.-Doz. Dr. Mohammad Reza Shaebani

Vorsitz: Prof. Dr. Albrecht Ott

EIDESSTATTLICHE ERKLÄRUNG

Hiermit versichere ich an Eides statt, dass ich die vorliegende Arbeit selbstständig und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Die aus anderen Quellen oder indirekt übernommenen Daten und Konzepte sind unter Angabe der Quelle gekennzeichnet. Die Arbeit wurde bisher weder im In- noch im Ausland in gleicher oder ähnlicher Form in einem Verfahren zur Erlangung eines akademischen Grades vorgelegt.

Ort, Datum

Unterschrift

ZUSAMMENFASSUNG

In dieser Dissertation untersuchen wir Modelle zur Entwicklung von Transportnetzwerken, die von dem paradigmatischen Organismus *Physarum polycephalum* inspiriert sind. Neben der Charakterisierung der erzeugten Netzwerke im deterministischen Fall analysieren wir, wie das Hinzufügen stochastischer Kräfte die topologischen Eigenschaften der entstehenden Netzwerke sowie deren Anpassungsfähigkeit an sich dynamisch ändernde Umweltbedingungen beeinflusst. Als zentrale Ergebnisse berichten wir von rauschinduzierten Resonanzen, die die Anpassungsfähigkeit an externe Faktoren sowie die Robustheit, Effizienz und Kosten der generierten Netzwerke bei endlicher Rauschamplitude optimieren. Darüber hinaus beleuchten wir die Rolle des Zusammenspiels von nichtlinearer Dynamik und stochastischen Kräften bei der Entstehung dieser rauschinduzierten Phänomene, was entscheidende Einblicke in Optimierungsalgorithmen im Allgemeinen bietet. Unsere Ergebnisse hinterfragen die traditionelle Betrachtung von Rauschen als Störfaktor und zeigen vielmehr, dass es sogar von Vorteil sein kann.

ABSTRACT

In this dissertation, we study models for the design of transport networks inspired by the paradigmatic organism *Physarum polycephalum*. Beyond characterizing the networks generated in a deterministic setting, we analyze how adding stochastic forces influences the topological properties of the emerging networks as well as their adaptability to dynamically changing environmental conditions. As key results, we report on noise-induced resonances, optimizing the adaptability to external factors as well as the robustness, efficiency, and cost of the resulting networks for finite noise amplitudes. Further, we shed light on the role of the interplay of nonlinear dynamics and stochastic forces in the emergence of these noise-induced phenomena, providing crucial insights for optimization algorithms in general. Our results challenge the traditional notion of noise being detrimental, demonstrating that it can even be used as a resource.

CONTENTS

Introduction	1
1 Interplay of periodic dynamics and noise: insights from a simple adaptive system	18
2 Physarum-inspired multi-commodity flow dynamics	28
3 Noise-induced network topologies	49
4 Self-organized transport in noisy dynamic networks	60
Conclusion and outlook	77
References	84

INTRODUCTION

I.1 Research context

I.1.1 Complex systems and networks

Over the last decades, the study of complex systems has gained interdisciplinary significance. These systems consist of numerous components whose interactions give rise to collective phenomena that are not manifest in the behavior of the individual constituents [1–4]. To describe complex systems, networks serve as paradigmatic models across various fields. For instance, networks are used in the following disciplines, with one of several possible applications given as an example for each:

- in social sciences, for mapping interactions among people, where individuals represent components and their relationships interactions [5–7], facilitating the analysis of information spread [8];
- in biology, for characterizing food webs, with species corresponding to components and predator-prey relations to interactions [9–11], revealing the dynamics of ecosystems [12];
- in neurosciences, for describing the structure of the human brain, consisting of neurons (components) that are connected through synapses (interactions) [13–15], helping to understand diseases [16];
- in physics, for assessing the distribution of galaxies, with the latter as components and gravitational forces as interactions [17, 18], unraveling the large-scale structure of the universe [19];
- in engineering, for designing power grids, where power plants, transformers, and consumers represent components and transmission lines interactions [20], enabling the optimization of electricity supply [21].

These examples are vastly different, yet the formation of the networks describing them is governed by universal principles [22]. For instance, in several situations, the networks evolve following the preferential attachment rule, which means that new components are more likely to establish interactions with components that already interact with many other components [23]. This mechanism leads to the emergence of common characteristics such as hubs—components interacting with a considerably larger number of other components than on average—and a scale-free degree distribution, where the number of interactions per component scales as a power law. Understanding these principles offers insights into the structure of a network, which sheds light on its function [24] and directly affects characteristics like resilience and adaptability [25–27]. Moreover, a network’s structure might give rise to collective phenomena, which further influence these properties. Examples of such phenomena include synchronization [28, 29], like the coordinated flashing of fireflies [30], symmetry breaking [31, 32], as occurs in ferromagnetic materials [33], and pattern

formation [34, 35], such as the emergence of oscillatory patterns [36]. For instance, in power grids, oscillations can arise from varying system loads, affecting the stability of the whole grid and potentially leading to blackouts [37]. This demonstrates how important careful network design is, especially for transport networks like power grids, as they form the backbone of the economy and daily life, enabling the distribution of people, goods, and resources, for example. In such real-world scenarios, a common objective is to design transport networks that achieve short transit times and resilience against disruptions while managing resources sustainably to reduce costs and environmental impact [38]. To address this key challenge, various approaches have been developed, ranging from agent-based models that simulate interactions between transported entities [39, 40], to genetic algorithms that iteratively evolve a network configuration [41, 42], and entropy maximization strategies that align network structure with maximum utility [43–45].

While most of these approaches are deterministic, real-world transport networks are typically exposed to random fluctuations, referred to as *noise* in the following. Noise can significantly impact their function and stability [46, 47]. Traditionally, it is viewed as detrimental because it often perturbs the flow of entities like information, goods, or resources, hindering their distribution. For example, stochastic fluctuations typically overlay analog signals and thereby lower the effectiveness of transmission. Yet, recent research indicates that noise can even play a beneficial role in transport networks, enhancing their efficiency and robustness [48–50]. This is particularly evident in transport networks formed by biological systems, which have evolved in noisy environments and are thus naturally optimized to work efficiently in them [51, 52]. Prominent examples include ant colonies [53–55] and primitive organisms like *Physarum polycephalum*, a representative of the true slime molds [56–58]. The latter is able to construct efficient and resilient networks of filaments for transporting nutrients with minimal use of resources—a capability acquired through hundreds of millions of years of natural selection [59]. In addition to that, *P. polycephalum* is highly adaptive to changing external conditions [60]. These features render the slime mold an ideal model organism to study the self-organization of networks in the presence of stochastic fluctuations. In particular, the slime mold’s remarkable performance in noisy environments raises the more general question of whether stochastic fluctuations could also be harnessed to optimize non-biological network formation. Inspired by this, we now shift to a more abstract, algorithmic perspective and ask: *Can adding noise to models for network design enhance topological properties of the resulting networks, such as robustness and efficiency, as well as their adaptability?*

In this dissertation, we focus on algorithms inspired by *P. polycephalum* with the goal of understanding whether and when stochastic fluctuations might lead to more robust and efficient self-organized network topologies than in the absence of noise. For this purpose, we analyze the networks generated by these algorithms when they are applied to so-called multi-commodity transport problems [61]. This analysis also has broader implications for optimization algorithms in general, addressing the question of how noise might be harnessed as a resource to enhance their performance.

To lay the basis for the discussion of the *Physarum*-inspired algorithms, we first formalize transport problems, especially those that involve the routing of different commodities such as goods, passengers, or information through a network. This requires a mathematical characterization of these networks, which can be achieved by abstracting them as graphs. Consequently, in Secs. I.1.2 and I.1.3, we introduce

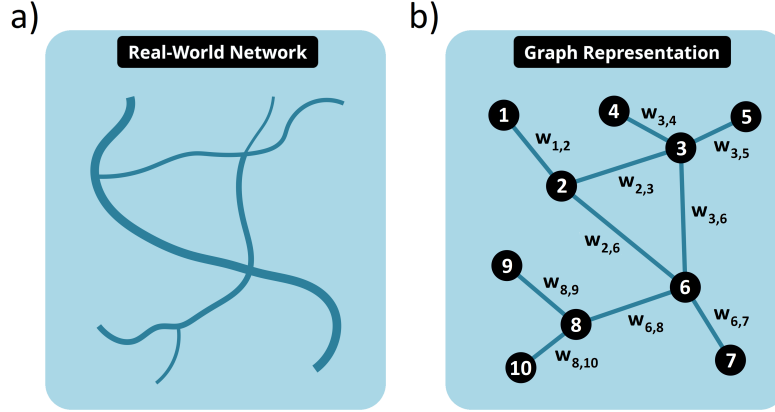


Figure 1: (a) Illustration of a real-world network, such as an urban transport system, a power grid, or a water supply infrastructure. (b) Corresponding graph representation, where edge weights are denoted as $w_{u,v}$.

the basic principles of graph theory. In Sec. I.1.4, we then define transport problems on graphs. Based on that, we discuss multi-commodity flow problems in Sec. I.1.5, providing the framework for our study of Physarum-inspired models for network design, which are introduced in Sec. I.1.6.

I.1.2 Graph representation of networks

Any real-world network can be abstracted as a graph [62], which is illustrated in Fig. 1 with a simple example. Figure 1(a) depicts a generic real-world network, such as an urban transport system, a power grid, or a water supply infrastructure. Figure 1(b) displays the graph representation of the network, where junctions and dead ends (components) of the real-world network correspond to nodes in the graph, and connections (interactions) between them to edges. In addition to that, edges might be assigned weights $w_{u,v}$ to quantify the relationship between the nodes u and v in the real-world network. For the examples mentioned above, the weights could represent transit times, the voltage across transmission lines, or the flow capacity of water pipes. This abstraction enables the application of the abundant toolbox that graph theory provides. For instance, using the graph representation of a network, one can determine its resilience against edge failures [63–65], identify the shortest paths between nodes [66], or calculate its total length [67].

In order to perform a solid mapping, it is necessary to first establish key definitions in graph theory, which we draw from West [68] in the following. We begin with the difference between undirected and directed graphs, as this distinction influences how interactions within the network are interpreted. An undirected graph is defined as a pair $G = (V, E)$, where V is a set of nodes and $E \subseteq \{\{u, v\} \mid u, v \in V, u \neq v\}$ a set of unordered pairs of these nodes, referred to as edges. Since $\{u, v\} = \{v, u\}$, the edge linking node u to node v is identical to the edge connecting node v to node u , implying a bidirectional relation between u and v . In contrast, for a directed graph, the set of edges is defined as $E \subseteq \{(u, v) \mid u, v \in V, u \neq v\}$, containing ordered pairs of nodes. As $(u, v) \neq (v, u) \forall u \neq v$, the edge from u to v and the edge from v to u are different, indicating that each edge corresponds to a unidirectional connection. Thus, to model a bidirectional relation between u and v , both edges (u, v) and (v, u) are required.

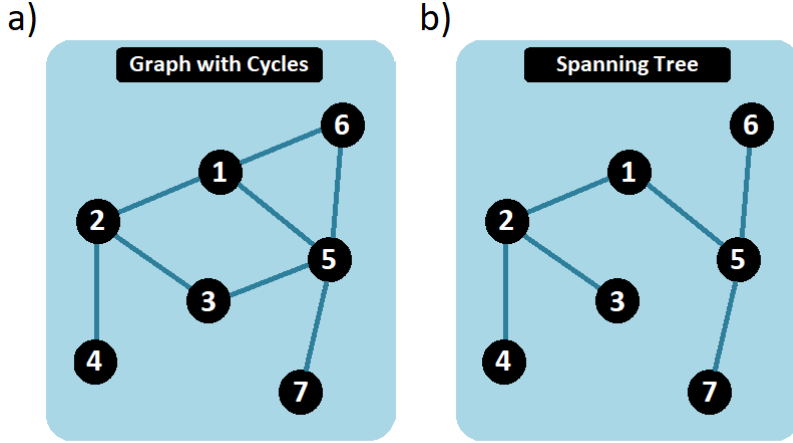


Figure 2: (a) A graph containing three cycles. (b) Spanning tree obtained by removing two edges from the graph.

For a detailed analysis of real-world networks, we need to introduce additional definitions to further characterize the structure of the underlying graphs. Therefore, we will discuss central concepts such as paths, cycles, and trees in the following section, providing important tools for evaluating topological properties like robustness and efficiency.

I.1.3 Paths, cycles and trees in graphs

In this section, we continue to use definitions from West [68], except where otherwise noted. First, we introduce a path in a graph as a sequence of edges (e_1, \dots, e_m) that connect a sequence of nodes (v_1, \dots, v_{m+1}) , where edge e_i links nodes v_i and v_{i+1} . For simplicity, the edge e_i is often denoted as $v_i \rightarrow v_{i+1}$ and the path (e_1, \dots, e_m) as $v_1 \rightarrow v_2 \rightarrow \dots \rightarrow v_{m+1}$. A path can be open or closed, depending on whether the first and the last node, v_1 and v_{m+1} , are distinct or identical. If there is at least one path that connects any pair of nodes, we refer to the graph as connected.

Another important concept in graph theory is the cycle, which is a closed path with $m \geq 2$, where v_1 and v_{m+1} are the only identical nodes. Cycles enhance the robustness and efficiency of networks by providing alternative paths between nodes [69], which is particularly relevant for transport systems. This redundancy ensures that removing an edge within a cycle does not disconnect the network. Figure 2(a) illustrates this with a graph containing three cycles, represented by the closed paths $1 \rightarrow 2 \rightarrow 3 \rightarrow 5 \rightarrow 1$, $1 \rightarrow 5 \rightarrow 6 \rightarrow 1$, and $1 \rightarrow 2 \rightarrow 3 \rightarrow 5 \rightarrow 6 \rightarrow 1$. Notably, the latter cycle is a combination of the former two, which leads us to the concept of the cycle space.

To introduce this concept, we first define a subgraph as a subset of the nodes and edges of the original graph. The set of all subgraphs where each node is connected to an even number of edges forms the cycle space of the graph. Let $B_c = \{c_1, \dots, c_k\}$ be a basis of the cycle space, meaning that any element of the cycle space can be expressed as a combination of the independent cycles c_1, \dots, c_k . The dimension of the cycle space is then given by

$$\begin{aligned} n_c &= \dim(B_c) \\ &= k. \end{aligned} \tag{0.1}$$

To determine the dimension n_c , we need to introduce the concepts of trees and spanning trees. A tree is defined as an undirected, connected graph with exactly one path between any pair of nodes, which means that there are no cycles. A spanning tree of a graph is defined as a subgraph that includes all nodes of the graph and is a tree. It always contains exactly $|V| - 1$ edges, where $|V|$ is the number of nodes. Therefore, in a graph with $|E|$ edges, the number of edges that are not included in a spanning tree is given by $|E| - |V| + 1$. This is illustrated in Fig. 2(b), which shows a spanning tree of the graph displayed in Fig. 2(a). The spanning tree has $|V| - 1 = 6$ edges, while $|E| - |V| + 1 = 2$ edges of the graph are not part of it.

Adding any of these remaining edges to the spanning tree creates a fundamental cycle, which consists of the added edge and the unique path between the nodes in the spanning tree that it connects. These fundamental cycles are independent, meaning that they cannot be expressed as a combination of other cycles. As the set of all fundamental cycles forms a basis of the cycle space, its dimension can be determined using Eq. (0.1), which yields

$$n_c = |E| - |V| + 1. \quad (0.2)$$

Hence, the graph depicted in Fig. 2(a) has the cycle space dimension $n_c = 2$.

Among all possible spanning trees of a graph, a minimum spanning tree (MST) has the smallest total edge weight, defined as the sum of all edge weights. When edge weights correspond to lengths, an MST minimizes the total length of the original graph. This property makes MSTs an important tool in network design, where minimizing connection costs is often a main objective [70, 71]. Various algorithms to determine MSTs have been developed and are widely used in practical applications [72, 73].

To further characterize cycles, we define the length of a cycle as the sum of the lengths L_e of its constituent edges,

$$l_c = \sum_{e \in \{e_1, \dots, e_m\}} L_e. \quad (0.3)$$

In real-world applications, the length is often associated with the cost to construct or maintain the structure.

To quantify the importance of cycles in a graph, we define the cycle edge fraction

$$e_c = \frac{|L|}{|E|}, \quad (0.4)$$

where L is the set of edges that are part of at least one cycle. This fraction can be used as an indicator for the redundancy of connections and, thus, the robustness of a network.

In the context of network science, cycles are often referred to as loops [62]. However, in graph theory, a loop denotes an edge that connects a node to itself. To align with the terminology commonly used in network analysis, we will use the term loops to refer to cycles in the rest of this dissertation.

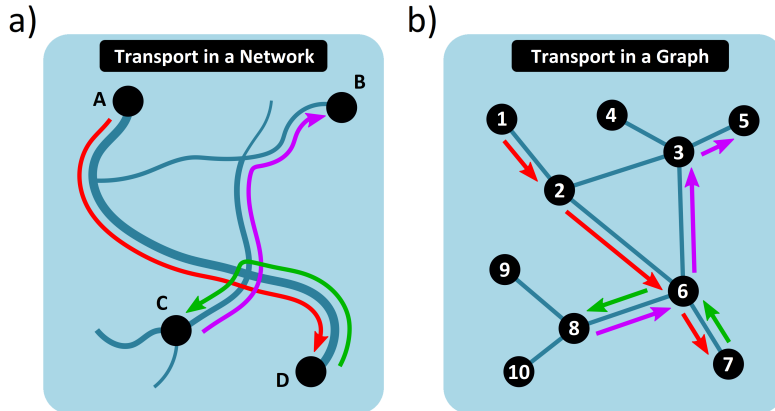


Figure 3: (a) Real-world network connecting four warehouses, labeled A to D, with transport routes indicated by arrows. Different colors represent distinct goods: red for electronics, green for furniture, and purple for groceries. (b) Corresponding transport problem in the graph representation.

I.1.4 Transport problems on graphs

Based on the graph representation of real-world networks and the structural properties of graphs defined in Secs. I.1.2 and I.1.3, we next introduce the simplest form of a transport problem: moving a single type of good, called a commodity, from one node to another [61, 74, 75]. We refer to the quantity of goods and the pair of nodes between which they are transported as a *demand*. However, more realistic scenarios require the simultaneous routing of multiple types of commodities between different pairs of nodes. As an example, Fig. 3(a) illustrates such a transport problem within the real-world network depicted in Fig. 1, assuming that it represents a road network where various goods need to be transported between four warehouses. While electronics (red) may need to be moved from A to D, furniture (green) may need to be carried from D to C, and groceries (purple) may need to be shipped from C to B. Possible transport routes are indicated by colored arrows.

In the graph representation of the network, as shown in Fig. 3(b), each warehouse is a distinct node, and the transport routes are paths. Importantly, paths along which different commodities are transported may share edges in the graph, underscoring the need to handle all demands simultaneously as real-world connections have limited capacity. This leads us to the concept of multi-commodity flow problems, which are commonly studied in computer science to model and solve such transport problems [76–78]. We will dedicate the next section to their formal introduction.

I.1.5 Multi-commodity flow problems

In this section, we follow definitions from Bertsekas [79]. The objective of multi-commodity flow problems is to find the optimal routing of a flow of k commodities through a network described as an undirected graph $G(V, E)$. To facilitate the mathematical formulation, we assign an arbitrary but fixed orientation to each undirected edge $\{u, v\} \in E$, denoting it as (u, v) . Each commodity i is associated with a source node s_+^i at which the commodity enters the network and a sink node s_-^i at which it leaves the network, representing a transport demand I_i between these nodes. Figure 4(a) illustrates these terms using a simple example with $k = 2$ commodities, four nodes (colored circles), and six edges (lines connecting the circles).

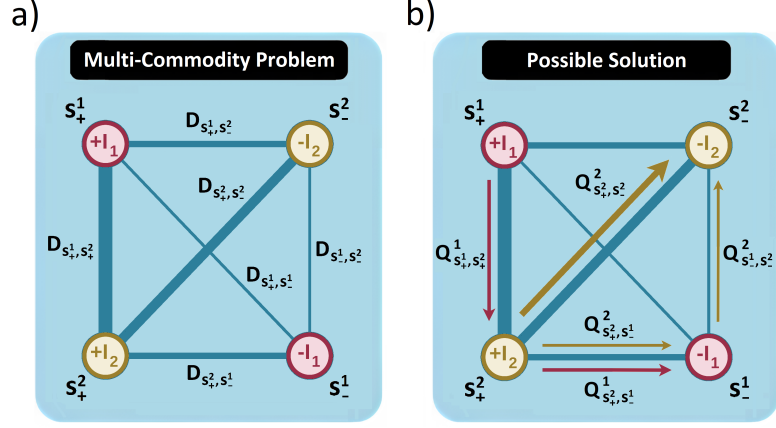


Figure 4: (a) Simple example of a multi-commodity flow problem. In a network of four nodes, two commodities need to be routed. Colored circles represent nodes, and lines denote edges with capacities indicated by their widths. (b) Possible solution of the multi-commodity flow problem. Arrows illustrate flow directions, with their thicknesses indicating flow magnitudes.

For any node $u \in V$, we define the set of all neighbors as

$$N_u = \{v \in V \mid (u, v) \in E\}. \quad (0.5)$$

For instance, in the example of Fig. 4(a), the neighbors of s_+^1 are $N_{s_+^1} = \{s_-^1, s_+^2, s_-^2\}$. The flow of commodity i through edge (u, v) is denoted as $Q_{u,v}^i \in [-I_i, I_i]$ and satisfies

$$Q_{u,v}^i = -Q_{v,u}^i. \quad (0.6)$$

Here, $Q_{u,v}^i > 0$ indicates flow from u to v and $Q_{u,v}^i < 0$ corresponds to flow from v to u . In addition to that, each edge is assigned a commodity-independent capacity $D_{u,v} > 0$, which is represented by the line width in Fig. 4(a).

The flow configuration $\{Q_{u,v}^i \mid (u, v) \in E, i \in \{1, \dots, k\}\}$ solves the problem if it satisfies the following constraints:

1. Flow conservation constraints

To account for the injection and removal of flow at the source and sink nodes and flow conservation at any other node, the relation

$$\sum_{v \in N_u} Q_{u,v}^i = \begin{cases} I_i, & \text{if } u = s_+^i, \\ -I_i, & \text{if } u = s_-^i, \\ 0 & \text{otherwise,} \end{cases} \quad \forall u \in V, \forall i \in \{1, \dots, k\} \quad (0.7)$$

needs to be satisfied.

2. Capacity constraints

To ensure that the total flow through an edge does not exceed its capacity, it must hold

$$\sum_{i=1}^k |Q_{u,v}^i| \leq D_{u,v} \quad \forall (u, v) \in E. \quad (0.8)$$

3. Additional constraints (optional)

By introducing additional constraints such as minimizing a cost function, which quantifies the cost associated with a flow configuration, an optimality criterion can be defined. For instance, a common goal is to minimize the total transport cost

$$C = \sum_{i=1}^k \sum_{(u,v) \in E} c_{u,v}^i |Q_{u,v}^i|, \quad (0.9)$$

where $c_{u,v}^i$ is the cost of transporting a unit flow of commodity i through the edge (u, v) .

Figure 4(b) demonstrates a possible solution for the multi-commodity flow problem presented in Fig. 4(a). The flow I_1 of the first commodity, which is injected at s_+^1 and removed at s_-^1 , is routed along the path $s_+^1 \rightarrow s_+^2 \rightarrow s_-^1$ with

$$Q_{s_+^1, s_+^2}^1 = Q_{s_+^2, s_-^1}^1 = I_1. \quad (0.10)$$

Similarly, the flow I_2 of the second commodity from s_+^2 to s_-^2 is split across the two paths $s_+^2 \rightarrow s_-^2$ and $s_+^2 \rightarrow s_-^1 \rightarrow s_-^2$ with

$$Q_{s_+^2, s_-^2}^2 = 2I_2/3 \quad \text{and} \quad Q_{s_+^2, s_-^1}^2 = Q_{s_-^1, s_-^2}^2 = I_2/3. \quad (0.11)$$

This setup ensures that the flow conservation constraint (0.7) is fulfilled. In Fig. 4(b), arrows indicate the direction of the flows, with their thicknesses being proportional to the flow magnitudes. The combined thickness of all arrows along each edge represents the sum of all flows, which stays within the edge's capacity indicated by the width of the respective line. This demonstrates that the capacity constraint (0.8) is met as well.

I.1.6 Physarum-inspired models for network design

Based on the mathematical framework for multi-commodity transport problems introduced in Sec. I.1.5, we now turn to a class of algorithms for routing optimization that are inspired by the behavior of *Physarum polycephalum*. This single-celled organism grows tubular networks to efficiently transport nutrients absorbed from food sources. Remarkably, despite lacking a nervous system or any other form of centralized control, *P. polycephalum* can find approximate solutions to well-known optimization problems, including the traveling salesman problem [80] and maze-solving tasks. The latter were investigated by Tero et al. [56], who demonstrated that a key principle underlying the slime mold's behavior is a local feedback mechanism that guides global self-organization.

These abilities of *P. polycephalum* have inspired the development of several algorithms in computer science addressing shortest-path problems [56, 81] and network design [57, 82–86]. In Sec. I.1.6.1, we discuss the seminal maze-solving experiment in detail and examine how it provided the basis for the shortest-path finding algorithm introduced in Sec. I.1.6.2. We then demonstrate in Secs. I.1.6.3 and I.1.6.4 how *P. polycephalum* can be harnessed to design efficient and robust transport networks, and extend the shortest-path finding algorithm to a multi-commodity flow model for network design. Further, in Sec. I.1.6.5, we discuss why adding noise to Physarum-inspired algorithms is a promising approach to enhance the resulting networks in terms of efficiency and robustness.

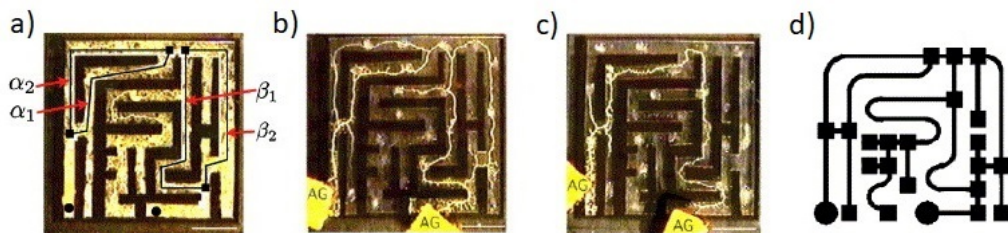


Figure 5: Adapted from Ref. [56]. Demonstration of the maze-solving capability of *Physarum polycephalum*. (a) Initial distribution of the slime mold (yellow) across the maze (black). (b) Retraction from dead ends and connection of the two food sources labeled AG. (c) Final state, where the slime mold converged to the shortest path between the two food sources. (d) Graph representation of the maze, with squares and circles indicating nodes and lines denoting edges.

I.1.6.1 Maze-solving capability of the slime mold

In the experiment by Tero et al. [56], the slime mold was initially distributed across the entire maze, as depicted in Fig. 5(a), where the walls of the maze are black and the organism is yellow. The entrance and the exit of the maze are represented by black circles. The figure also indicates two sets of alternative paths, α_1 and α_2 as well as β_1 and β_2 , between distinct points within the maze, illustrating the choices the organism has to make to optimize its transport network. Food sources, which are labeled AG, were then placed at the entrance and the exit of the maze. The response of the slime mold, shown in Fig. 5(b), involved retracting from dead ends and concentrating its mass around the food sources. After a few hours, the organism converged to the shortest path connecting the entrance and the exit of the maze, thereby identifying the most efficient transport route between the two locations, as depicted in Fig. 5(c).

At the core of this behavior is a phenomenon known as protoplasmic shuttle streaming, where cytoplasm, the fluid inside the cell, flows rhythmically back and forth within the organism’s network of tubes [87]. This oscillatory movement with a period of approximately two minutes [88] is driven by mechanical contractions of actin-myosin fibers that form the tube walls, occurring throughout the whole organism and increasing in amplitude at food sources [89]. This process allows the slime mold to continuously redistribute its mass. Tero et al. [56] identified a feedback loop that increases the diameter of a tube for large flows, and decreases it for small flows. Based on these observations, in the same study, they developed a mathematical model that reproduces the ability of the slime mold to find the shortest path between two food sources. In their model, they assume that the mechanical contractions are larger in amplitude at one food source, leading to an effective net flow of cytoplasm to the other food source. While this is a strong simplification—implying the injection and removal of cytoplasm at the food sources, which is biologically not plausible—it is reasonable for modeling purposes because the feedback mechanism only depends on the magnitude of the flow, not on its direction. Additionally, the model considers time-averaged flows, which is reasonable given that the oscillations driving the protoplasmic shuttle streaming have a much shorter period than the time scale on which the tubes evolve.

In the following section, we formalize the algorithm based on the framework of multi-

commodity flow problems introduced in Sec. I.1.5. However, we set $k = 1$ to capture the single-demand scenario of the maze, modeling the effective net flow between the two food sources.

I.1.6.2 Model for finding the shortest path through a maze

For this purpose, we represent the maze as a graph $G(V, E)$, where junctions, as well as the entrance and exit, are nodes, and the tubes connecting them are edges. In the slime mold's maze-solving process, junctions are decision points for choosing which path to follow. The graph abstraction is shown in Fig. 5(d), with squares indicating the entrance and exit nodes, circles representing all other nodes, and lines depicting the edges. The model by Tero et al. [56] applies the following modifications compared to the standard multi-commodity flow problems discussed in Sec. I.1.5:

1. Flow is determined by edge capacity and node potentials

The capacity constraint (0.8), which merely provides an upper limit to the flow through the edge (u, v) , is replaced by the relation

$$Q_{u,v} = \frac{D_{u,v}}{L_{u,v}}(p_u - p_v), \quad (0.12)$$

rendering the flow proportional to the edge capacity. Here, $L_{u,v}$ denotes the length of the edge and p_x the potential at node x with $x = u, v$. Due to the analogy between fluid flow and electrical current, Eq. (0.12) resembles Ohm's law of electrical circuits, with $D_{u,v}/L_{u,v}$ corresponding to the inverse of the resistance and $(p_u - p_v)$ representing the voltage. The potential at each node is induced by the flow conservation constraints (0.7), which remain unchanged. Together with Eq. (0.12), they form a linear system of equations for the node potentials for given edge capacities. The node potentials can then be used to determine the flows through all edges.

2. Dynamic edge capacity depending on the flow

The model assumes that the edge capacity $D_{u,v}$ is a dynamic variable whose evolution is governed by the set of coupled differential equations

$$\partial_t D_{u,v} = f(|Q_{u,v}|) - \gamma_{u,v} D_{u,v}. \quad (0.13)$$

Here, the first term describes the nonlinear feedback of the flow through the edge (u, v) on its capacity, determined by an activation function f that is strictly monotonously increasing and satisfies $f(0) = 0$. This process competes with dissipation at a rate $\gamma_{u,v}$, modeling the reported feedback mechanism where the edge capacity decreases when the flow through it diminishes and instead increases as the flow grows.

Figure 6 illustrates the numerical solution of the model given by Eqs. (0.7), (0.12) and (0.13) for the same setup as considered by Tero et al. in their experiment [56]. While Fig. 6(a) shows the initial setup, where all edges share the same capacity, subfigures (b–d) demonstrate the final edge configurations for different activation functions. Figure 6(b) corresponds to the case $f(x) = x$, in which the model converges to the shortest path between the source and the sink node, independent of the initial conditions. In particular, Bonifaci et al. have proven this mathematically by showing that the system minimizes the total energy dissipation in the network, which is achieved by routing the flow along the shortest path [81]. However, while the system

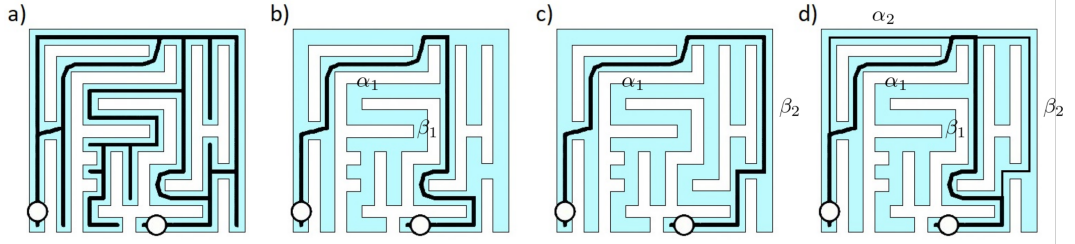


Figure 6: Adapted from Tero et al. [56]. Numerical simulation of their maze-solving experiment using the model given by Eqs. (0.7), (0.12) and (0.13). The maze is represented by the blue area, with edge capacities indicated by the width of the black lines. (a) Initially, all edges share the same capacity. (b–d) Final configurations for different activation functions: (b) $f(x) = x$, (c) $f(x) = |x|^\mu$ with $\mu > 1$, and (d) $f(x) = |x|^\mu$ with $\mu < 1$.

may still converge to the shortest path for certain initial conditions when considering nonlinear activation functions, this is not guaranteed in general [90].

Tero et al. [56] also considered the case $f(x) = x^\mu$, where the exponent $\mu > 0$ controls how the edge capacities react to the flow magnitude. For $\mu > 1$, shown in Fig. 6(c), the system further enhances the positive feedback of large flows on the edge capacities while small flows have less influence. Consequently, the algorithm prefers dominant paths that carry most of the flow. In contrast, for $0 < \mu < 1$, illustrated in Fig. 6(d), the effect of large flows is diminished and the impact of small flows is increased. This leads to a more uniform distribution of the flow across the network, meaning that loops can form.

This behavior is reminiscent of the phase transition identified by Banavar et al. [91], who studied transport networks that minimize a global cost function of the form

$$\tilde{J} = \sum_{u,v} |Q_{u,v}|^\Gamma. \quad (0.14)$$

At the critical point $\Gamma_{\text{crit}} = 1$, these networks switch from tree-like structures ($\Gamma < 1$) to structures with loops ($\Gamma > 1$). As discussed in Sec. I.1.6.2, this corresponds to a shift in key network characteristics such as robustness and efficiency. For an extensive characterization of this phase transition, see Ref. [92].

While minimizing a cost function of the form of Eq. (0.14) shares the power-law scaling of the flow with the algorithm outlined in Eqs. (0.7), (0.12) and (0.13), there are two key differences: First, the Physarum-inspired model dynamically adapts the edge capacities based on the flow feedback, generating networks that may or may not minimize \tilde{J} . In contrast, Banavar et al. [91] considered stationary networks achieving the lowest value of \tilde{J} across all possible topologies. Second, the occurrence of loops as a function of the exponent is reversed for both approaches. The model by Tero et al. [56] reinforces large flows for $\mu > 1$ as they further increase the edge capacities, resulting in a single dominant path. Conversely, the cost function \tilde{J} penalizes large flows for $\Gamma > 1$, making it cheaper to split the flow across several edges and thus form loops. Analogous arguments hold for the case $\mu < 1$ and $\Gamma < 1$.

Other common choices for $f(x)$, although not shown in Fig. 6, include saturating activation functions [56, 57, 93, 94]. Due to their asymptotic behavior for $x \rightarrow \infty$, these functions introduce a maximum edge capacity, representing a more realistic scenario.

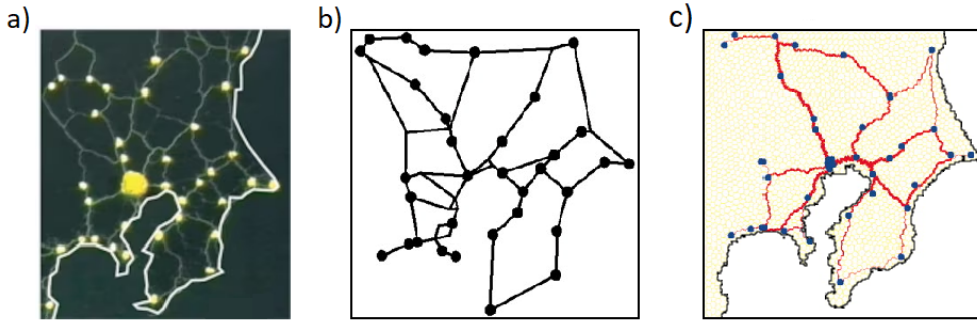


Figure 7: Adapted from Ref. [57]. Reprinted with permission from AAAS. (a) Transport system formed by the slime mold after food sources were geometrically arranged to reflect the relative locations of major cities around Tokyo. (b) Graph abstraction of the Tokyo railway system. (c) Network generated by the model described by Eqs. (0.7), (0.12) and (0.13).

I.1.6.3 Slime mold networks mimicking the Tokyo railway system

In a second, seminal experiment, Tero et al. demonstrated how *Physarum polycephalum* can be harnessed to solve network design problems [57]. As illustrated in Fig. 7(a), food sources were distributed to represent urban centers around Tokyo. While each major city was assigned a single food source, Tokyo itself was indicated by seven food sources to account for its importance in the region. This setup effectively simulated a civil engineering challenge, leading the slime mold to create a transport network that connects these urban centers. To consider geographical constraints, certain regions were selectively illuminated, penalizing growth into these areas as the slime mold minimizes its exposure to light. As a result, the slime mold formed a transport network that resembled the Tokyo railway system shown in Fig. 7(b), closely matching its cost, measured by the total length of the network, and its transport efficiency, quantified by the average length of the shortest paths connecting the different cities. In addition to that, the transport network generated by the slime mold exhibited comparable fault tolerance, as indicated by the probability of disconnecting the network by removing a random single edge.

In the same study [57], Tero et al. were able to reproduce these experimental observations by extending their shortest-path finding algorithm [56], given by Eqs. (0.7), (0.12) and (0.13), to consider time-varying source and sink pairs. For this purpose, they assumed a transport demand between any pair of cities. The restriction of the algorithm to a single demand and, thus, a single source and sink pair, was addressed by randomly selecting a different pair of cities as the source and sink nodes at each time step of the simulation. This random selection was performed with a uniform probability distribution over all pairs of cities. By iterating through multiple source and sink pairs, the model effectively accounts for multiple transport demands. Using this algorithm, the system self-organized into a network that closely mirrored the Tokyo railway system, demonstrating the model's ability to design robust and efficient transport networks based on simple local rules.

Figure 7(c) illustrates a network generated using this approach, where the yellow shaded area represents the available space, the blue dots denote the urban centers, and the red lines indicate the connections, with their widths corresponding to the capacity of the edges [57]. Remarkably, topological properties of the emerging network, such as robustness and efficiency, could be controlled by the flow I_0 that

is injected and removed at the source and sink nodes. For specific choices of I_0 , the generated network exhibited a higher fault-tolerance-to-cost ratio than both the slime mold network and the Tokyo railway system.

I.1.6.4 Model for multi-commodity transport

Despite the success of the extended shortest-path finding algorithm that we discussed in Sec. I.1.6.3 in reproducing the Tokyo railway system, this approach has inherent limitations. While it effectively approximates multi-commodity transport by iterating through different source and sink pairs, it does not account for all transport demands at once. This constraint demonstrates the need for a generalization of the single-commodity model defined by Eqs. (0.7), (0.12) and (0.13), as real-world systems typically require the simultaneous transport of various commodities. Such a generalization is introduced in Chapter 2 of this dissertation. Independently, Lonardi et al. [84] proposed an algorithm that, while similar to ours, differs in key aspects. To provide context for the subsequent discussion of our own model, we will briefly outline their algorithm in the following. Note that we adapt the notation so that it is consistent with the definitions in Secs. I.1.5 and I.1.6.2.

In the model by Lonardi et al. [84], the flow conservation constraints (0.7) of the single-commodity algorithm remain unchanged. While the single-commodity model assumes a single flow per edge as in Eq. (0.12), the generalized approach considers independent flows for each commodity i , which are induced by commodity-specific node potentials:

$$Q_{u,v}^i = \frac{D_{u,v}}{L_{u,v}}(p_u^i - p_v^i). \quad (0.15)$$

Here, the edge capacity $D_{u,v}$ is the same for all demands. Furthermore, the dynamics of the edge capacities, as defined in Eq. (0.13) of the single-commodity algorithm, are modified. Specifically, the first term is scaled with a power of the edge capacity:

$$\partial_t D_{u,v} = D_{u,v}^{2-\beta} f(Q_{u,v}) - \gamma_{u,v} D_{u,v}, \quad (0.16)$$

where $\beta > 0$ and $f(Q_{u,v}) = \|Q_{u,v}\|_2^2$ is the square of the two-norm of the vector $Q_{u,v} = (Q_{u,v}^1, \dots, Q_{u,v}^k)$ containing all commodity flows through the edge (u, v) .

Lonardi et al. demonstrated that their model is equivalent to an optimization problem [84]. Specifically, the stationary solutions of Eqs. (0.16) minimize the total transportation cost

$$J = \sum_{(u,v)} L_{u,v} \|Q_{u,v}\|_2^\Gamma, \quad (0.17)$$

with $\Gamma = 2(2 - \beta)/(3 - \beta)$. Remarkably, in the single-commodity case, Eq. (0.17) reduces to the cost function (0.14) considered by Banavar et al. [91]. As discussed in Sec. I.1.6.3, they observed a phase transition between tree-like topologies ($\Gamma < 1$) and structures containing loops ($\Gamma > 1$). In the multi-commodity generalization provided by Lonardi et al. [84], the same phase transition persists, but loops can emerge even in regimes where tree-like structures are optimal for single-commodity transport. This phenomenon arises from the indirect interaction among the commodity-specific flows, which is a consequence of sharing the same edge capacities. In contrast to this algorithm, our generalized model aligns more closely with the original single-commodity algorithm by Tero et al. [56], particularly with respect to modeling the nonlinear flow feedback. For further details, we refer to Chapter 2.

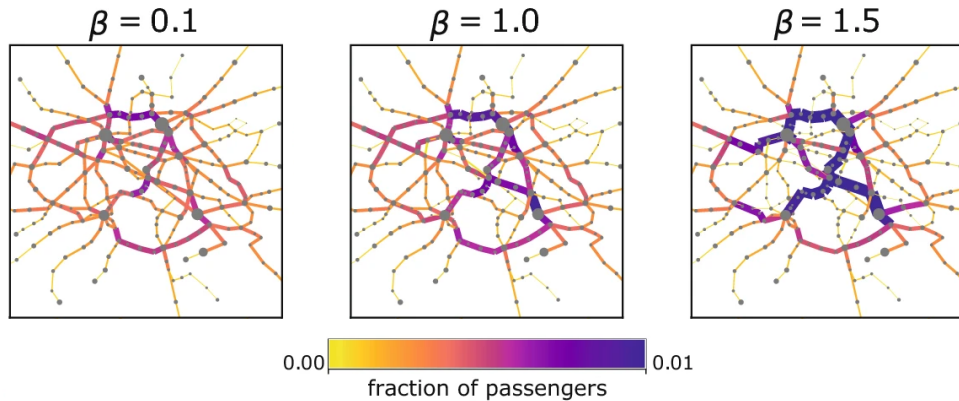


Figure 8: Adapted from Ref. [85]. Simulation of passenger flows through the Paris metro using the multi-commodity algorithm outlined in Eqs. (0.7), (0.15) and (0.16), for $\beta = 0.1$ ($\Gamma \approx 1.31$), $\beta = 1.0$ ($\Gamma = 1$) and $\beta = 1.5$ ($\Gamma \approx 0.66$). The width and color of each line represent the total passenger flow through the corresponding edge. Licensed under CC BY 4.0, <https://creativecommons.org/licenses/by/4.0/>.

To analyze their generalized model in a real-world context, Lonardi et al. applied it to optimizing passenger flows through the Paris metro [85]. By modeling the passenger demands between different stations as multi-commodity flows, they demonstrated that their algorithm can yield flow distributions allowing for resilient and efficient transport. Figure 8 shows the resulting passenger flows for various values of the exponent β . Here, the fraction of the total passenger flow carried by each edge is indicated by the width and color of the line representing the edge. As β increases, which corresponds to a decreasing Γ , loops become less prominent. However, even at the maximum value of $\beta = 1.5$ ($\Gamma \approx 0.66$), loops are still present. This is consistent with the earlier observation that loops can persist in multi-commodity settings, even in regimes where trees are optimal in the single-commodity case, i.e., for $\Gamma < 1$. Beyond civil engineering tasks such as determining optimal passenger flows, the model defined by Eqs. (0.7), (0.15) and (0.16) has been applied to other contexts, including multi-layer transport problems [95] and image processing problems [96].

1.1.6.5 Incorporating stochastic fluctuations

The Physarum-inspired models for network design that we discussed in Secs. I.1.6.2 and I.1.6.4 provide a deterministic approach for transport optimization both in single- and multi-commodity scenarios. However, as highlighted in Sec. I.1.1, real-world systems are naturally exposed to stochastic fluctuations. While Lonardi et al. [84] did not consider the effect of noise on their multi-commodity model, prior research has explored its impact on simpler setups. For instance, Meyer et al. [94] applied the single-commodity model defined by Eqs. (0.7), (0.12) and (0.13) to a system of two paths connecting the source and the sink node. These paths have equal length, which means that the algorithm for shortest-path finding would assign the same edge capacity to both of them. To analyze adaptability to dynamically changing environmental conditions, they introduced time-dependent and path-specific dissipation rates, breaking the symmetry of the problem. The response of the system was measured by the path preference c , where $c = -1$ indicates that the path with higher average dissipation was selected, and $c = 1$ indicates that the path with smaller average dissipation was chosen.

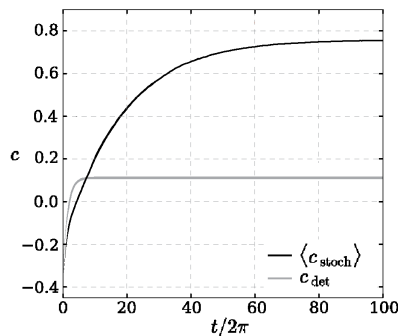


Figure 9: Adapted from Ref. [94]. The shortest-path finding algorithm given by Eqs. (0.7), (0.12) and (0.13) is applied to a system of two equally long paths but differing time-dependent dissipation rates. The quantity c measures the path preference and is displayed as a function of time. Shown are the deterministic solution c_{det} and the ensemble average $\langle c_{\text{stoch}} \rangle$ in the presence of stochastic fluctuations. Licensed under CC BY 4.0, <https://creativecommons.org/licenses/by/4.0/>.

Figure 9 shows the time evolution of c , with the deterministic value being denoted as c_{det} , and the ensemble average in the stochastic case as $\langle c_{\text{stoch}} \rangle$. Remarkably, for the chosen initial conditions, the model was only able to reliably select the path with the smaller average dissipation when stochastic forces were added to the dynamics of the edge capacities given by Eq. (0.13). While this study conceptually demonstrates how stochastic fluctuations can enhance the adaptability of a system, it remains limited to a fixed frequency of the periodically changing dissipation and noise amplitude as well as a highly simplified setup.

Motivated by these insights, in this dissertation, we incorporate noise into Physarum-inspired algorithms to systematically analyze how it might be harnessed for optimizing real-world transport systems. A detailed outline of the dissertation is given in the following section.

I.2 Outline

This cumulative dissertation comprises four publications, each of which is presented as an independent chapter. They address several open questions related to Physarum-inspired models for network design and the impact of noise, focusing on adaptability enhancement, model generalization, network optimization, and the interplay of nonlinear dynamics and stochastic forces. The following overview provides a description of the central research question of each chapter as well as their interrelation.

Chapter 1 (Ref. [93]): In this chapter, we apply the single-commodity model defined by Eqs. (0.7), (0.12) and (0.13) to a system exposed to stochastic fluctuations. We analyze how varying key parameters such as the noise amplitude influences the system's response to periodically changing dissipation, asking the question: *Under which conditions can stochastic fluctuations enhance adaptability to dynamic environments?*

Chapter 2 (Ref. [82]): This chapter addresses the limitations of applying the single-commodity algorithm considered in Chapter 1 to multi-commodity transport prob-

lems. We seek to answer the question: *Can Physarum-inspired dynamics be extended to design multi-commodity transport networks?*

Chapter 3 (Ref. [83]): In this chapter, we investigate whether the network topologies generated by the generalized model introduced in Chapter 2 can be enhanced by adding stochastic fluctuations. We focus on a simple setup with two transport demands and pose the question: *Can noise optimize the robustness, efficiency, and cost of transport networks?* In addition to that, we study how the stochastic fluctuations influence the convergence speed of the algorithm.

Chapter 4 (Ref. [86]): This chapter is divided into two parts. In the first part, based on the analysis in Chapter 3, we extend our approach to designing transport networks that connect major cities around Tokyo. We consider various classes of activation functions and ask the question: *How does the interplay of nonlinear dynamics and stochastic forces influence network topologies?* In the second part, which consists of unpublished results, we analyze the effect of noise on the number of loops, their lengths, and the fraction of edges that are part of at least one loop.

Outlook: We investigate whether the results presented in this dissertation can benefit a wider range of applications, especially other well-known optimization algorithms. In particular, we consider the gradient descent method for finding the global minimum of a cost function as a case study. Using various approaches, we add nonlinear dynamics and stochastic forces to this algorithm and analyze whether these modifications can enhance its convergence.

I.3 Contributions of the author

In this section, we detail the aspects of the publications contained in this dissertation to which the author contributed. Note that the co-authors of the publications have significantly contributed to the respective aspects as well.

Chapter 1:

Interplay of periodic dynamics and noise: insights from a simple adaptive system

Published in: Phys. Rev. E 104, 054215 (2021)

Authors: Frederic Folz, Kurt Mehlhorn, Giovanna Morigi

Contributions of the author:

- Developed the model together with Kurt Mehlhorn (KM) and Giovanna Morigi (GM).
- Performed the numerical simulations and analytical studies.
- Analyzed the data together with KM and GM.
- Wrote the article together with KM and GM.

Chapter 2:

Physarum-inspired multi-commodity flow dynamics

Published in: Theoretical Computer Science 920 (2022) 1–20

Authors: Vincenzo Bonifaci, Enrico Facca, Frederic Folz, Andreas Karrenbauer,

Pavel Kolev, Kurt Mehlhorn, Giovanna Morigi, Golnoosh Shahkarami, Quentin Vermande

Contributions of the author:

- Analyzed data together with Kurt Mehlhorn and Giovanna Morigi.
- Created a figure for the publication.

Chapter 3:

Noise-induced network topologies

Published in: Phys. Rev. Lett. 130, 267401 (2023)

Authors: Frederic Folz, Kurt Mehlhorn, Giovanna Morigi

Contributions of the author:

- Developed the model together with Kurt Mehlhorn (KM) and Giovanna Morigi (GM).
- Performed the numerical simulations and analytical studies.
- Analyzed the data together with KM and GM.
- Wrote the article together with KM and GM.

Chapter 4:

Part 1: Self-organized transport in noisy dynamic networks

Published in: Phys. Rev. E 110, 044310 (2024)

Authors: Frederic Folz, Kurt Mehlhorn, Giovanna Morigi

Contributions of the author:

- Developed the model together with Kurt Mehlhorn (KM) and Giovanna Morigi (GM).
- Performed the numerical simulations and analytical studies.
- Analyzed the data together with KM and GM.
- Wrote the draft of the article, which was finalized together with KM and GM.

Part 2: Unpublished material

Contributions of the author:

- Developed the model together with Kurt Mehlhorn (KM) and Giovanna Morigi (GM).
- Performed the numerical simulations and analytical studies.
- Analyzed the data together with KM and GM.

CHAPTER I

INTERPLAY OF PERIODIC DYNAMICS AND NOISE: INSIGHTS FROM A SIMPLE ADAPTIVE SYSTEM

Interplay of periodic dynamics and noise: insights from a simple adaptive system

Phys. Rev. E 104, 054215 – Published 29 November 2021

© 2021 American Physical Society

DOI: 10.1103/PhysRevE.104.054215

Authors: Frederic Folz¹, Kurt Mehlhorn², Giovanna Morigi¹

¹ *Theoretische Physik, Universität des Saarlandes, 66123 Saarbrücken, Germany*

² *Algorithms and Complexity Group, Max-Planck-Institut für Informatik, Saarland Informatics Campus, 66123 Saarbrücken, Germany*

Abstract:

We study the dynamics of a simple adaptive system in the presence of noise and periodic damping. The system is composed by two paths connecting a source and a sink, and the dynamics is governed by equations that usually describe food search of the paradigmatic *Physarum polycephalum*. In this work we assume that the two paths undergo damping whose relative strength is periodically modulated in time, and we analyze the dynamics in the presence of stochastic forces simulating Gaussian noise. We identify different responses depending on the modulation frequency and on the noise amplitude. At frequencies smaller than the mean dissipation rate, the system tends to switch to the path which minimizes dissipation. Synchronous switching occurs at an optimal noise amplitude which depends on the modulation frequency. This behavior disappears at larger frequencies, where the dynamics can be described by the time-averaged equations. Here we find metastable patterns that exhibit the features of noise-induced resonances.

Interplay of periodic dynamics and noise: Insights from a simple adaptive system

Frederic Folz¹, Kurt Mehlhorn² and Giovanna Morigi¹

¹Theoretische Physik, Universität des Saarlandes, 66123 Saarbrücken, Germany

²Algorithms and Complexity Group, Max-Planck-Institut für Informatik, Saarland Informatics Campus, 66123 Saarbrücken, Germany



(Received 23 August 2021; accepted 10 November 2021; published 29 November 2021)

We study the dynamics of a simple adaptive system in the presence of noise and periodic damping. The system is composed by two paths connecting a source and a sink, and the dynamics is governed by equations that usually describe food search of the paradigmatic *Physarum polycephalum*. In this work we assume that the two paths undergo damping whose relative strength is periodically modulated in time, and we analyze the dynamics in the presence of stochastic forces simulating Gaussian noise. We identify different responses depending on the modulation frequency and on the noise amplitude. At frequencies smaller than the mean dissipation rate, the system tends to switch to the path which minimizes dissipation. Synchronous switching occurs at an optimal noise amplitude which depends on the modulation frequency. This behavior disappears at larger frequencies, where the dynamics can be described by the time-averaged equations. Here we find metastable patterns that exhibit the features of noise-induced resonances.

DOI: 10.1103/PhysRevE.104.054215

I. INTRODUCTION

Patterns are ubiquitous in nature; metastable spatiotemporal structures are observed from the microscopic to the astrophysical scale [1–4]. The systematic characterization of their onset and stability is a formidable challenge of theoretical physics. Theoretical descriptions are often based on coupled nonlinear equations for macroscopic variables, whose fixed points often capture essential features of the metastable dynamics; see, for instance, [1,5–8].

A prominent example are the equations modeling the dynamics of biological systems, such as food search of *Physarum polycephalum*, a representative of the so-called true slime molds [9]. *Physarum polycephalum* is a single-celled organism that, despite its lack of any form of nervous system, is able to solve complex tasks like finding the shortest path through a maze [10–12] and creating efficient and fault-tolerant networks [13,14]. These dynamics are qualitatively reproduced by the noise-free coupled nonlinear equations of motion [9], which are a reference model system for optimization algorithms and deep learning [15].

Most theoretical descriptions of *Physarum* are noise-free and do not include the effect of a thermal bath in which *Physarum* is naturally immersed. On the other hand, tasks such as finding the optimal path in a maze are solved in contact with the external environment [10,13], which shows that *Physarum* dynamics is robust and probably even optimized for a certain level of noise, as typical of adaptive systems [16,17].

Motivated by this question, in this work we consider the coupled nonlinear equations describing an adaptive system that can choose between two paths in the presence of Gaussian noise and that is inspired by the dynamics of *P. polycephalum*. We analyze the dynamics when the relative dissipation rate between the two paths is modulated in time and as a function

of the modulation frequency and of the noise amplitude. We benchmark our results with the work of Ref. [18], who studied this model for a fixed value of frequency and noise amplitude.

This paper is organized as follows. In Sec. II A we introduce the physical model, and in Sec. III we discuss the response as a function of the frequency and noise amplitude and identify the regime where stochastic resonance characterizes the dynamics. In Sec. IV we then turn to the regime outside the stochastic resonance condition when the frequency is sufficiently large and characterize the system dynamics as a function of the noise amplitude. The conclusions are drawn in Sec. V. The Appendixes provide details on the model and on the calculations in Secs. III and IV.

II. TWO PATHS WITH MODULATED DISSIPATION

A. The model

The model we consider is a simplified network, where a source and a sink are connected through two paths of equal and constant length L as illustrated in Fig. 1(a) and whose dynamics is inspired by *P. polycephalum*. Here the capability to connect two food sources is modeled with the flow of gel inside the cell body along the network edges and is quantified by the conductivity D_i of path $i = 1, 2$ [9,19]. This variable increases monotonically with the gel's flow and vanishes when the flow vanishes according to the deterministic equation ($i = 1, 2$)

$$[\partial_t D_i]_0 = f(\tilde{D}_i) - \gamma_i(t) D_i, \quad (1)$$

where $f(x)$ is the nonlinear force with argument $\tilde{D}_i = D_i/(D_1 + D_2)$:

$$f(x) = \Gamma(1 + \epsilon)x^2/(\epsilon + x^2), \quad (2)$$

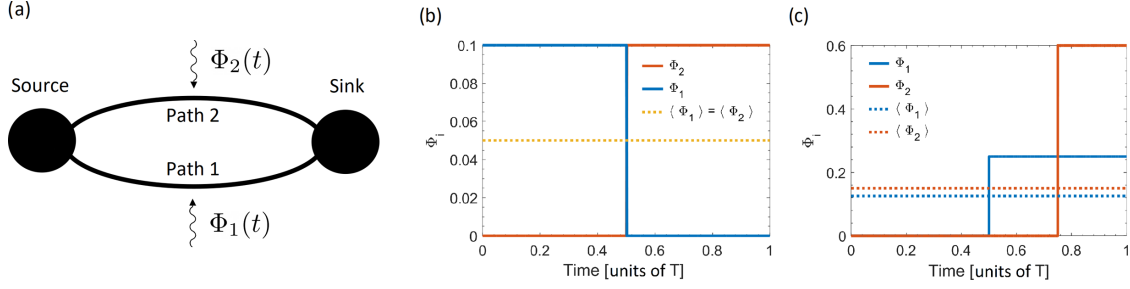


FIG. 1. (a) Illustration of a system consisting of two different paths connecting a source and a sink. The paths are exposed to a periodically modulated dissipation. (b, c) The periodic function $\Phi_i(t)$, determining the temporal behavior of the damping rate of path i , over one period T and for the two examples discussed in this work [symmetric (b) and biased (c) case]. The time is in units of $1/\gamma$; the dotted lines are the time averages over one period.

with $\epsilon = 0.2$ and γ_i is the damping rate of path i [18]. The dissipation is here periodically modulated in time with period T ,

$$\gamma_i(t) = \gamma + \Phi_i(t), \quad (3)$$

with $\Phi_i(t) = \Phi_i(t + T)$ a periodic and positive function; see Fig. 1. Let $\omega = 2\pi/T$ denote the angular frequency. Furthermore, we set $\Gamma = \gamma$ in the following. From now on, the time will be given in units of γ^{-1} unless otherwise stated.

In this work we characterize the system response to the dynamically changing environment in the presence of the stochastic force $\xi_i(t)$:

$$\partial_t D_i = [\partial_t D_i]_0 + \tilde{\alpha} \xi_i(t). \quad (4)$$

The stochastic force is scaled by the positive parameter $\tilde{\alpha} = \alpha\gamma$ and describes white noise. It has vanishing expectation value, $\langle \xi_i(t) \rangle = 0$, and correlations

$$\langle \xi_i(t) \xi_j(t') \rangle = \kappa \delta_{i,j} \delta(t - t'), \quad (5)$$

where $\kappa = \gamma^{-1}$ and $\langle \cdot \rangle$ is the average taken over a sufficiently large number of independent realizations of the stochastic process $\xi_i(t)$ [20,21]. Note that, since this force can take negative value, we need to regularize the behavior of Eq. (4) for small values of D_i so as to keep $D_i \geq 0$. This in turn reflects the physical constraint that D_i represents a conductivity. For this purpose, when $D_i = 0$ and the right-hand side of Eq. (4) becomes negative, we set it equal to 0. In what follows we numerically simulate Eq. (4) using the Euler-Maruyama scheme [22] with a step size $\Delta t = 0.05\gamma^{-1}$, unless otherwise specified.

As in Ref. [18] we quantify the system's response by means of the quantity

$$c(t) = \frac{D_1(t) - D_2(t)}{D_1(t) + D_2(t)}, \quad (6)$$

which we denote by risk function. The risk function varies in the interval $[-1, 1]$. The extremal values $c = +1$ and $c = -1$ correspond to the system being in the path $i = 1$ and $i = 2$, respectively. We note that the function minimizing the risk takes the form

$$c_r(t) = 2 \left\{ \theta[\Phi_2(t) - \Phi_1(t)] - \frac{1}{2} \right\},$$

with $\theta(x)$ Heaviside's function. The corresponding dynamics follows the path whose instantaneous dissipation is minimal.

B. The biological system

Equation (1) has been proposed in Ref. [9] for simulating the dynamics of *P. polycephalum* for a constant dissipation γ . This model was extended in Ref. [18] for the purpose of analyzing the effect of noise on the adaptivity of *P. polycephalum* to dynamically varying environmental conditions. The dynamically varying environment was there modeled by a periodically modulated dissipation. The latter simulated the effect of light, which inhibits the growth of *P. polycephalum*.

In the present work we start from that analysis and extend it by investigating the dynamics as a function of the noise strength and of the modulation frequency. This allows us to identify whether and when there is an optimal noise strength for which the system optimally adapts to the external environment.

III. SYMMETRIC CONFIGURATION

In the following we determine the dynamics as a function of α , the noise strength, and ω , the frequency at which dissipation is modulated. We assume that the two functions $\Phi_i(t)$ are step functions shifted by half period with respect to one another: $\Phi_1(t) = \Phi_2(t + T/2)$. Over one period we choose $\Phi_1(t) = \gamma_0 \theta(t) \theta(T/2 - t)$, with $\gamma_0 = 0.1\gamma$. Figure 2 displays the evolution of the conductivity D_1 for different, increasing values of α . Among the three examples displayed, the flow seems to best adapt to the periodic changes of the external parameters for $\alpha \sim 0.1$ (we emphasize that for $\alpha = 0$ the system does not switch paths).

We quantify the capability of the system to adapt to the changes of dissipation over the evolution time $[0, t_{\text{end}}]$ by means of the normalized correlation function

$$g(\tau) = \frac{2}{t_{\text{end}}} \int_0^{t_{\text{end}}} \left\{ \theta[c(t)] - \frac{1}{2} \right\} c_r(t - \tau) dt, \quad (7)$$

which quantifies the overlap between the signal $c(t)$ and the function $c_r(t)$, as a function of the delay $\tau > 0$. Perfect correlation (anticorrelation) between dynamics and dissipation

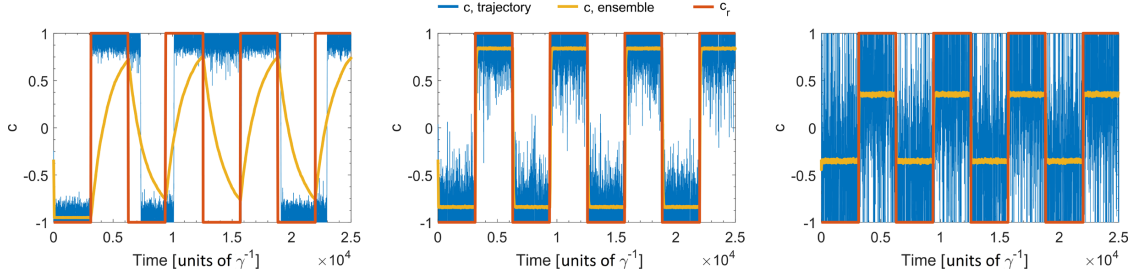


FIG. 2. Time evolution of the risk function c for $\omega = 10^{-3}\gamma$ and (a) $\alpha = 0.051$, (b) $\alpha = 0.099$, and (c) $\alpha = 0.222$. The blue line corresponds to one trajectory, the yellow line is the average over 5000 trajectories, and the red line displays $c_r(t)$. The initial conditions are $D_1^0 = 0.5, D_2^0 = 1$ corresponding to $c^0 = -1/3$. In Appendix A we report a zoom of panel (a).

minimizes (maximizes) the risk and corresponds to $g = 1$ and $\tau \rightarrow 0$ ($\tau = T/2$).

Figure 3(a) displays the color plot of $g(0)$ as a function of the noise strength α and of the modulation frequency ω . For the full range of angular frequencies ω shown in the figure, it holds $g(0) \approx \max_\tau [g(\tau)]$. We observe a region for which $g(0) \approx 1$, indicating synchronous behavior at an optimal noise

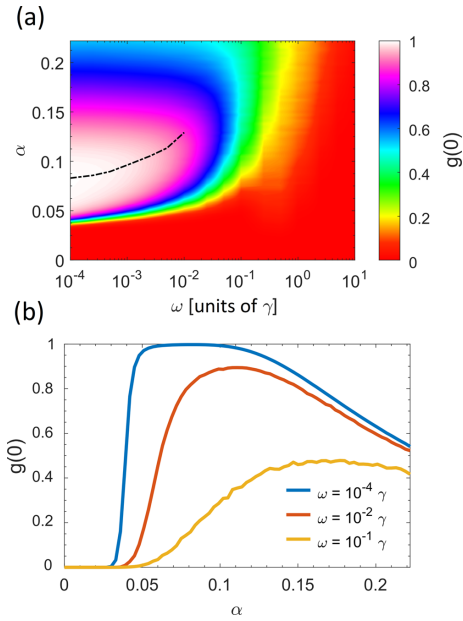


FIG. 3. (a) Color plot of the measure $g(0)$, calculated using Eq. (7), as a function of the noise strength α and of the modulation frequency ω . The dotted line represents the noise strengths at which we expect stochastic resonance to occur according to the matching condition of Eq. (8). For the displayed range of angular frequencies ω it holds $g(0) \approx \max_\tau [g(\tau)]$. (b) Measure $g(0)$ as a function of α for three values of the frequency ω ; see legend. We used the initial conditions $D_1^0 = 0.5, D_2^0 = 1$ corresponding to $c^0 = -1/3$. The data were generated by evaluating an ensemble of 5000 trajectories with simulation time of $t_{\text{end}} = 8\pi/\omega$ and step size $\Delta t = 10^{-4}/\omega$.

strength within a frequency range $\omega \lesssim 10^{-2}\gamma$. In Fig. 3(b) we display $g(\tau)$ for three values of ω and as a function of α : The dynamics exhibits the features of stochastic resonance [23]; the optimal noise strength where $g(\tau)$ is maximum can be found from the relation

$$t_s = \frac{\pi}{\omega}, \quad (8)$$

where t_s is the average switching time between the two stable fixed points $c = -1$ and $c = 1$. We estimate the switching time using a one-dimensional model; the details are reported in Appendix B. The resulting curve is the dotted line of Fig. 3(a) and reproduces the position of the resonance in the α - ω plane, which we find numerically.

The resonance behavior as a function of α broadens as ω increases. For $\omega \gtrsim \gamma$ there seems to be no correlation between $c(t)$ and the temporal modulation of the dissipation. In the next section we analyze the effect of noise in this regime.

IV. SECULAR REGIME

In this section we analyze the dynamics as a function of the noise strength at large frequencies, such that $\omega/\gamma \gg 1$. In this regime we expect that the effect of the time-dependent dissipation on the dynamics can be replaced by its average. We choose an asymmetric modulation of dissipation between the two paths and fix $\omega = 10\gamma$. In order to compare with Ref. [18] we define

$$\Phi_1(t) = \gamma_{01}\theta(t - T/2)\theta(T - t), \quad (9)$$

$$\Phi_2(t) = \gamma_{02}\theta(t - 3T/4)\theta(T - t) \quad (10)$$

for $t \in [0, T]$, with $\gamma_{01} = \gamma/4$ and $\gamma_{02} = 3\gamma/5$. According to this choice we expect a bias towards path 1, since it is characterized by the minimal (time-averaged) dissipation. We note that in Ref. [18] the authors considered the specific noise strength $\alpha = 0.05$ and simulated the dynamics till time $t_{\text{end}} = 200\pi\gamma^{-1}$. In the following we study the dynamics as a function of α . Moreover, we analyze the convergence of the simulated trajectories by taking different times and by comparing the results with the stationary state, which we analytically determine.

Figure 4 displays the evolution of the conductivities D_i by numerically integrating Eq. (4) assuming that initially

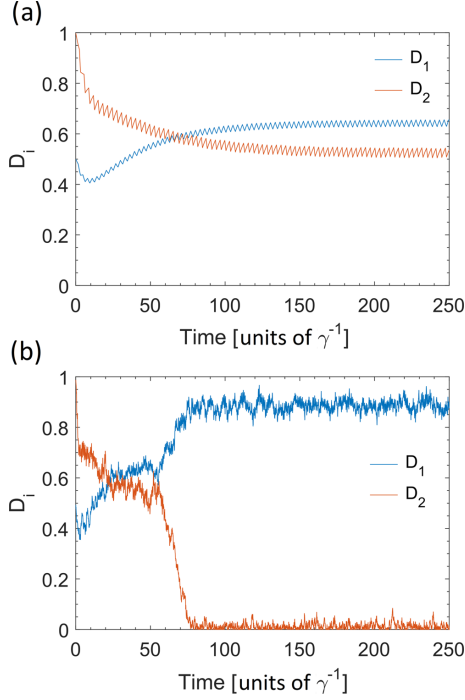


FIG. 4. Evolution of the conductivities D_i as a function of time (in units of $1/\gamma$) for $\omega = 10\gamma$ and (a) $\alpha = 0$, (b) $\alpha = 0.05$. In the deterministic case, $\alpha = 0$, the system converges to the fixed point of the secular equation; see the blue solid circle of Fig. 5(a). In the presence of noise this fixed point is metastable: the system quickly converges to the stable fixed point at which the system favors the path exposed to minimal averaged dissipation. The data shown represent a single trajectory. The initial conditions are $D_1^0 = 0.5$ and $D_2^0 = 1$.

$D_2 > D_1$. In the deterministic case, for $\alpha = 0$, the curves exhibit fast, small-amplitude oscillations at a frequency close to the modulation frequency ω , and their time average over a period varies over a significantly longer timescale. After a transient, the conductivity D_1 reaches a value larger than D_2 . The corresponding dynamics in the presence of white noise is shown in Fig. 4(b): we observe a timescale separation as in the deterministic case, while the frequency of the fast oscillations become chaotic. The slow dynamics reaches a metastable state (corresponding to the asymptotic state of the noise-free dynamics) where it is trapped for a relatively short time. It then quickly reaches the steady state, with the value of the conductivity D_1 approaching unity, while the flow along the second path is almost suppressed.

A. Fixed points of the secular dynamics

In order to perform an analytical study, we consider the secular dynamics, where we average Eq. (4) over a period T of the oscillations and replace $\gamma_i(t)$ with the time-averaged dissipation coefficients:

$$\gamma_i^{\text{eff}} = \gamma + \langle \Phi_i \rangle_T,$$

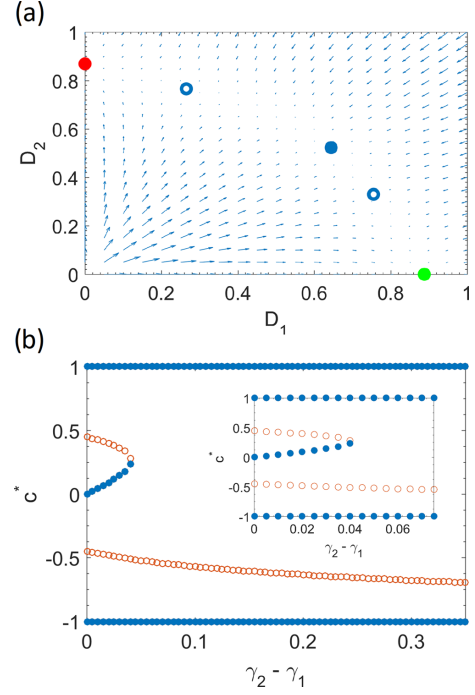


FIG. 5. (a) Fixed points (D_1^* , D_2^*) of the secular equations for $\langle \Phi_1 \rangle_T = 0.125$ and $\langle \Phi_2 \rangle_T = 0.15$. The arrows indicate the flow. Stable (unstable) fixed points are represented by solid (hollow) circles. The green (red) stable fixed point correspond to the system choosing the path with minimal (maximal) average dissipation. (b) Fixed points $c^* = (D_1^* - D_2^*)/(D_1^* + D_2^*)$ as a function of $\gamma_2^{\text{eff}} - \gamma_1^{\text{eff}}$ with fixed $\gamma_1^{\text{eff}} = 1.125$. Full (hollow) circles indicate stable (unstable) fixed points. The inset zooms into the parameter region of the intermediate, metastable fixed point.

with

$$\langle A \rangle_T \equiv \frac{1}{T} \int_t^{t+T} d\tau A(\tau).$$

According to our parameter choice $\gamma_1^{\text{eff}} < \gamma_2^{\text{eff}}$ and in the regime of validity of this secular approximation path 1 is favored. We then set $\alpha = 0$ and study the fixed points of the dynamics D_i^* , fulfilling $\partial_t D_i^* = 0$. We report the details in Appendix C.

Figure 5(a) displays the fixed points (D_1^* , D_2^*): stable (unstable) solutions are represented by solid (hollow) circles. The vector fields illustrate the flow. The path with average minimal dissipation corresponds to the green circle. The blue circle is the metastable path to which the deterministic dynamics of Fig. 4(a) converges for the given initial condition. This solution still characterizes the transient dynamics in the presence of noise, Fig. 4(b). On longer timescales, however, the stochastic dynamics brings the system to the path minimizing dissipation. Figure 5(b) shows the variable $c^* = (D_1^* - D_2^*)/(D_1^* + D_2^*)$ as a function of the average dissipation γ_2^{eff} , which is varied starting from the symmetric case $\gamma_2^{\text{eff}} = \gamma_1^{\text{eff}}$ while keeping γ_1^{eff} fixed. The blue solid (red hollow)

symbols indicate the stable (unstable) solutions. The red hollow symbols are the borders of the basins of attraction of the nearby stable solution for the deterministic dynamics. The two extremal solutions—where the system settles on one of the two paths—are independent of γ_2^{eff} . Instead, the solution characterized by finite conductivities along both paths exists solely below a threshold value γ_2^{eff} . As γ_2^{eff} increases from γ_1^{eff} towards the threshold, this fixed point moves towards positive values, indicating that the symmetry of the solutions is broken. Correspondingly, its basin of attraction becomes asymmetric by increasing γ_2^{eff} and biased towards the region $c > 0$. At the threshold, the basin of attraction of $c^* = 1$ undergoes a discontinuous jump and extends to the neighborhood of the other fixed point $c^* = -1$.

Despite the fact that this analysis has been performed using a secular approximation, we note that the fixed points of Fig. 5(a) allow us to understand the steady state of Fig. 4 as well as the metastable dynamics in Fig. 4(b). According to this picture, one would expect that by increasing α the rate of convergence to the stable state shall accordingly increase. In the following section we will show that this is only partly true.

B. Convergence to the stationary state

The analysis of the fixed point of the deterministic dynamics provides some insight into the behavior observed in Fig. 4. Here metastable fixed points correspond to metastable configurations whose lifetime is limited by noise. Fluctuations, in general, allow the system to explore a large configuration space.

In order to characterize the convergence to the steady state as a function of the noise strength α we first determine the variable c for sufficiently large integration times t_{end} . We single out the slowly varying value by averaging out the fast oscillations over a period,

$$\bar{c}(t_{\text{end}}) = \frac{1}{T} \int_{t_{\text{end}}-T}^{t_{\text{end}}} c(t') dt', \quad (11)$$

where $c(t)$ in the integrand is the ensemble average over the individual trajectories. Figure 6 displays $\bar{c}(t_{\text{end}})$ as a function of the noise strength α for different integration times t_{end} . The choice $t_{\text{end}} = 200\pi/\gamma$ corresponds to the same integration time of Ref. [18]. Comparison with longer simulation times shows that at this time and $\alpha \lesssim 0.07$ the dynamics has not yet converged to the stationary state. For $\alpha \lesssim 0.02$, moreover, the system is still trapped in the metastable configuration at $c^* \sim 0$ even for the longest integration time here considered.

The behavior in the interval $0.02 \lesssim \alpha \lesssim 0.07$ is remarkable, as it exhibits local minima which have the form of resonances. We note that the lifetime of the metastable configurations at $\alpha \sim 0.04$ exceeds $t_{\text{end}} \sim 10^5/\gamma$.

Figure 7 shows the time evolution of c about these special values of α . Figure 7(a) displays the time evolution for values of $\alpha \lesssim 0.02$. Here the system is trapped in the metastable configuration corresponding to the fixed point at $c^* \gtrsim 0$: the residence time visibly decreases as α increases. This trend is still visible at larger values of α in Fig. 7(b). At longer timescales, however, the curves in Fig. 7(b) show a metastable regime close to path 1 where the system remains trapped and whose lifetime is maximal for $\alpha \sim 0.04$. This metastable

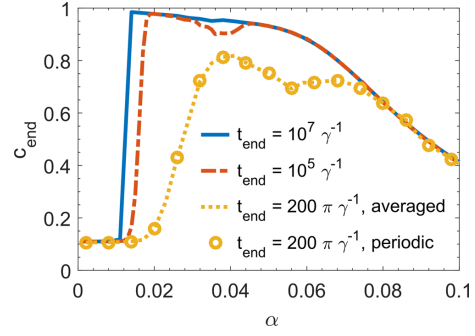


FIG. 6. Average value \bar{c} at the simulation time t_{end} as a function of the noise strength α for $t_{\text{end}} = 200\pi\gamma^{-1}$, $10^5\gamma^{-1}$, $10^7\gamma^{-1}$. The data shown were generated by averaging over a total of 5000 trajectories. The initial conditions are $D_1^0 = 0.5$, $D_2^0 = 1$ corresponding to $c^0 = -1/3$. For all values of t_{end} we evolve the secular equation. We also show the result of the full, periodic dynamics (circles) for $t_{\text{end}} = 200\pi\gamma^{-1}$. Here, to generate comparable data, we use the same set of random numbers for the trajectories.

configuration is not captured by the fixed point analysis and seems to crucially depend on the noise strength. It has thus the form of a noise-induced resonance.

We study these resonances by inspecting the variation Δc of the curve $c(t)$ at the extremal of the interval of time $\mathcal{I}(\delta) = [\delta \cdot t_{\text{end}}, t_{\text{end}}]$ with $\delta \in (0, 1)$:

$$\Delta c = \left(c_{\text{end}} - \frac{1}{T} \int_{\delta \cdot t_{\text{end}}-T}^{\delta \cdot t_{\text{end}}} c(t') dt' \right) c_{\text{end}}^{-1}, \quad (12)$$

where c_{end} is the stationary value for $t \rightarrow \infty$. When the system dynamics does not change over the interval $\mathcal{I}(\delta)$, then the variation $\Delta c = 0$. Figure 7(c) shows Δc as a function of α for $\delta = 0.6$ and different integration times t_{end} . Metastable configurations appear as resonances. We observe several resonances for relatively short integration times, while for increasing t_{end} the number of metastable states decreases. The small resonance at $\alpha \sim 0.04$ signals the metastable configuration of Fig. 7(b). The largest resonance at $\alpha \sim 0.02$ separates the regime where the system is still trapped in the metastable fixed point from the regime where the system has already escaped this region.

C. Steady state

We complete this study by discussing the steady state of the dynamics as a function of α . We apply the approach implemented in Ref. [18] and extend it to determine the dependence on α . We review here the basic steps. The approach consists in determining the time-averaged dynamics of the single variable \bar{c} , assuming that it undergoes a time-continuous Markov process in the presence of a drift $\mu(\bar{c})$ and an Itô diffusion with amplitude $\sigma(\bar{c}, t)$, which are determined by means of an equation-free analysis; see Appendix D. We verify the validity of this approximation by comparing the predictions of the full dynamics at $t_{\text{end}} = 10^7/\gamma$ with the one of the stochastic equation for the single variable \bar{c} ; see Fig. 8(a). For the one-dimensional case we take noise strengths $\alpha > 0.01$ since

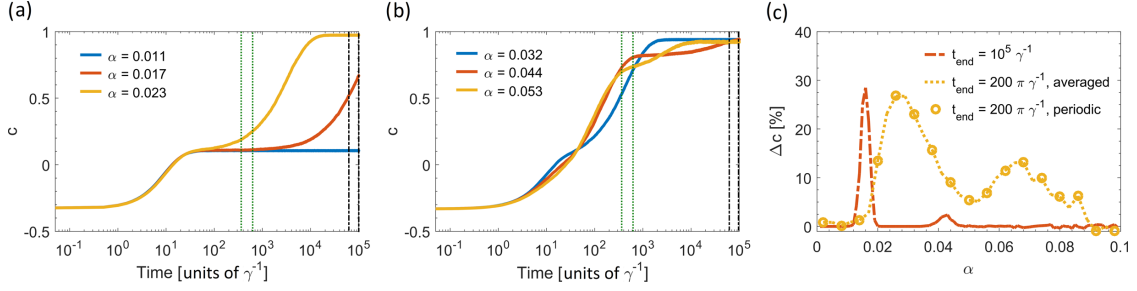


FIG. 7. Time evolution of the average value of c , Eq. (6), for the noise strengths (a) $\alpha = 0.011$, $\alpha = 0.017$, $\alpha = 0.023$, and (b) $\alpha = 0.032$, $\alpha = 0.044$, $\alpha = 0.053$. In each panel the dotted green (black) lines label the times $0.6 \times 200\pi\gamma^{-1}$ ($0.6 \times 10^5/\gamma$) and $t = 200\pi\gamma^{-1}$ ($10^5\gamma^{-1}$). (c) Variation Δc as a function of α for the simulation times $t_{\text{end}} = 200\pi\gamma^{-1}$ and $t_{\text{end}} = 10^5\gamma^{-1}$. For $t_{\text{end}} = 200\pi\gamma^{-1}$ the trajectories for periodic and time-averaged dynamics are shown. To generate comparable data, in both of these cases the same set of random numbers was used to simulate the trajectories. The integration method, trajectories, and initial conditions are the same as in Fig. 6.

for smaller values the numerical algorithm does not converge within the simulation times we considered. The corresponding Fokker-Planck equation for the probability distribution $p(\tilde{c}, t)$ takes the form

$$\partial_t p(\tilde{c}, t) = -\frac{\partial}{\partial \tilde{c}} \mathcal{J}(\tilde{c}, t), \quad (13)$$

with the current

$$\mathcal{J}(\tilde{c}, t) = \mu(\tilde{c})p(\tilde{c}, t) - \frac{1}{2} \frac{\partial}{\partial \tilde{c}} \sigma^2(\tilde{c})p(\tilde{c}, t). \quad (14)$$

We denote the steady state of $p(\tilde{c}, t)$ by $p_0(\tilde{c})$. The steady-state distribution $p_0(\tilde{c})$ fulfills $\partial_t p_0(\tilde{c}) = 0$, which corresponds to a constant current \mathcal{J} and explicitly reads

$$p_0(\tilde{c}) = \frac{\mathcal{N}}{\sigma^2(\tilde{c})} \exp[-\varphi(\tilde{c})] \quad (15)$$

with \mathcal{N} a normalization coefficient, warranting $\int_{-1}^1 d\tilde{c} p(\tilde{c}) = 1$, and

$$\varphi(\tilde{c}) = -\int_{-1}^{\tilde{c}} \frac{2\mu(y)}{\sigma^2(y)} dy \quad (16)$$

being the potential associated with the stationary solution. Potential and steady-state probability are shown in Figs. 8(b) and

8(c), respectively, as a function of \tilde{c} and α . For small values of α the minima and maxima are localized at the fixed points of the deterministic equation. The position of the minima are shifted towards the center of the interval as α is increased. Increasing the noise, moreover, decreases the barrier between minima: at sufficiently large α the system explores the full interval of values of \tilde{c} with longer residence times in the region at $\tilde{c} = 1$. At large α the effect of noise is to diffuse the solution across both paths keeping a bias towards $c = 1$. We remark that the noise-induced resonances observed in Fig. 6 are not captured by the equilibrium potential calculated from the one-dimensional Fokker-Planck equation.

V. CONCLUSION

We have analyzed the dynamics of a simple adaptive system as a function of the strength of a stochastic force. The system is composed by two paths connecting a sink and a source and subject to a periodic modulation of the dissipation rate between the two paths.

When the dissipation modulation frequency is smaller than the mean value of the dissipation rate, the system dynamics exhibit stochastic resonance, with the system periodically

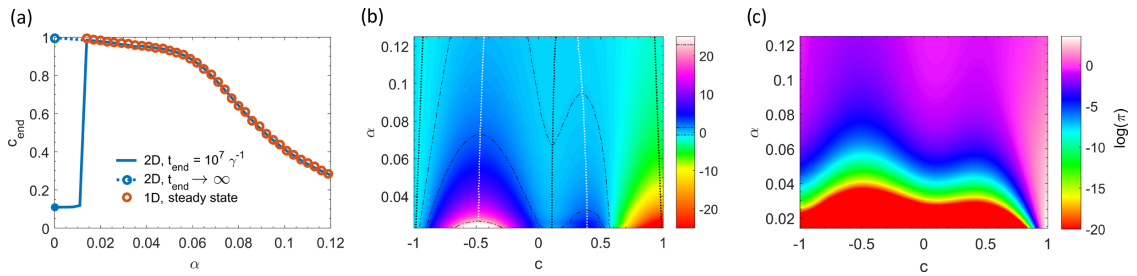


FIG. 8. (a) The risk function c_{end} as a function of the noise strength α for the parameters of Fig. 6. The prediction of the model of Eq. (4) at $t_{\text{end}} = 10^7\gamma^{-1}$ is compared with the one of the Fokker-Planck equation of Eq. (13) for $t \rightarrow \infty$. For the Fokker-Planck approach, the value of c_{end} is given by the average of the steady-state distribution $c_{\text{end}} = \int_{-1}^1 \tilde{c} p_0(\tilde{c}) d\tilde{c}$. Furthermore, we take only $\alpha > 0.01$ for the one-dimensional case as for smaller noise strengths the numerical algorithm does not converge within the considered simulation time. (b) Color plot of the potential φ , Eq. (16). The dotted white lines indicate the local maxima, the dotted black lines the local minima. (c) Color plot of the logarithm of the probability density distribution p_0 , Eq. (15), as a function of α and \tilde{c} .

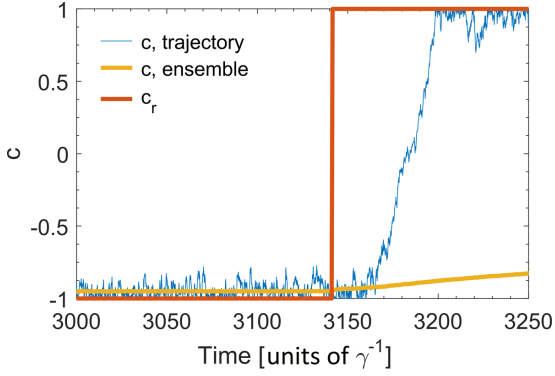


FIG. 9. Time evolution of the risk function c for $\omega = 10^{-3}\gamma$ and $\alpha = 0.051$. The blue line corresponds to one trajectory, the yellow line is the average over 5000 trajectories, and the red line displays $c_r(t)$. See Fig. 2(a).

switching to the path minimizing dissipation. At large frequencies, instead, the dynamics is reproduced by the secular equations, obtained by taking the time average of the damping coefficient over a period. The steady state is the fixed point of the secular equations which minimizes dissipation. The metastable configurations are in general the other fixed points of the secular equations, and the net effect of noise is to limit their lifetime. Nevertheless, the dynamics also exhibits other metastable configurations at certain values of the noise amplitude that are neither captured by stochastic resonance nor by the fixed point analysis. They exhibit the features of noise-induced resonances.

Our study suggests that noise could play a nontrivial role in both developing and optimizing algorithms for search problems, network design, and artificial intelligence. In the future we will extend this investigation to a network such as the configuration considered in Ref. [19] in the presence of a dynamically changing environment.

ACKNOWLEDGMENTS

The authors acknowledge support from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) Project No. 429529648, TRR 306 QuCoLiMa (Quantum Cooperativity of Light and Matter).

APPENDIX A: TIME EVOLUTION OF THE RISK FUNCTION

Figure 2(a) shows a single trajectory of the time evolution of the risk function c for $\omega = 10^{-3}\gamma$ and $\alpha = 0.051$. Figure 9 zooms over the behavior during one period.

APPENDIX B: SWITCHING TIME

We determine $t_s(\alpha, \omega)$, Eq. (8), using the first-passage time. This is the time to reach a metastable point, say, $c = -1$, when starting close to the other metastable point, say, $c = 1 - \epsilon$, in an interval with reflecting boundaries. This corresponds to the first passage time of the one-dimensional model of Sec. IV C

and takes the form [21,24]

$$t_s(\alpha, \epsilon) = \int_{1-\epsilon}^{-1} dy \frac{2}{\sigma^2(y)p(y)} \int_1^y dz p(z). \quad (\text{B1})$$

We evaluate $t_s(\alpha, \epsilon)$ numerically and use it in Eq. (8) in order to find the stochastic resonance condition.

APPENDIX C: FIXED POINTS AND LINEAR STABILITY ANALYSIS

In the following, we will perform a linear stability analysis of the system of differential Eq. (4). In the stationary regime, the conductivities D_1 and D_2 oscillate with the period $T = \frac{2\pi}{\omega}$ of the dissipation around the stationary values $\langle D_1 \rangle = D_1^*$ and $\langle D_2 \rangle = D_2^*$ and it holds

$$\begin{aligned} \left\langle \frac{\partial D_i}{\partial t} \right\rangle &= \lim_{t \rightarrow \infty} \left(\frac{1}{T} \int_t^{t+T} \frac{\partial D_i}{\partial t}(t') dt' \right) \\ &= 0. \end{aligned} \quad (\text{C1})$$

At this point, we remind the reader that we use $\tilde{D}_i = D_i/(D_1 + D_2)$. Applying this averaging procedure to Eq. (4) we obtain

$$0 = \langle f(\tilde{D}_i) \rangle - \langle \gamma_i(t) D_i \rangle. \quad (\text{C2})$$

We assume that the conductivities D_1 and D_2 are approximately constant during the period T . Thus, in the stationary regime the following relation holds:

$$\begin{aligned} \langle \gamma_i(t) D_i \rangle &= \lim_{t \rightarrow \infty} \left(\frac{1}{T} \int_t^{t+T} \gamma_i(t') D_i dt' \right) \\ &\approx \lim_{t \rightarrow \infty} \left(\frac{1}{T} D_i^* \int_t^{t+T} \gamma_i(t') dt' \right) = \gamma_i^{\text{eff}} D_i^*. \end{aligned} \quad (\text{C3})$$

This leads us to the equation

$$0 = \langle f(\tilde{D}_i) \rangle - \gamma_i^{\text{eff}} D_i^*. \quad (\text{C4})$$

Performing a Taylor expansion of the first term in this equation around the stationary values D_1^* and D_2^* to first order, we get

$$\begin{aligned} \langle f(\tilde{D}_i) \rangle &\approx f(\tilde{D}_i^*) + \left(\frac{\partial f}{\partial \tilde{D}_i} \frac{\partial \tilde{D}_i}{\partial D_1} \right) \Big|_{D_1=D_1^*, D_2=D_2^*} \langle (D_1 - D_1^*) \rangle \\ &\quad + \left(\frac{\partial f}{\partial \tilde{D}_i} \frac{\partial \tilde{D}_i}{\partial D_2} \right) \Big|_{D_1=D_1^*, D_2=D_2^*} \langle (D_2 - D_2^*) \rangle. \end{aligned} \quad (\text{C5})$$

We notice that the terms of first order vanish due to $\langle (D_i - D_i^*) \rangle = \langle D_i \rangle - D_i^* = 0$. Using this expression, Eq. (C4) yields

$$0 = f(\tilde{D}_i^*) - \gamma_i^{\text{eff}} D_i^*. \quad (\text{C6})$$

In the following, we introduce the deviation $y_i = D_i - D_i^*$ of the conductivities from their stationary value. We can cast Eq. (4) in the form

$$\frac{\partial y_i}{\partial t} = f(\tilde{D}_i) - \gamma_i(t)(y_i + D_i^*). \quad (\text{C7})$$

Assuming that the period T of the illumination is much smaller than the timescale on which the network structure

changes, we can cast Eq. (C7) in the form

$$\frac{\partial y_i}{\partial t} = f(\tilde{D}_i) - \gamma_i^{\text{eff}}(y_i + D_i^*). \quad (\text{C8})$$

Performing a Taylor expansion of the first term in this equation around the stationary values D_1^* and D_2^* to first order, we get

$$\begin{aligned} f(\tilde{D}_i) &\approx f(\tilde{D}_i^*) + \left(\frac{\partial f}{\partial \tilde{D}_i} \frac{\partial \tilde{D}_i}{\partial D_1} \right) \Big|_{D_1=D_1^*, D_2=D_2^*} y_1 \\ &\quad + \left(\frac{\partial f}{\partial \tilde{D}_i} \frac{\partial \tilde{D}_i}{\partial D_2} \right) \Big|_{D_1=D_1^*, D_2=D_2^*} y_2 \\ &= f(\tilde{D}_i^*) + A_i y_1 + B_i y_2, \end{aligned} \quad (\text{C9})$$

with

$$\begin{aligned} A_i &= \left(\frac{\partial f}{\partial \tilde{D}_i} \frac{\partial \tilde{D}_i}{\partial D_1} \right) \Big|_{D_1=D_1^*, D_2=D_2^*}, \\ B_i &= \left(\frac{\partial f}{\partial \tilde{D}_i} \frac{\partial \tilde{D}_i}{\partial D_2} \right) \Big|_{D_1=D_1^*, D_2=D_2^*}. \end{aligned} \quad (\text{C10})$$

Combining Eq. (C7) and Eq. (C10) with Eq. (C9) yields

$$\begin{aligned} \frac{\partial y_i}{\partial t} &= A_i y_1 + B_i y_2 - \gamma_i^{\text{eff}}(y_i + D_i^*) + \gamma_i^{\text{eff}} D_i^* \\ &= A_i y_1 + B_i y_2 - \gamma_i^{\text{eff}} y_i. \end{aligned} \quad (\text{C11})$$

In order to solve the system of differential Eq. (C11), we create the ansatz $y_i(t) = Y_i e^{\lambda t}$ with $\lambda \in \mathbb{C}$. This yields

$$\underbrace{\begin{pmatrix} \lambda - A_1 + \gamma_1^{\text{eff}} & -B_1 \\ -A_2 & \lambda - B_2 + \gamma_2^{\text{eff}} \end{pmatrix}}_{A:=} \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \quad (\text{C12})$$

To find nontrivial solutions, it must hold $\det(A) = 0$, which gives

$$\begin{aligned} 0 &= \det(A) \\ \Leftrightarrow 0 &= (\lambda - A_1 + \gamma_1^{\text{eff}})(\lambda - B_2 + \gamma_2^{\text{eff}}) - A_2 B_1, \\ \Leftrightarrow \lambda_{1,2} &= \frac{A_1 + B_2 - (\gamma_1^{\text{eff}} + \gamma_2^{\text{eff}})}{2}, \\ &\quad \pm \left(\frac{A_1 + B_2 - (\gamma_1^{\text{eff}} + \gamma_2^{\text{eff}})^2}{4} \right. \\ &\quad \left. - (A_1 - \gamma_1^{\text{eff}})(B_2 - \gamma_2^{\text{eff}}) + A_2 B_1 \right)^{1/2}. \end{aligned} \quad (\text{C13})$$

The fixed points (D_1^*, D_2^*) are given by Eq. (C6). They are stable if the corresponding values λ_i are negative. The corresponding fixed points c^* of this quantity can be calculated from the fixed points (D_1^*, D_2^*) . A fixed point c^* is considered stable if (D_1^*, D_2^*) is a stable fixed point. Figure 10 shows the fixed points for different choices of the average dissipation along the two paths.

APPENDIX D: FOKKER-PLANCK EQUATION

The variable \tilde{c} is assumed to undergo a time-continuous Markov process in the presence of a drift $\mu(\tilde{c})$ and a Itô

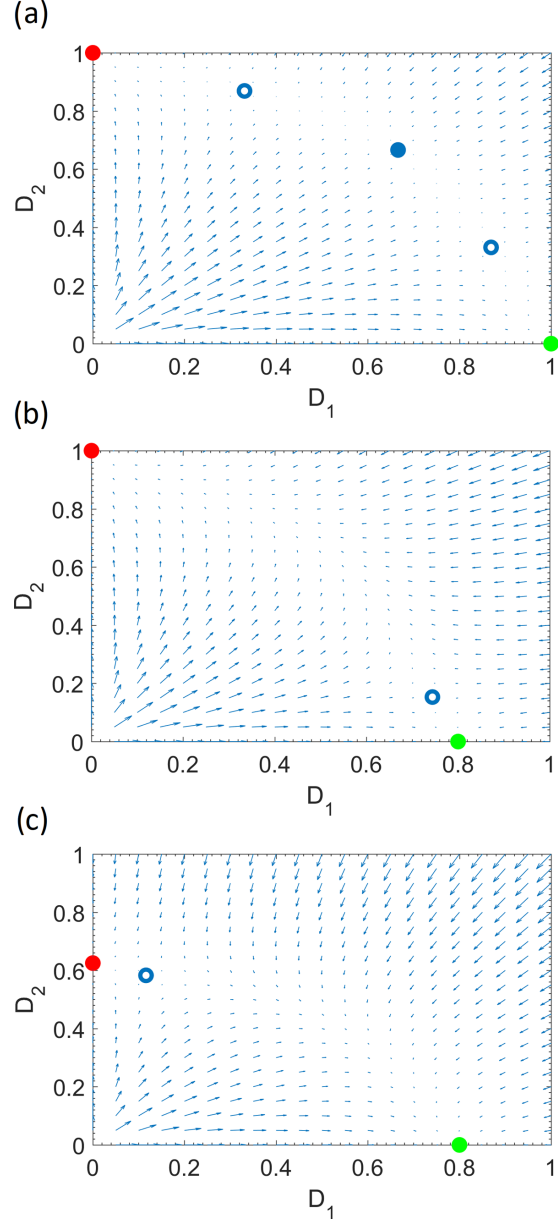


FIG. 10. Fixed points (D_1^*, D_2^*) of the secular equations. The arrows indicate the flow. Stable (unstable) fixed points are represented by solid (hollow) circles. The green (red) stable fixed points correspond to the system choosing the path with minimal (maximal) average dissipation. In (b) $\langle \Phi_1 \rangle = 0$ and $\langle \Phi_2 \rangle = 0.6$, while in (c) $\langle \Phi_1 \rangle = 0.25$ and $\langle \Phi_2 \rangle = 0.6$.

diffusion with amplitude $\sigma(\tilde{c}, t)$:

$$\frac{\partial \tilde{c}}{\partial t} = \mu(\tilde{c}) + \sigma(\tilde{c})\xi(t) \quad (\text{D1})$$

with $\xi(t)$ describing white noise. Drift and the diffusion coefficient are determined by means of equation-free analysis:

$$\mu(\bar{c}) = \frac{\langle c(t + \delta t) - c(t) | c(t) = \bar{c} \rangle_E}{\delta t}, \quad (\text{D2})$$

$$\sigma^2(\bar{c}) = \frac{\langle [c(t + \delta t) - c(t) - \mu(\bar{c})\delta t]^2 | c(t) = \bar{c} \rangle_E}{\delta t}, \quad (\text{D3})$$

where $\langle \cdot \rangle_E$ indicates a sample average and $c(t)$ is calculated from Eq. (6) from the values of $D_i(t)$, that are obtained by numerical integration of Eq. (4). We verify the validity of this approach by comparing the values of c_{end} we obtain by numerically integrating Eq. (D1) with the ones of Eq. (4); see Fig. 8(a). The Fokker-Planck equation corresponding to Eq. (D1) is given in Eq. (13).

-
- [1] M. C. Cross and P. C. Hohenberg, Pattern formation outside of equilibrium, *Rev. Mod. Phys.* **65**, 851 (1993).
- [2] D. Sherrington, Physics and complexity, *Phil. Trans. R. Soc. A* **368**, 1175 (2010).
- [3] T. Chou, K. Mallick, and R. K. P. Zia, Non-equilibrium statistical mechanics: From a paradigmatic model to biological transport, *Rep. Prog. Phys.* **74**, 116601 (2011).
- [4] A. V. Kravtsov and S. Borgani, Formation of galaxy clusters, *Annu. Rev. Astron. Astrophys.* **50**, 353 (2012).
- [5] M. Kardar, G. Parisi, and Y.-C. Zhang, Dynamic Scaling of Growing Interfaces, *Phys. Rev. Lett.* **56**, 889 (1986).
- [6] J. von Hardenberg, E. Meron, M. Shachak, and Y. Zarmi, Diversity of Vegetation Patterns and Desertification, *Phys. Rev. Lett.* **87**, 198101 (2001).
- [7] A.-K. Malchow, A. Azhand, P. Knoll, H. Engel, and O. Steinbock, From nonlinear reaction-diffusion processes to permanent microscale structures, *Chaos* **29**, 053129 (2019).
- [8] F. Folz, L. Wettmann, G. Morigi, and K. Kruse, *Phys. Rev. E* **99**, 050401(R) (2019).
- [9] A. Tero, R. Kobayashi, and T. Nakagaki, A mathematical model for adaptive transport network in path finding by true slime mold, *J. Theor. Biol.* **244**, 553 (2007).
- [10] T. Nakagaki, H. Yamada, and A. Tero, Maze-solving by an amoeboid organism, *Nature (London)* **407**, 470 (2000).
- [11] T. Nakagaki, H. Yamada, and M. Hara, Smart network solutions in an amoeboid organism, *Biophys. Chem.* **107**, 1 (2004).
- [12] C. Oettmeier, T. Nakagaki, and H. G. Döbereiner, Slime mold on the rise: The physics of *Physarum polycephalum*, *J. Phys. D* **53**, 310201 (2020).
- [13] A. Tero, S. Takagi, T. Saigusa, K. Ito, D. P. Bebber, M. D. Fricker, K. Yumiki, R. Kobayashi, and T. Nakagaki, Rules for biologically inspired adaptive network design, *Science* **327**, 5964 (2010).
- [14] A. Boussard, A. Fessel, C. Oettmeier, L. Briard, H.-G. Döbereiner, and A. Dussutour, Adaptive behaviour and learning in slime moulds: The role of oscillations, *Philos. Trans. R. Soc. B* **376**, 20190757 (2021).
- [15] C. Gao, C. Liu, D. Schenz, X. Li, Z. Zhang, M. Jusup, Z. Wang, M. Beekman, and T. Nakagaki, Does being multi-headed make you better at solving problems? A survey of *Physarum*-based models and computations, *Phys. Life Rev.* **29**, 1 (2019); *Physarum* inspires research beyond biomimetic algorithms, *ibid.* **29**, 51 (2019).
- [16] C. Gross, *Complex and Adaptive Dynamical Systems* (Springer, Berlin, 2013).
- [17] J. Noetel, V. L. S. Freitas, E. E. N. Macau, and L. Schimansky-Geier, Optimal noise in a stochastic model for local search, *Phys. Rev. E* **98**, 022128 (2018).
- [18] B. Meyer, C. Ansorge, and T. Nakagaki, The role of noise in self-organized decision making by the true slime mold *Physarum polycephalum*, *PLoS ONE* **12**, e0172933 (2017).
- [19] V. Bonifaci, E. Facca, F. Folz, A. Karrenbauer, P. Kolev, K. Mehlhorn, G. Morigi, G. Shahkarami, and Q. Vermande, *Physarum* multi-commodity flow dynamics, [arXiv:2009.01498v4](https://arxiv.org/abs/2009.01498v4) (2020).
- [20] H. Haken, *Synergetics, an Introduction: Nonequilibrium Phase Transitions and Self-Organization in Physics, Chemistry, and Biology* (Springer-Verlag, New York, 1983).
- [21] N. G. Van Kampen, *Stochastic Processes in Physics and Chemistry* (Elsevier, Amsterdam, 2007).
- [22] P. E. Kloeden and E. Platen, *Numerical Solution of Stochastic Differential Equations* (Springer, Berlin, 1992).
- [23] L. Gammaitoni, P. Hänggi, P. Jung, and F. Marchesoni, Stochastic resonance, *Rev. Mod. Phys.* **70**, 223 (1998).
- [24] B. Meyer, Optimal information transfer and stochastic resonance in collective decision making, *Swarm Intell.* **11**, 131 (2017).

CHAPTER II

PHYSARUM-INSPIRED MULTI-COMMODITY FLOW DYNAMICS

Physarum-inspired multi-commodity flow dynamics

Theoretical Computer Science 920 (2022) 1–20

© 2022 Elsevier B.V. All rights reserved.

DOI: <https://doi.org/10.1016/j.tcs.2022.02.001>

Authors: Vincenzo Bonifaci^a, Enrico Facca^b, Frederic Folz^c, Andreas Karrenbauer^d, Pavel Kolev^d, Kurt Mehlhorn^{d,*}, Giovanna Morigi^c, Golnoosh Shahkarami^e, Quentin Vermande^f

^a *Dipartimento di Matematica e Fisica, Università degli Studi Roma Tre, Roma, Italy*

^b *Scuola Normale Superiore, Pisa, Italy*

^c *Fachbereich Physik, Universität des Saarlandes, Saarbrücken, Germany*

^d *Max Planck Institute for Informatics, Saarbrücken, Germany*

^e *Max Planck Institute for Informatics and Fachbereich Informatik, Universität des Saarlandes, Germany*

^f *École Normale Supérieure, Paris, France*

* *Corresponding author. E-mail address: mehlhorn@mpi-inf.mpg.de (K. Mehlhorn).*

Abstract:

In wet-lab experiments, the slime mold *Physarum polycephalum* has demonstrated its ability to tackle a variety of computing tasks, among them the computation of shortest paths and the design of efficient networks. For the shortest path problem, a mathematical model for the evolution of the slime is available and it has been shown in computer experiments and through mathematical analysis that the dynamics solves the shortest path problem. In this paper, we generalize the dynamics to the network design problem. We formulate network design as the problem of constructing a network that efficiently supports a multi-commodity flow problem. We investigate the dynamics in computer simulations and analytically. The simulations show that the dynamics is able to construct efficient and elegant networks. In the theoretical part we show that the dynamics minimizes an objective combining the cost of the network and the cost of routing the demands through the network. We also give alternative characterizations of the optimum solution.



Contents lists available at ScienceDirect

Theoretical Computer Science

www.elsevier.com/locate/tcs



Physarum-inspired multi-commodity flow dynamics

Vincenzo Bonifaci^a, Enrico Facca^b, Frederic Folz^c, Andreas Karrenbauer^d,
Pavel Kolev^d, Kurt Mehlhorn^{d,*}, Giovanna Morigi^c, Golnoosh Shahkarami^e,
Quentin Vermande^f



^a Dipartimento di Matematica e Fisica, Università degli Studi Roma Tre, Roma, Italy

^b Scuola Normale Superiore, Pisa, Italy

^c Fachbereich Physik, Universität des Saarlandes, Saarbrücken, Germany

^d Max Planck Institute for Informatics, Saarbrücken, Germany

^e Max Planck Institute for Informatics and Fachbereich Informatik, Universität des Saarlandes, Germany

^f École Normale Supérieure, Paris, France

ARTICLE INFO

Article history:

Received 23 October 2020

Received in revised form 2 February 2022

Accepted 8 February 2022

Available online 17 February 2022

Communicated by C. Blum

Keywords:

Physarum

Network design

Multi-commodity flow

Dynamical system

ABSTRACT

In wet-lab experiments, the slime mold *Physarum polycephalum* has demonstrated its ability to tackle a variety of computing tasks, among them the computation of shortest paths and the design of efficient networks. For the shortest path problem, a mathematical model for the evolution of the slime is available and it has been shown in computer experiments and through mathematical analysis that the dynamics solves the shortest path problem. In this paper, we generalize the dynamics to the network design problem. We formulate network design as the problem of constructing a network that efficiently supports a multi-commodity flow problem. We investigate the dynamics in computer simulations and analytically. The simulations show that the dynamics is able to construct efficient and elegant networks. In the theoretical part we show that the dynamics minimizes an objective combining the cost of the network and the cost of routing the demands through the network. We also give alternative characterizations of the optimum solution.

© 2022 Elsevier B.V. All rights reserved.

1. Introduction

Physarum polycephalum is a slime mold in the Mycetozoa group [11]. Its cells can grow to considerable size and it can form networks. In wet-lab experiments, the slime mold *Physarum polycephalum* was applied to a diverse variety of computing problems: computation of shortest paths in a network [36], computation of minimum risk paths [35], design of efficient networks [42,44,9], computation of Voronoi and Delaunay diagrams [5,37], computing circuits and electronics [24] and many more. We refer the reader to [7,6] for a survey of the many problems that can be attacked using live *Physarum polycephalum* and for which the slime is able to find good or even optimal solutions to instances of limited size. Fig. 1 illustrates the shortest path and the network formation experiments in [36,42].

* Corresponding author.

E-mail address: mehlhorn@mpi-inf.mpg.de (K. Mehlhorn).

<https://doi.org/10.1016/j.tcs.2022.02.001>

0304-3975/© 2022 Elsevier B.V. All rights reserved.

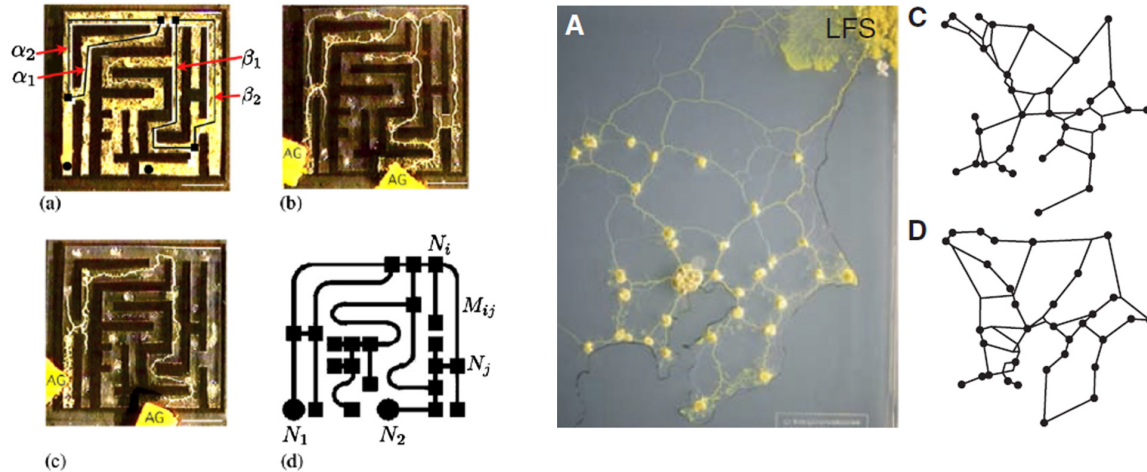


Fig. 1. The figure on the left shows the shortest path experiment. It is reprinted from [36]. The edges of a graph were uniformly covered with Physarum (subfigure (a)) and food in the form of oat-meal was provided at the locations labeled AG in subfigure (b). After a while the slime retracted to the shortest path connecting the two food source (subfigure (c)). The underlying graph is shown in (d). The figure on the right shows the network design experiment. It is reprinted from [42]. Food was provided at many places (the larger dots in the picture) and the slime was constrained to live in an area that looks similar to the greater Tokyo region. The large dot in the center corresponds to Tokyo and the empty region below it corresponds to Tokyo bay. The slime formed a network connecting the food sources. The two graphs on the right compare the network built by the slime (label C) with the railroad network in the Tokyo region (label D).

There is also considerable work aimed at understanding the inner workings of *Physarum polycephalum*, for example, how global synchronization can result from random peristaltics [1], how information can be transported and a memory can exist in an organism without a nervous system [2,25], and whether tubes of the mold can transfer electricity [43].

It is important to stress that the plasmodium of *Physarum polycephalum* is not an automaton.¹ The papers [30,32] clearly demonstrate the limits of slime-mold computations even for the shortest path problem and [6,30] argue convincingly that conventional computing terminology should be applied with great care when discussing biological systems. After all, the solutions constructed in wet-lab experiments strongly depend on the initial conditions, e.g., how much food is provided and how the plasmodium is distributed initially, the solutions are not strictly optimal but only approximately optimal, the outcomes of the experiments are not deterministic and hence hard to reproduce, and the maze in Fig. 1 has fairly narrow edges and hence guides the slime towards building nearly straight connections. The papers [38,22] discuss more generally the question what it means for a biological or physical device to compute. According to their definition, to which we subscribe, *Physarum polycephalum* does not compute. A comprehensive survey of analog computing models is given in [15] and [33,34] discuss the differences and commonalities of biology and computing.

The experimental work mentioned above instigated the development of *Physarum-inspired algorithms* mimicking (parts of) the behavior of the slime mold. This is akin to algorithms mimicking ant colonies [17], neural networks [8], simulated annealing [26], and other bio-inspired computing paradigms. *Physarum-inspired algorithms* have been used to solve a variety of computational tasks, for example, the design of transportation networks [42,45,46] and supply-chain networks [47]. For the shortest path problem, a mathematical model in the form of a coupled system of differential equations was given for the evolution of the slime, the biological relevance of the model was argued, and the model was shown to solve shortest path problems in computer simulations [41]. Mathematical proofs that the model solves (it is explained below what is meant by solves) shortest path problems can be found in [31,12]. The *Physarum* dynamics is also able to solve more general linear programs [39,40,23,27]. It is important to emphasize that the *Physarum* dynamics is inspired by the behavior of the mold, it captures – at best – parts of the behavior of the mold.

The paper [41] is the starting point for this work. Tero et al. model the slime network as an electrical network $G = (V, E)$ with time varying resistors. Each edge e of the network has a fixed positive length c_e and a time-varying diameter $x_e(t)$. In this paper, we will refer to c_e as the cost of the edge and to x_e as the capacity of the edge. The resistance of e at time t is then $r_e(t) = c_e/x_e(t)$. Let s_0 and s_1 be two fixed vertices in the network; they represent the two food sources. One unit of electrical current is sent from s_0 to s_1 . Let $q_e(t)$ be the current flowing across e . Then the capacity of e evolves according to the differential equation

$$\dot{x}_e = \frac{d}{dt}x_e(t) = |q_e| - x_e \quad \text{for all } e \in E, \quad (1)$$

¹ There is a small community of researchers that think differently, see [21] for example.

i.e., the diameter of an edge grows (shrinks, stays unchanged) if the flow is larger than (smaller than, equal to) the current diameter. The equations for the different edges are coupled because the flow through an edge e depends on the resistance of all other edges. As is customary, we write \dot{x} for the derivative with respect to time and drop the time-argument of x and q . Tero et al. showed in computer simulations that a discretization of the model converges to the shortest path connecting the source and the sink in the following sense: $x_e(\infty) = 1$ for the edges on the shortest path and $x_e(\infty) = 0$ for the other edges. This assumes that the shortest path is unique. Bonifaci et al. [12] proved that the dynamics converges to the shortest source-sink path, i.e., that the Physarum dynamics (1) solves the shortest path problem. A related dynamics

$$\dot{x}_e = |q_e|^\mu - x_e \quad \text{for all } e \in E,$$

where μ is a constant larger than one, may converge to a path different from the shortest path depending on the initial conditions [32].

In this paper, we generalize the model of Tero et al. [41] to network design. We introduce a simple model, again in the form of a system of differential equations,

- that for the case of the shortest path problem agrees with the model proposed in Tero et al.,
- that in computer simulations qualitatively reconstructs the behavior observed in the wet-lab network design experiments [36,42], and
- that is amenable to theoretical analysis.

We do not argue biological plausibility and we do not claim any biological relevance. We also do not try to describe a general model of the Physarum that fits all experimental setups; less ambitiously, we focus on the network design experiments. This is a paper in algorithm design and analysis.

The shortest path problem can be viewed as a network design problem. Given two vertices in a graph, the goal is to construct the cheapest network connecting the given vertices. The solution is the shortest path connecting the vertices. The shortest path problem can also be viewed as a minimum cost flow problem. We want to send one unit of flow between the given vertices and the cost of sending a certain amount across an edge is equal to the cost of the edge times the amount sent. The solution is the shortest path connecting the given vertices.

Networks are designed for a particular purpose. For this paper, the purpose is multi-commodity flow. Suppose that we have many pairs of vertices between which we want to send flow. We want to construct a network that satisfies the many demands in an economical way. Economical could mean many things: minimum cost of the network (that's the Steiner tree problem), shortest realization of each demand (then the network is the union of the shortest paths), or something in the middle, i.e., some combination of the total cost of the network and the cost of routing the demands in the network. We assume *economies of scale*, i.e., that there is some benefit in sharing a connection, i.e., the cost of sending one unit each of two commodities across an edge is lower than two times the cost of sending one unit of one commodity across the edge. In Section 5, we give examples of how sharing is encouraged by our model. The principles of our model are simple. As in Tero et al. each edge has a cost and a capacity. We have demands between pairs of vertices; this could be passengers entering the network at some station and leaving the network at some other station. The demand i between vertices $s_i^{(1)}$ and $s_i^{(2)}$ leads to an electrical flow q^i . For each edge e , we aggregate the individual flows $q^i(e)$ to an overall flow $q(e)$. This flow is then used in equation (1). For the aggregation we use either the one-norm or the two-norm of the vector $(q^1(e), q^2(e), \dots)$ and find that the two-norm aggregation is to be preferred. We mention that one-norm aggregation is used in [45].

This paper is organized as follows. In Section 2 we introduce our model and in Section 3 we review our results. In Section 4 we discuss related work. In Section 5, we report about paper-and-pencil and computer experiments. The analytical part starts with Section 6. We review basic facts about electrical flows. In subsequent sections, we prove the existence of a solution defined for $t \in [0, \infty)$, characterize the fixed points, introduce a Lyapunov function \mathcal{L} for the dynamics, derive further properties of the Lyapunov minimum, show convergence to the Lyapunov minimum, and finally make a connection to mirror descent. Section 13 offers conclusions.

2. The model

Before presenting our model we want to fix some notations. Given a matrix $M \in \mathbb{R}^{n \times m}$, we will denote with M_i , M^j , and $M_{i,j}$ its i^{th} -row, its j^{th} -column, and its entry i, j , respectively. Moreover, we denote with $\text{Tr}(M) = \sum_{i=1}^n M_{i,i}$ the trace of a square matrix $M \in \mathbb{R}^{n,n}$.

Our model for the multi-commodity network design problem is inspired by the Physarum model for the shortest path problem and its generalization to linear programming. Let $A \in \mathbb{R}^{n \times m}$ be an arbitrary real matrix and let b^1 to b^k in \mathbb{R}^n be k right-hand sides such that each of the linear systems $Af = b^i$ is solvable.

The reader may want to think of A as the node-arc incidence matrix of a connected undirected graph G with n nodes and m edges, i.e., for each $e = (u, v) \in E$, the column $(A^T)_e$ has an entry $+1$ in position u and entry -1 in position v ; the orientation of the edge is arbitrary, but fixed. We have k different source-sink pairs $(s_i^{(1)}, s_i^{(2)})$, $1 \leq i \leq k$. Let $b^i \in \mathbb{R}^n$ be the vector with entry $+1$ in position $s_i^{(1)}$ and entry -1 in position $s_i^{(2)}$. All other entries of b^i are zero. Since G is assumed to be connected, the linear system $Af = b^i$ admits solutions for all i . We refer to this setting as the *multi-commodity flow setting*.

Now, for any non-negative vectors $x \in \mathbb{R}^m$ and $c \in \mathbb{R}^m$, we define the following matrices

$$X = \text{diag}(x) \quad C = \text{diag}(c) \quad L(x) = AX C^{-1} A^T. \tag{2}$$

Given a solution f of $Af = b^i$, we use

$$E_x(f) = \begin{cases} \sum_e c_e / x_e f_e^2 & \text{if } \text{supp } f \subseteq \text{supp } x, \\ \infty & \text{if } \text{supp } f \not\subseteq \text{supp } x, \end{cases}$$

to denote the *energy* of f with respect to x . Let $q^i(x) \in \mathbb{R}^m$, or simply q^i , be the minimum energy solution i.e.,

$$q^i = \underset{f \in \mathbb{R}^m}{\text{argmin}} \left\{ E_x(f) : Af = b^i \right\}. \tag{3}$$

The optimal solution of the optimization problem (3) (see Section 6 for details) is given by

$$q^i(x) = X C^{-1} A^T p^i(x), \quad i = 1, \dots, k, \tag{4}$$

where $p^i(x)$, or simply p^i , is defined as any solution to

$$L(x)p^i = b^i, \quad i = 1, \dots, k.$$

In the multi-commodity flow setting, the minimal energy solution q^i is simply the electrical flow realizing the demand b^i and p^i are the corresponding node potentials. The node potentials are not unique; they can be made unique by defining a particular node as ground, i.e., giving it potential zero. The electrical flow is induced by the potential drops $A^T p^i$ multiplied by the conductivity $X C^{-1}$. If we now define the matrix $B \in \mathbb{R}^{n,k}$ by

$$B = (B^1, \dots, B^k) := (b^1, \dots, b^k),$$

we can express the potentials, the potential drops per unit cost, and the fluxes corresponding to the different commodities in the following matrix form

$$\begin{aligned} P &= (P^1, \dots, P^k) &:= (p^1, \dots, p^k) & \text{with } L(x)P = B \\ \Lambda &= (\Lambda^1, \dots, \Lambda^k) &:= (\lambda^1, \dots, \lambda^k) &= C^{-1} A^T P \\ Q &= (Q^1, \dots, Q^k) &:= (q^1, \dots, q^k) &= X \Lambda. \end{aligned} \tag{5}$$

Note that we use P^i and p^i interchangeably and similarly for Λ and Q .

We are now ready to define our model. We let the vector Q_e of values $Q_{e,i}$ for any edge e determine the capacity of an edge and study different ways of combining the individual solutions, in particular, one-norm and two-norm.² This leads to the following dynamics:

$$\dot{x}_e = -x_e + \sum_i |Q_{e,i}| = x_e \left(-1 + \sum_i \frac{|Q_{e,i}|}{x_e} \right) = x_e \left(-1 + \sum_i |\Lambda_{e,i}| \right) = x_e (\|\Lambda_e\|_1 - 1), \tag{6}$$

$$\dot{x}_e = -x_e + \sqrt{\sum_i Q_{e,i}^2} = x_e \left(\sqrt{\sum_i \left(\frac{Q_{e,i}}{x_e} \right)^2} - 1 \right) = x_e \left(\sqrt{\sum_i \Lambda_{e,i}^2} - 1 \right) = x_e (\|\Lambda_e\|_2 - 1). \tag{7}$$

In (6), we form the one-norm $\|\Lambda_e\|_1$ of the different normalized potential drops across any edge e , and in (7), we form the two-norm $\|\Lambda_e\|_2$. For $k = 1$, the one-norm and the two-norm dynamics coincide. The results of this paper suggest that the two-norm dynamics is the appropriate generalization to larger k .

The following *generalized Physarum dynamics* introduced in [13] subsumes the two-norm dynamics as a special case. For each $e \in E$, let g_e be a non-negative, increasing and differentiable function with $g_e(1) = 1$:

$$\dot{x}_e = x_e (g_e(\|\Lambda_e\|_2) - 1). \tag{8}$$

² In the multi-commodity flow setting, the q_i 's are flows in the network G . The fact that flows from different demand pairs on the same edge do not cancel each other (not even partially) seems a bit strange at the microscopic level. After all, physically, only the cytoplasm is being transported. How does an edge “distinguish” between the cytoplasm of pair i and the cytoplasm of pair i' ? For this reason, we do not claim biological plausibility for our model. When $k = 1$, clearly this was not an issue.

The two-norm dynamics is a special case with $g_e(z) = 1 + (z - 1)$. Other examples are $g_e(z) = 1 + d_e(z - 1)$ and $g_e(z) = 1 + d_e(z^2 - 1)$ where $d_e > 0$ is the “reactivity” [27] of edge e , $g_e(z) = z^{\mu_e}$ for some $\mu_e > 0$ or $g_e(z) = (1 + \alpha_e)z^{\mu_e} / (1 + \alpha_e z^{\mu_e})$ for some $\mu_e, \alpha_e > 0$.

The right hand sides of (6) to (8) are defined for any

$$x \in \Omega = \left\{ x \in \mathbb{R}_{\geq 0}^m : \exists P \in \mathbb{R}^{n,k} \text{ solving } L(x)P = B \right\}.$$

3. Our results

In the analytical part of the paper, we ask and answer the following questions for the generalized Physarum dynamics. We have little to say about the one-norm dynamics.

- Do the dynamics have a solution $x(t)$ with $t \in [0, \infty)$?
- Do the dynamics converge?
- What are the fixed points and the limit points of the dynamics?
- What do the dynamics optimize?
- How can we characterize the limit points?

In the experimental part of the paper, we perform computer and pencil-and-paper simulations of the dynamics and address the following questions:

- How strong are the sharing effects of the dynamics? How far deviate individual flows from their shortest realization in order to benefit from sharing edges with other flows?
- Do the dynamics construct “nice” networks? Does it qualitatively reconstruct the wet-lab experiments in [42]?

Our first result concerns the existence of solutions with domain $[0, \infty)$ for the generalized Physarum dynamics.

Theorem 1. *Let $x(0) \in \mathbb{R}_{> 0}^m$. The generalized Physarum dynamics has a solution $t \mapsto x(t) \in \mathbb{R}_{> 0}^m$ for $t \in [0, \infty)$.*

The cost of a capacity vector x is defined as

$$\mathcal{C}(x) = c^T x = \sum_e c_e x_e.$$

The energy dissipation for a single demand b induced by a capacity vector x is defined as

$$\min_{f: Af=b} E_x(f) = \sum_e r_e / x_e q_e^2 = b^T p = p^T L(x)p,$$

where q is the minimum energy solution of $Af = b$ with respect to x and p is the corresponding node potential. We will show the second equality in Section 6. The last equality follows from $L(x)p = b$. The energy dissipation $\mathcal{E}(x)$ for a set of k demands b^1, \dots, b^k is the sum of the energy dissipations for the individual demands, i.e.,

$$\mathcal{E}(x) = \sum_i E_x(q^i) = \sum_i (b^i)^T p^i = \text{Tr}(P^T L(x)P),$$

where p^i is the node potential with respect to the minimum energy solution q^i to the i -th demand.

The fixed points of a dynamics are the points x with $\dot{x}_e = 0$ for all e . We use \mathcal{F}_1 and \mathcal{F}_g to denote the fixed points (also called equilibrium points) of the one-norm and the generalized dynamics.

Lemma 1 (The fixed points of the one-norm dynamics). $x \in \mathcal{F}_1$ iff for all e either $x_e = 0$ or $\|\Lambda_e\|_1 = 1$. The latter condition is equivalent to $\|Q_e\|_1 = x_e$ as well as to $\|(A^T)_e P\|_1 = c_e$.

The fixed points of the generalized dynamics have a remarkable property. For a fixed point $x \in \mathcal{F}_g$, the cost $\mathcal{C}(x)$ equals the dissipated energy $\mathcal{E}(x)$.

Lemma 2 (The fixed points of the generalized Physarum dynamics). $x \in \mathcal{F}_g$ iff for all e either $x_e = 0$ or $\|\Lambda_e\|_2 = 1$. The latter condition can be expressed equivalently by $\|(A^T)_e P\|_2 = c_e$ and also by $x_e = \|Q_e\|_2$. Further, for every $x \in \mathcal{F}_g$ we have $x \geq 0$, $AQ = B$, and

$$\text{Tr}(B^T P) = \mathcal{E}(x) = \mathcal{C}(x) = c^T x,$$

i.e., for fixed points of the generalized Physarum dynamics the cost equals the energy dissipation.

The beauty goes further. The dynamics follows a path along which the sum of cost and energy dissipation decreases and, under mild additional assumptions, minimizes the sum in the limit of $t \rightarrow \infty$. Let

$$\mathcal{L}(x) = \frac{1}{2}(\mathcal{C}(x) + \mathcal{E}(x)) = \frac{1}{2}(c^T x + \sum_i (b^i)^T p^i)$$

be one-half of the sum of the cost and the energy dissipation of the network. We show in Section 9 that the function \mathcal{L} is a Lyapunov function for the generalized Physarum dynamics, in particular, $\mathcal{L}(x(t))$ is a non-negative decreasing function of t . Formally, the conditions for a Lyapunov function are: $\mathcal{L}(x) \geq 0$ for $x \in \Omega$, $\frac{d}{dt} \mathcal{L}(x(t)) = \langle \nabla \mathcal{L}, \dot{x} \rangle \leq 0$ for all t . In the case $k = 1$, \mathcal{L} is also a Lyapunov function for the one-norm dynamics as shown in [27]. Let

$$\mathcal{V} = \{x : \langle \nabla \mathcal{L}, \dot{x} \rangle = 0\}$$

be the set of points in which the dynamics does not decrease the Lyapunov function any further. It follows from general theorems about dynamical systems that the dynamics converges to the set \mathcal{V} . We show that \mathcal{V} is equal to the set of fixed points \mathcal{F}_g , and that under mild additional assumptions, the dynamics converges to the minimizer of the Lyapunov function.

Theorem 2. $\mathcal{F}_g = \mathcal{V}$ and the generalized Physarum dynamics converges to \mathcal{V} . Moreover, if the set \mathcal{F}_g is finite and any two points in \mathcal{F}_g have distinct values of \mathcal{L} , the dynamics $x(t)$ converges to $x^* = \operatorname{argmin}_{x \in \mathbb{R}_{\geq 0}^m} \mathcal{L}(x)$.

The minimum of the Lyapunov function can also be characterized in alternative ways.

Theorem 3. The following quantities MinQ , MaxP , and MinL are equal.

$$\operatorname{MinQ} = \min_{Q \in \mathbb{R}^{m \times k}} \left\{ \sum_e c_e \|Q_e\|_2 : A Q = B \right\}, \tag{9}$$

$$\operatorname{MaxP} = \max_{P \in \mathbb{R}^{n \times k}} \left\{ \operatorname{Tr}(B^T P) : \|(A^T)_e P\|_2 \leq c_e \text{ for all } e \right\}, \tag{10}$$

$$\operatorname{MinL} = \min_{x \in \mathbb{R}_{\geq 0}^m} \mathcal{L}(x). \tag{11}$$

Moreover, there are optimizers Q^* , R^* and x^* such that

$$\begin{aligned} x_e^* &= \|Q_e^*\|_2 \quad \text{for all } e, \\ L(x^*)P^* &= B, \\ Q^* &= X^* C^{-1} A^T P^*. \end{aligned}$$

It is instructive to interpret the theorem for the case $k = 1$, A the node-arc incidence matrix of a directed graph, and b a vector with one entry $+1$ and one entry -1 and all other entries equal to zero. Then $\operatorname{MinQ} = \min_{q \in \mathbb{R}^m} \{ \sum_e c_e |q_e| : A q = b \}$ is the minimum cost of a flow realizing b in the underlying undirected network and $\operatorname{MaxP} = \max_{p \in \mathbb{R}^n} \{ b^T p : |p_v - p_u| \leq c_e \text{ for all } e = (u, v) \}$ is the maximum distance between the two nodes designated by b for any distance function on the nodes satisfying the cost constraints imposed by c . Both values are equal to the cost of the minimum cost path connecting the two designated nodes and hence $\operatorname{MinQ} = \operatorname{MaxP}$. The third characterization via $\operatorname{MinL} = \min_{x \geq 0} \mathcal{L}(x)$ is non-standard. Note that $\mathcal{L}(x) = (c^T x + b^T p)/2$, where p are node potentials driving a current of 1 between the nodes designated by b in the network with edge resistances c_e/x_e . Then $b^T p$ is the potential difference between the two designated nodes which, since the driven current is one, is the effective resistance between the two designated nodes. In Lemma 10, we will show $\frac{\partial}{\partial x_e} \mathcal{L}(x) = \frac{c_e}{2}(1 - \|\Lambda_e\|_2^2)$, i.e., the minimizer x^* of $\mathcal{L}(x)$ must satisfy $x_e^* \neq 0 \Rightarrow |(A^T)_e p^*| = c_e$, where p^* are node potentials corresponding to x^* . Note that for an edge $e = (u, v)$, $|(A^T)_e p^*| = |p_v^* - p_u^*|$ is the potential drop on e . Orient all edges such that potential drops are positive and consider any path W (W for Weg) in $\operatorname{supp}(x^*)$ connecting the two designated nodes. Then

$$b^T p^* = \sum_{e \in W} (A^T)_e p^* = \sum_{e \in W} c_e,$$

since the potential difference between the two designated nodes is the sum of the potential drops along W . Thus any two paths in $\operatorname{supp}(x^*)$ connecting the two designated nodes must have the same cost and hence (assuming that any two such paths have distinct cost) $\operatorname{supp}(x^*)$ contains a single path connecting the two designated nodes. In fact, $\operatorname{supp}(x^*)$ is equal to such a path. Now $\sum_{e \in W} c_e x_e + \sum_{e \in W} c_e/x_e = \sum_{e \in W} (c_e x_e + c_e/x_e)$ is minimized for $x_e = 1$ for all $e \in W$ and then is equal to

twice the cost of W . Of course, the cost of W is minimized for the shortest undirected path connecting the two designated nodes.

We turn to the result of our computer experiments. We performed three *case studies*, two small and the third inspired by the wet-lab experiment by [42]. The first example (Section 5.1) can be treated analytically, we consider a ring with three nodes with a demand of one between any pair of nodes. We will see that a solution using all three edges is superior to a solution using only two edges. Also, we see confirmed that for fixed points of the two-norm dynamics the cost of the network and the total energy dissipation is the same. The second example (Section 5.2) concerns flow in the Bow-Tie graph shown in Fig. 3. We will investigate the incentive for sharing links. In this example, the demands can share a link at the cost of increasing the distance between the terminals. We will see that sharing pays off. The third Example 5.3 is based on the example in [42]. We will see that the dynamics forms nice networks similar to the networks in [42].

4. Related work

This paper is inspired by [41], [45], and [42]. We already explained the connection to these papers in detail in the previous sections.

Shortly after this work was posted on arXiv, a closely related paper [28] was posted. It considers the multi-commodity transportation problem in graphs. Let A be the node-arc incidence matrix of a directed graph and vectors b_1 to b_k with $1^T b_i = 0$ for all i be k supply-demand vectors. Each arc of the graph has a fixed cost c_e and a capacity x_e . This is what we called the multi-commodity flow setting in Section 2. They model the interaction between the different commodities in exactly the same way as we do, i.e., for each i , a minimum energy solution q_i is a minimum energy solution with respect to the resistances c_e/x_e of the system $Aq = b_i$. The different flows on each edge e are combined by forming their two-norm. The difference lies in the dynamics. The paper considers the dynamics

$$\dot{x}_e = x_e^\beta \|\Lambda_e\|_2^2 - x_e, \tag{12}$$

where $\beta \in (0, 2)$ is a parameter. For $\beta = 1$, this dynamics is a special case of our generalized dynamics obtained by setting $g_e(z) = 1 + (z^2 - 1)$.

The paper investigates the dynamics analytically and experimentally. For the experimental evaluation, the paper uses the Paris metro. In the analytical part, the paper shows that the fixed points satisfy $x_e^{3-\beta} = \|Q_e\|_2^2$ and that the solution to the optimization problem

$$\text{minimize } \sum_e \frac{c_e}{x_e} \|Q_e\|_2^2 \text{ subject to } \sum_e c_e x_e^{2-\beta} = K \text{ and } A Q = B,$$

where K is a positive constant, satisfies $x_e^{3-\beta} = C \cdot \|Q_e\|_2^2$, where C is a constant, i.e., fixed points and optimal solutions to the optimization problem exhibit the same relation between x_e and Q_e . The paper also contains an extensive discussion of the simulation of the dynamics and of the numerical solution of the optimization problem above.

Convergence of the dynamics is not shown. However, a slight modification of the Lyapunov function used in this paper also works for their dynamics. Assume $\beta \in (0, 2)$ and define

$$\mathcal{L}(x) = \frac{1}{2} \left(\frac{1}{2-\beta} c^T x^{2-\beta} + \sum_{i=1}^k (b^i)^T p^i \right).$$

Lemma 3 (Gradient of \mathcal{L}). For all $e \in E$,

$$\frac{\partial}{\partial x_e} \mathcal{L}(x) = \frac{c_e}{2} (x_e^{1-\beta} - \|\Lambda_e\|_2^2). \tag{13}$$

Proof. The derivative $\frac{\partial}{\partial x_e} \sum_{i=1}^k (b^i)^T p^i$ is computed in Lemma 10 and $\frac{\partial}{\partial x_e} c^T x^{2-\beta} = (2-\beta)c_e x_e^{1-\beta}$. \square

Theorem 4. The function $\mathcal{L} : \Omega \mapsto \mathbb{R}$ is a Lyapunov function for the dynamics (12), i.e., $\frac{d}{dt} \mathcal{L}(x(t)) \leq 0$ for all t . Let

$$\mathcal{V} = \{x \in \Omega : \langle \nabla \mathcal{L}(x), \dot{x} \rangle = 0\}.$$

Then \mathcal{V} is equal to the fixed points of (12).

Proof. Since $\frac{d}{dt} \mathcal{L}(x(t)) = \langle \nabla \mathcal{L}(x), \dot{x} \rangle$, we obtain

$$\frac{d}{dt} \mathcal{L}(x(t)) = \sum_e \frac{c_e}{2} (x_e^{1-\beta} - \|\Lambda_e\|_2^2) \cdot (x_e^\beta \|\Lambda_e\|_2^2 - x_e) = - \sum_e \frac{c_e}{2} x_e^\beta (x_e^{1-\beta} - \|\Lambda_e\|_2^2)^2 \leq 0.$$

We have equality if and only if for all e either $x_e = 0$ or $\|\Lambda_e\|_2^2 = x_e^{1-\beta}$. Thus $x \in \mathcal{V}$ if and only if x is a fixed point of (12). \square

Lemma 4. For fixed points x of (12), $\sum_e c_e x_e^{2-\beta} = \sum_i (b^i)^T p_i$.

Proof.

$$\sum_i (b^i)^T p^i = \sum_e \sum_i \frac{c_e}{x_e} q_{ei}^2 = \sum_e \sum_i \frac{c_e}{x_e} (x_e \Lambda_{ei})^2 = \sum_e c_e x_e \|\Lambda_e\|_2^2 = \sum_e c_e x_e^{2-\beta}. \quad \square$$

We mentioned in the result section that our generalized Physarum dynamics converges to a solution for which the cost $\sum_e c_e x_e$ is equal to the dissipated energy $\sum_i (b^i)^T p_i$. The dynamics (12) allows a wider choice of equilibrium points.

5. Case studies

5.1. Multi-commodity flow in a ring

Consider a graph consisting of three vertices a, b , and c and three edges connecting them into a 3-cycle. All edges have cost one and we have a demand of one between any pair of nodes. An equilibrium uses either two edges or three edges.

5.1.1. Two edge solution

We will see below that, for each of the dynamics, the solution is symmetric, i.e., both edges have the same capacity in equilibrium, say z . The flow across both edges is two. For each demand, the potential drop on each edge is $1/z$. So the total energy spent is $\mathcal{E} = 2/z + 2 \cdot 1/z = 4/z$ (one demand uses two edges for a energy dissipation of $2/z$ and two demands use one edge for a energy dissipation of $1/z$ each) and the total cost $C = 2z$. Thus $C + \mathcal{E} = 4/z + 2z$.

One-norm dynamics: The current across each edge is 2 and hence $z = 2$ for each of the existing edges. Thus $C = c^T z = 4$, $\mathcal{E} = \sum_i (b^i)^T p^i = 4/z = 2$, and $C + \mathcal{E} = 6$.

Two-norm dynamics The current across each edge is $1 + 1$ and hence $z = \sqrt{2}$. Thus $C = c^T z = 2\sqrt{2}$, $\mathcal{E} = \sum_i (b^i)^T p^i = 4/\sqrt{2} = 2\sqrt{2}$ and $C + \mathcal{E} = 4\sqrt{2}$. Note that $C = 2\sqrt{2} = \mathcal{E}$. This is not a coincidence as we show in Lemma 9.

Optimum: We have $C + \mathcal{E} = 4/z + 2z$. The optimum is attained for $z = \sqrt{2}$. Note that this corresponds to the equilibrium of the two-norm. This is not a coincidence as we show in Theorem 5.

5.1.2. Three edge solution

We will see below that, for each of the dynamics, the solution is symmetric, i.e., all edges have the same capacity in equilibrium, say z , and hence the same resistance $1/z$. Then $C = 3z$. Each demand is routed partly the short way and partly the long way. Since the long way has twice the resistance, the amount routed the short way is twice the amount routed the long way, i.e., $2/3$ of each demand is routed the short way and $1/3$ is routed the long way.

For each demand, let Δ be the potential drop between source and sink. The total energy spent is 3Δ . The potential drop Δ must be such that it can drive a current of $2/3$ across a wire of conductance z . Thus $\Delta = 2/(3z)$. We obtain $C + \mathcal{E} = 3z + 2/z$.

One-norm dynamics: z is equal to the total current flowing across an edge and hence $z = 2/3 + 2 \cdot 1/3 = 4/3$ and $\Delta = 1/2$. So $C = c^T z = 4$, $\mathcal{E} = \sum_i (b^i)^T p^i = 3/2$, and $C + \mathcal{E} = 11/2$. This is better than for the two-edge equilibrium.

Two-norm dynamics For each edge, we have one flow of value $2/3$ and two flows of value $1/3$ and hence $z^2 = 4/9 + 2 \cdot 1/9 = 6/9$. Thus $z = \sqrt{2/3}$. Δ must be such that it can drive a current of $2/3$ across a wire of conductance $\sqrt{2/3}$ and hence $\Delta = \sqrt{2/3}$.

Hence $C = c^T z = 3 \cdot \sqrt{2/3} = \sqrt{6}$ and $\mathcal{E} = \sum_i (b^i)^T p^i = 3 \cdot \sqrt{2/3} = \sqrt{6}$. Note that again we have the same value for the cost C and the total energy spent \mathcal{E} . For the sum, we obtain $C + \mathcal{E} = 2\sqrt{6}$. This is better than the two-edge equilibrium.

Optimum: For a general value of z , we have $C + \mathcal{E} = 3z + 2/z$. This is minimized for $z = \sqrt{2/3}$, i.e., the equilibrium of the two-norm is equal to the minimum combined cost solution.

5.1.3. Computer simulations

Table 1 shows the results of a typical simulation. For the simulation we discretized the differential equation and applied an Euler forward scheme.

5.2. The Bow-Tie Graph

Consider the graph shown in Fig. 2; we refer to this graph as a bow-tie. The edge costs are as shown and we are sending one unit each between nodes 0 and 1 and nodes 4 and 5, i.e., $b_0 = (1, -1, 0, 0, 0, 0)$ and $b_1 = (0, 0, 0, 0, 1, -1)$. For each pair the direct path connecting the pair has length 10, the path using the middle edge has length $L + 2$ and the path using the edge connecting the other pair has length 14. Figs. 3 and 4 show the results of a simulation. Initial x -values were chosen randomly in the interval $[1, 10]$. We observe:

Table 1
The initial z -values were chosen randomly between $1/1000$ and 1 . In all cases, the system converged to the 3-edge equilibrium. Note that $0.82 \approx \sqrt{2/3}$ and $1.33 \approx 4/3$.

	the final z -values of the three edges		
two-norm dynamics	0.8160	0.8167	0.8166
one-norm dynamics	1.331	1.327	1.342

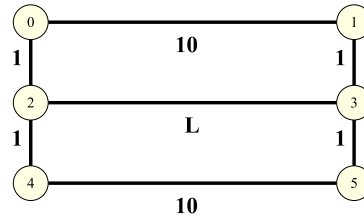


Fig. 2. The top and the bottom horizontal edge have cost 10, the middle horizontal edge has cost L , and all other edges have cost 1. We are sending one unit between nodes 0 and 1 and one unit between nodes 4 and 5.

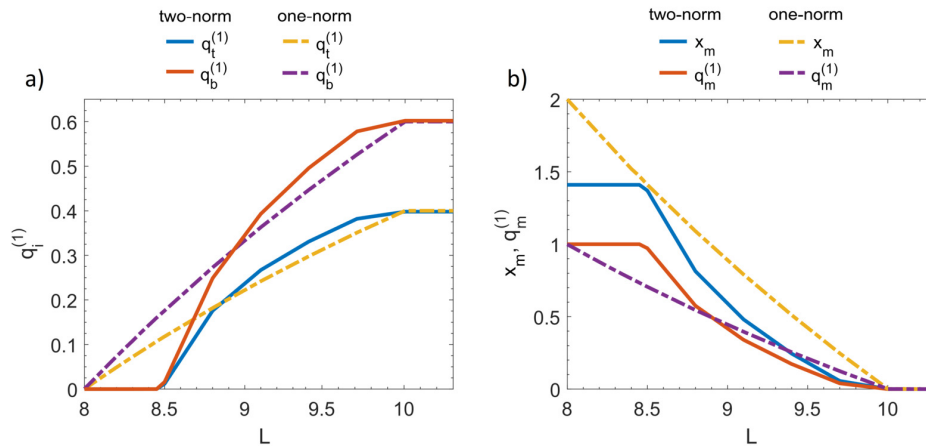


Fig. 3. Simulation of the Bow-Tie Graph: The plot depicts the quantities $q_b^{(1)}$, $q_m^{(1)}$, and $q_t^{(1)}$ (= the split-up of the flow from node 0 to node 1 across the three horizontal edges bottom, middle, and top) and x_m (= the capacity of the middle edge) as a function of L in the range $[8, 10.3]$. For $L < 8$, the quantities are as for $L = 8$, and for $L > 10.3$, the quantities are as for $L = 10.3$. For $q^{(2)}$ the flow across the middle edge is the same and the flow across the other edges is reversed. For all L , $q_b^{(1)} + q_m^{(1)} + q_t^{(1)} = 1$. (1) The image on the left shows $q_b^{(1)}$ and $q_t^{(1)}$. (2) The image on the right shows the capacity x_m and the flow $q_m^{(1)}$ across the middle edge. We have $x_m = \sqrt{2(q_m^{(1)})^2} = \sqrt{2} \cdot q_m^{(1)}$ in the case of the two-norm and $x_m = 2q_m^{(1)}$ in the case of the one-norm. (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)

- For $L \leq 8$, both dynamics generate essentially the same solution. All flow is essentially routed through the middle edge.
- For the two-norm dynamics: For $L < 8.5$, the sharing effect is strong and basically all flow is routed through the middle edge. Note that for $L > 8$, the path through the middle edge is not the shortest path for either demand. Starting at $L = 8.5$, the top and the bottom edge are also used. For $L \geq 10$, only the top and the bottom edge are used and this may give the impression that there is no sharing effect for large L . This is not the case. The solution for $L = \infty$ is easily computed analytically. Because of symmetry, a fraction a of each flow is routed the short way (length 10) and a fraction $1 - a$ is routed the long way (length 14). So through each edge, we have a flow of value a and a flow of value $1 - a$ and hence all edges will have the same capacity in equilibrium; call it x . Therefore the flows must be in the same ratio as the costs, i.e., $a/(1 - a) = 10/14$. This solves to $a = 7/12$. Then $x = \sqrt{a^2 + (1 - a)^2} = \sqrt{74}/12 \approx 8.6023$. The cost of the network is then $24 \cdot \sqrt{74}/12 = 2\sqrt{74} \approx 17.2$ and the dissipated energy is the same. Assume now that we delete the vertical edges. Then each demand is routed separately and the bottom and the top edge will have a capacity of one each. The cost of the network will be 20 and the dissipated energy will also be 20. This is considerably more than the cost of the network constructed by our dynamics.
- For the one-norm dynamics: Starting at $L = 8.05$, the top and the bottom edge are also used. For $L \geq 10.3$, only the top and the bottom edge are used.

L	6.5	6.8	7.1	7.4	7.7	8.0	8.3	8.6	8.9	9.2	9.5	9.8
C	13.2	13.6	14.0	14.5	14.9	15.3	15.7	16.1	16.4	16.6	16.7	16.7
\mathcal{E}	13.2	13.6	14.0	14.5	14.9	15.3	15.7	16.1	16.4	16.6	16.7	16.7

Fig. 4. Simulation results for the two-norm dynamics for the bow-tie graph. The cost $C = c^T x$ and the energy $\mathcal{E} = \sum_i (b^i)^T p^i$ for the limit states for different values of L . Note that $C = \mathcal{E}$ always.

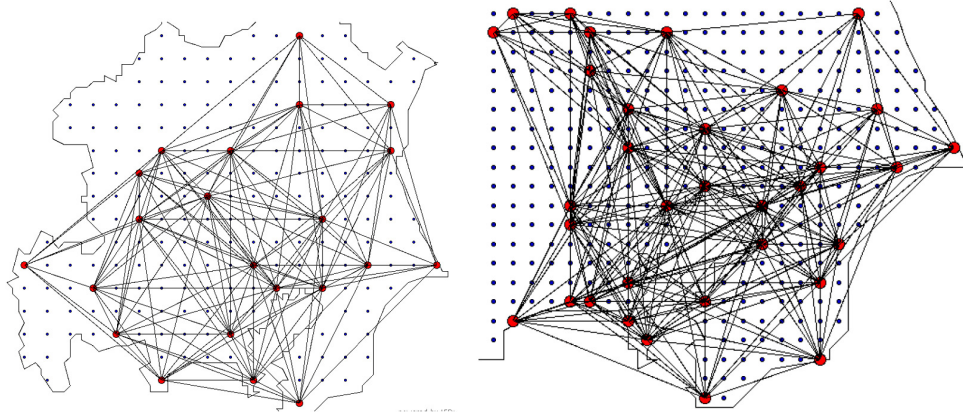


Fig. 5. The polygonal region on the left is a digitization of the Greater Tokyo Region. The red dots indicate major cities. We set up 140 demands. For each red city, we created a demand of one unit to any other red city within a certain distance threshold. The threshold is about 1/2 the distance between the topmost and the bottommost red point. The region on the right is approximately the right lower quadrant of the region on the left. For the placement of the terminals we tried to copy the placement shown in Fig. 1. We set up 282 demands, again between cities below a certain distance threshold. The demands are one, except if one of the terminals corresponds to Tokyo. Then the demand is seven; this is as in [42].

- For the two-norm dynamics, the cost C and the dissipated energy \mathcal{E} are equal in the limit; see Fig. 4.

5.3. A case study inspired by [42]

In [42] the slime molds ability to construct elegant networks in investigated. The slime is allowed to grow in a region that is shaped according to the greater Tokyo region and food is provided at many different places. Fig. 1 shows the results of the wet-lab experiment and compares a network constructed by the slime to the railroad network around Tokyo. The paper also reports about a computer experiment. Repeatedly a pair of food sources was chosen at random and a step of the shortest path dynamics was executed. Figure 4 in [42] shows the results of the computer experiment. No details are given in the paper and also the positions of the food sources are not given in detail.

We tried to repeat the experiment with the two-norm dynamics. For this purpose, we digitized the boundary of the Greater Tokyo region in the form of a polygonal region and overlaid a regular grid in which each node is connected to its up to eight neighbors (north, northwest, west, southwest, south, southeast, east, northeast) inside the region. The edge lengths are 1 for the horizontal and vertical edges and 1.41 for the diagonal edges. We perturbed the edge lengths slightly by adding $r \cdot 0.05$ for a random integer $r \in [-3, 3]$ so as to avoid many equal length path. We chose the terminals in two different ways.

First choice: We chose the largest 25 cities in Greater Tokyo region according to Wikipedia and generated 140 demands. Each city was connected to all other cities whose distance is below a certain threshold. For the threshold we chose about 1/2 times the diameter of the region. The left side of Fig. 5 shows the input and Fig. 6 shows the output of a computer simulation.

Second choice: We mimicked the choice of sites used in [42]. We generated 282 demands again between any pair of sites whose distance is below a certain threshold. The demands are 1, except if one of the terminals corresponds to Tokyo. Then the demand is seven; this is as in [42]. The right side of Fig. 5 shows the input and Fig. 7 shows the output of a computer simulation.

6. Preliminaries

We recall the definition of energy dissipation and cost. For a capacity vector $x \in \mathbb{R}_{\geq 0}^m$ and a vector $f \in \mathbb{R}^m$ with $\text{supp}(f) \subseteq \text{supp}(x)$, we use

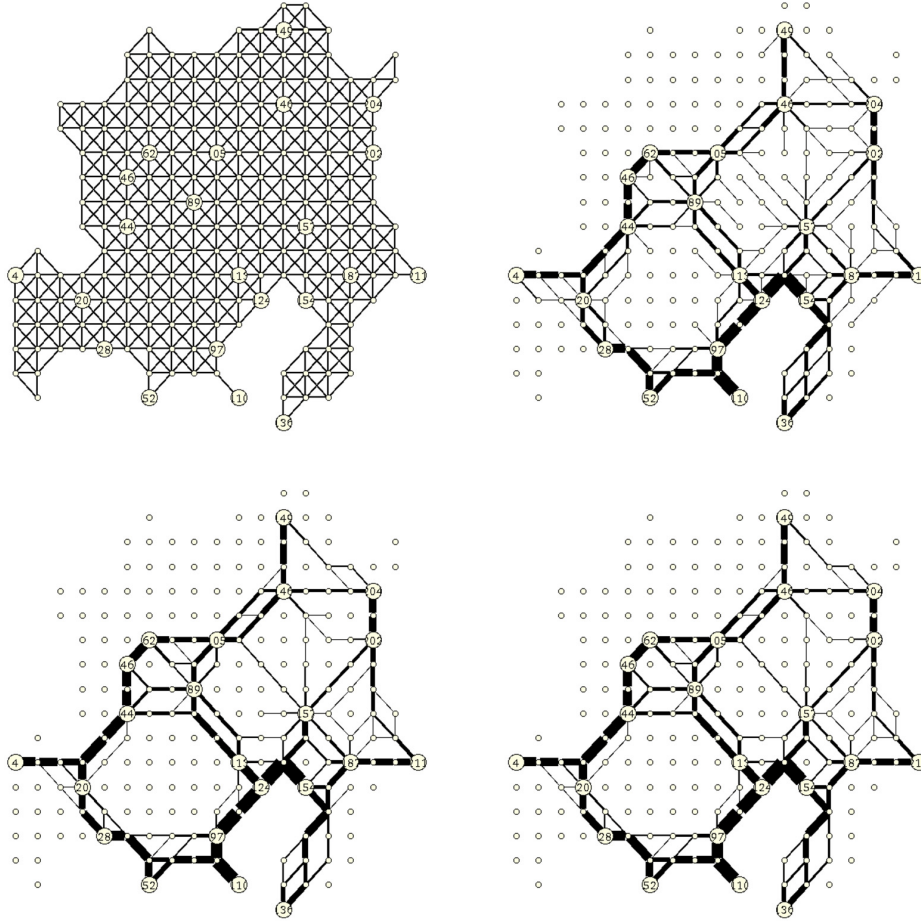


Fig. 6. An output of a simulation of the two-norm dynamics on the left instance in Fig. 5. The graph in the upper left corner shows the initial graph. Each node is connected to its up to 8 neighbors. The length of the horizontal and vertical edges is approximately 1, the length of the diagonals is approximately 1.41. All capacities are 0.5 initially and the capacity of an edge is indicated by its thickness. The following figures show the state after 1950 and 4875 iterations. For the situation after 4875 iterations, we also show the reduced graph where we iteratively removed nodes of degree one (which are not terminals). The numbers inside the nodes are unique identifiers; they have no meaning beyond this.

$$E_x(f) = \begin{cases} \sum_e (c_e/x_e) f_e^2 & \text{if } \text{supp } f \subseteq \text{supp } x, \\ \infty & \text{if } \text{supp } f \not\subseteq \text{supp } x, \end{cases}$$

to denote the *energy dissipation* of f with respect to x . Strictly speaking we should sum only over the e in $\text{supp } x$. We use the convention $0^2/0 = 0$ to justify summing over all edges e . Further, we use

$$\mathcal{C}(f) = \sum_e c_e |f_e| = c^T |f|$$

to denote the *cost* of f . Note that

$$E_x(x) = \sum_e (c_e/x_e) x_e^2 = \sum_e c_e x_e = \mathcal{C}(x).$$

We use R to denote the diagonal matrix with entries c_e/x_e . Energy-minimizing solutions are induced by node potentials $p \in \mathbb{R}^n$ according to the following equations:

$$b = Aq, \tag{14}$$

$$q = R^{-1} A^T p, \tag{15}$$

$$AR^{-1} A^T p = b. \tag{16}$$

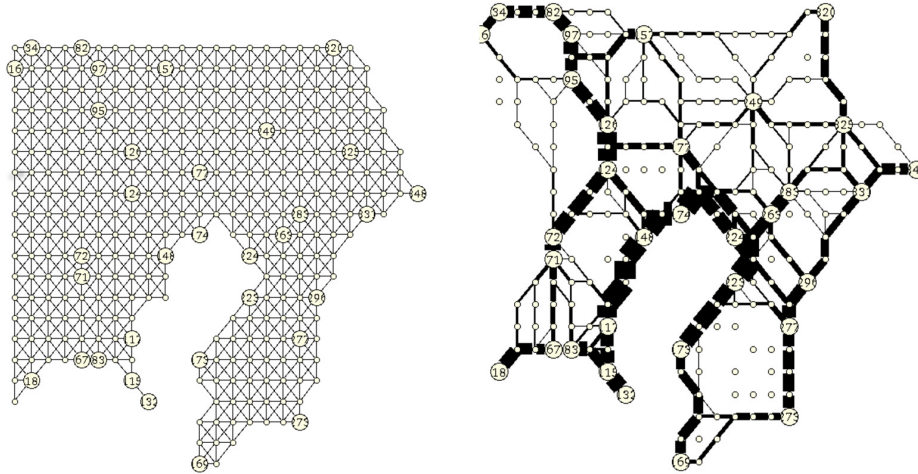


Fig. 7. An output of a simulation of the two-norm dynamics on the right instance in Fig. 5. The graph in the upper left corner shows the initial graph. Each node is connected to its up to 8 neighbors. The length of the horizontal and vertical edges is approximately 1, the length of the diagonals is approximately 1.41. All capacities are 0.5 initially and the capacity of an edge is indicated by its thickness. The figure on the right show the state after 16000 iterations where we iteratively removed nodes of degree one.

We give a short justification why the equations above characterize the energy minimizing solution to the linear system. The energy minimizing solution q minimizes the quadratic function $\sum_e (c_e/x_e)q_e^2$ subject to the constraints $Aq = b$ and $\text{supp}(q) \subseteq \text{supp}(x)$. The KKT conditions (see [16, Subsection 5.5]) state that at the optimum, the gradient of the objective is a linear combination of the gradients of the constraints, i.e.,

$$2(c_e/x_e)q_e = \sum_i p^i A_{i,e} \quad \text{for all } e \in \text{supp}(x)$$

for some vector $p \in \mathbb{R}^n$ and $q_e = 0$ for $e \notin \text{supp}(x)$. Absorbing the factor 2 into p yields equation (15). Substitution of (15) into (14) gives (16). The energy-minimizing solution is unique. It exists if and only if $b \in \text{Im}A$. Node potentials p are not unique, but the values of $b^T p$ and $p^T L(x)p$ are equal for all solutions of (16).

Lemma 5. Assume $x > 0$. Then $\text{Ker}L(x) = \text{Ker}A^T$ and $\text{Im}L(x) = \text{Im}A$. The values $b^T p$, $p^T L(x)p$ and $q = XC^{-1}A^T p$ do not depend on the particular solution of $L(x)p = b$.

Proof. Clearly, $\text{Ker}A^T \subseteq \text{Ker}L(x)$. So assume $z \in \text{Ker}L(x)$. Then $L(x)z = 0$ and hence $z^T L(x)z = 0$. Let $D^{1/2}$ be the diagonal matrix with entries $\sqrt{x_e/c_e}$. Then

$$0 = z^T L(x)z = z^T A D^{1/2} D^{1/2} A^T z = \|D^{1/2} A^T z\|_2^2$$

and hence $D^{1/2} A^T z = 0$ and further $0 = A^T z$. So $z \in \text{Ker}A^T$.

Clearly, $\text{Im}L(x) \subseteq \text{Im}A$. So assume $b \notin \text{Im}L(x)$. Then the rank of the matrix obtained by augmenting $L(x)$ by the column b is larger than the rank of $L(x)$ (Rouché-Capelli theorem) and hence there is a vector r such that $r^T b \neq 0$ and $r^T L(x) = 0$. Since $L(x)$ is symmetric, $L(x)r = 0$ and hence $r \in \text{Ker}L(x) = \text{Ker}A^T$. So $0 = A^T r = (r^T A)^T$. Thus r also proves $b \notin \text{Im}A$.

Let p and \bar{p} be node potentials. Then $L(x)p = b = L(x)\bar{p}$ and hence $\bar{p} - p \in \text{Ker}L(x)$. Then

$$b^T \bar{p} = b^T p + b^T (\bar{p} - p) = b^T p + p^T L(x)^T (\bar{p} - p) = b^T p + p^T L(x) (\bar{p} - p) = b^T p$$

and

$$XC^{-1}A^T \bar{p} = XC^{-1}A^T p + XC^{-1}A^T (\bar{p} - p) = XC^{-1}A^T p.$$

Finally, $b^T p = p^T L(x)p$. \square

For the arc-node incidence matrix A of a connected graph, the kernel $\text{Ker}A^T$ consists of the all-ones vector in \mathbb{R}^n . We can make the node potential unique by requiring $p_v = 0$ for some fixed node v , i.e., by grounding node v .

Lemma 6. Let ℓ be the dimension of $\text{Ker}A^T$ and let $K \in \mathbb{R}^{n \times \ell}$ be a matrix whose columns form a basis of $\text{Ker}A^T$. Let $V' \subseteq [n]$ with $|V'| = \ell$ be such that the submatrix of K with rows selected by V' is nonsingular. Then the solution p to $L(x)p = b$ with $p_v = 0$ for all $v \in V'$ is unique, i.e. “grounding all nodes in V' makes the potential unique”.

Proof. Observe first that such a solution exists. Let p be an arbitrary solution to $L(x)p = b$. Then there is a vector $\lambda \in \mathbb{R}^\ell$ such that $(K\lambda)_v = p_v$ for all $v \in V'$ and hence $p - K\lambda$ is the desired node potential. Assume now that we have two solutions p and p' with $p_v = p'_v$ for all $v \in V'$. Then $p - p' \in \text{Ker}L(x) = \text{Ker}A^T$ and $(p - p')_v = 0$ for all $v \in V'$. Since $p - p' \in \text{Ker}A^T$ there is a $\lambda \in \mathbb{R}^\ell$ such that $p - p' = K\lambda$. Then $(K\lambda)_v = 0$ for all $v \in V'$. Since the columns of K are independent, this implies $\lambda = 0$ and hence $p = p'$. \square

The next Lemma gives alternative expressions for the energy $E_x(q)$ of the minimum energy solution.

Lemma 7. $E_x(q) = \sum_e (c_e/x_e)q_e^2 = b^T p = p^T L(x)p$, where p is any solution of (16).

Proof. This holds since

$$E_x(q) = q^T Rq = p^T AR^{-1}RR^{-1}A^T p = p^T AR^{-1}A^T p = p^T L(x)p = p^T b. \quad \square$$

Finally, we recapitulate a bound on the components of q established in [40] and slightly improved form in [10, Lemma 3.3].

Lemma 8. Let D be the maximum absolute value of a square submatrix of A . Then $|q_e| \leq D\|b\|_1$ for every $e \in [m]$.

7. Existence of a solution

We prove Theorem 1. The right-hand side (8) is locally Lipschitz-continuous in x . The function g_e is locally Lipschitz by assumption, the q^i 's are infinitely often differentiable rational functions in the x_e and hence locally Lipschitz. Furthermore, locally Lipschitz-continuous functions are closed under additions and multiplications. Thus $x(t)$ is defined and unique for $t \in [0, t_0)$ for some t_0 .

Since g_e is non-negative, we have $\dot{x}_e \geq -x$ and thus $x_e \geq x_e(0)e^{-t}$. Hence, $x(t) > 0$ for all t . By assumption $b^i \in \text{Im}A$ for all i , and hence whenever $x(t) > 0$, we have solutions q^i with $\text{supp}(q^i) \subseteq \text{supp}(x)$.

In Section 9, we will show that \mathcal{L} is a Lyapunov function for the dynamics (8). Thus

$$c^T x \leq \mathcal{L}(x) \leq \mathcal{L}(x(0))$$

and hence x stays in a bounded domain.

It now follows from general results about the solutions of ordinary differential equations [20, Corollary 3.2] that $t_0 = \infty$.

8. Fixed points

A point x is a fixed point iff $\dot{x} = 0$. We use \mathcal{F}_g for the set of fixed points of (8).

Lemma 9 (The fixed points of the generalized Physarum dynamics). $x \in \mathcal{F}_g$ iff for all e either $x_e = 0$ or $\|\Lambda_e\|_2 = 1$. The latter condition is equivalent to $x_e = \|Q_e\|_2$ or $\|(A^T)_e P\|_2 = c_e$. For $x \in \mathcal{F}_g$, $\mathcal{C}(x) = \mathcal{E}(x)$.

Proof. We have $\dot{x} = 0$ iff we have $x_e = 0$ or $g_e(\|\Lambda_e\|_2) = 1$ for all e . Since g_e is increasing and $g_e(1) = 1$, the latter condition is tantamount to $\|\Lambda_e\|_2 = 1$ which expands to $\sum_i ((A^T)_e P^i)^2 = c_e^2$. Multiplying both sides by $(x_e/c_e)^2$ yields $x_e^2 = \sum_i (Q_{e,i})^2$.

For $x \in \mathcal{F}_g$, we have

$$\mathcal{E}(x) = \sum_e \sum_i \frac{c_e}{x_e} (Q_{e,i})^2 = \sum_e \frac{c_e}{x_e} \cdot x_e^2 = \sum_e c_e x_e = \mathcal{C}(x). \quad \square$$

9. Lyapunov function

Let

$$\mathcal{L}(x) = \frac{1}{2} \left(c^T x + \sum_{i=1}^k (b^i)^T p^i \right).$$

We will show that \mathcal{L} is a Lyapunov function for the dynamics (8). The function \mathcal{L} was introduced in [19]. For $k = 1$, [18] shows that \mathcal{L} is a Lyapunov function for the one-norm dynamics and [27] shows that this holds true also for the generalized Physarum dynamics. The calculations below generalize the calculations in these papers. They are similar to the calculations in [14, Lemma 2.6].

Lemma 10 (Gradient of \mathcal{L}). For all $e \in E$,

$$\frac{\partial}{\partial x_e} \mathcal{L}(x) = \frac{c_e}{2} (1 - \|\Lambda_e\|_2^2). \quad (17)$$

Proof. Recall $L(x) = AX C^{-1} A^T$. Let $e \in [m]$ be arbitrary. Then $\frac{\partial}{\partial x_e} L(x) = \frac{1}{c_e} A^e (A^T)_e$. From $L(x)p = b$ and $\frac{\partial}{\partial x_e} b = 0$, we obtain

$$0 = \frac{\partial}{\partial x_e} L(x)p = \frac{\partial L(x)}{\partial x_e} p + L(x) \frac{\partial p}{\partial x_e}$$

and thus

$$L(x) \frac{\partial p}{\partial x_e} = -\frac{1}{c_e} A^e (A^T)_e p.$$

Hence, we have

$$\frac{\partial}{\partial x_e} b^T p = b^T \frac{\partial p}{\partial x_e} = p^T L(x) \frac{\partial p}{\partial x_e} = -\frac{1}{c_e} p^T A^e (A^T)_e p = -c_e \left(\frac{(A^T)_e p}{c_e} \right)^2,$$

and more generally,

$$\frac{\partial}{\partial x_e} \sum_i (b^i)^T p^i = -c_e \sum_i \left(\frac{(A^T)_e p^i}{c_e} \right)^2 = -c_e \|\Lambda_e\|_2^2.$$

The claim follows. \square

Theorem 5. The function $\mathcal{L} : \Omega \mapsto \mathbb{R}$ is a Lyapunov function for the dynamics (8), i.e., $\mathcal{L}(x) \geq 0$ for all $x \in \Omega$, and $\frac{d}{dt} \mathcal{L}(x(t)) \leq 0$ for all t . Let

$$\mathcal{V} = \{x \in \Omega : \langle \nabla \mathcal{L}(x), \dot{x} \rangle = 0\}.$$

Then $\mathcal{V} = \mathcal{F}_g$.

Proof. $\mathcal{L}(x) \geq 0$ for all $x \in \Omega$ is obvious.

Since $\frac{d}{dt} \mathcal{L}(x(t)) = \langle \nabla \mathcal{L}(x), \dot{x} \rangle$, we obtain

$$\frac{d}{dt} \mathcal{L}(x(t)) = \sum_e \frac{c_e}{2} (1 - \|\Lambda_e\|_2^2) \cdot x_e (g_e(\|\Lambda_e\|_2) - 1) \leq 0,$$

where the inequality holds since $g_e(\|\Lambda_e\|_2) - 1$ and $\|\Lambda_e\|_2 - 1$ have the same sign, as g_e is a non-negative and increasing function with $g_e(1) = 1$.

We have equality if and only if for all e either $x_e = 0$ or $\Lambda_e = 1$. Thus $x \in \mathcal{V}$ if and only if $x \in \mathcal{F}_g$. \square

10. Further properties of the Lyapunov minimum

We give two alternative characterizations for the minimum of the Lyapunov function. This extends [18, Proposition 2] from $k = 1$ to arbitrary k .

Theorem 6. The following quantities MinQ , MaxP , and MinL are equal.

$$\text{MinQ} = \min_{Q \in \mathbb{R}^{m \times k}} \left\{ \sum_e c_e \|Q_e\|_2 : A Q = B \right\}, \quad (18)$$

$$\text{MaxP} = \max_{P \in \mathbb{R}^{n \times k}} \left\{ \text{Tr}[B^T P] : \|(A^T)_e P\|_2 \leq c_e \text{ for all } e \right\}, \quad (19)$$

$$\text{MinL} = \min_{x \in \mathbb{R}_{\geq 0}^m} \mathcal{L}(x). \quad (20)$$

Moreover, there are optimizers Q^* , P^* and x^* such that

$$\begin{aligned} x_e^* &= \|Q_e^*\|_2 \quad \text{for all } e, \\ L(x^*)P^* &= B, \\ Q^* &= X^* C^{-1} A^T P^*. \end{aligned}$$

Lemma 11. Let Q^* be a minimizer of (18) and let x^* be defined by $x_e^* = \|Q_e^*\|_2$ for all e . Then $x^* \in \mathcal{F}_g$. Moreover, there is a potential matrix $P \in \mathbb{R}^{n \times k}$ such that $L(x^*)P = B$ and $Q = X^*C^{-1}A^T P$, $\sum_e c_e \|Q_e\|_2 = \text{Tr}[B^T P] = \mathcal{L}(x^*)$, and $\|(A^T)_e P\|_2 \leq c_e$ for all e . The objective values of (18) to (20) satisfy $\text{MinL} \leq \text{MinQ} \leq \text{MaxP}$.

Proof. We start by slightly reformulating the minimization problem (18). This is necessary since the function $Q_e \mapsto \|Q_e\|_2$ is not differentiable for $Q_e = 0$ and hence the KKT-conditions cannot be applied. We formulate equivalently:

$$\min \sum_e c_e x_e \text{ subject to } A Q = B, x_e^2 \geq \|Q_e\|_2^2, x_e \geq 0 \text{ for all } e,$$

with variables $Q \in \mathbb{R}^{m \times k}$ and $x \in \mathbb{R}^m$. Let Q^* and x^* be an optimal solution. Then clearly $x_e^* = \|Q_e^*\|_2$ for all e . Using the Lagrange multipliers $P \in \mathbb{R}^{n \times k}$ for the equations $A Q = B$, and $\alpha \in \mathbb{R}_{\geq 0}^m$ and $\beta \in \mathbb{R}_{\geq 0}^m$ for the inequalities, the KKT conditions [16, Subsection 5.5] become

$$c_e - 2\alpha_e x_e^* - \beta_e = 0 \quad \text{for all } e, \quad (21)$$

$$P^T (A^T)_e + 2\alpha_e Q_e^* = 0 \quad \text{for all } e, \quad (22)$$

$$\alpha_e (x_e^*)^2 - \|Q_e^*\|_2^2 = 0 \quad \text{for all } e, \quad (23)$$

$$\beta_e x_e^* = 0 \quad \text{for all } e. \quad (24)$$

Here the first two conditions state that at the optimum, the gradient of the objective with respect to the variables x_e and $Q_{e,i}$ must be linear combinations of the gradients of the active constraints and the last two conditions are complementary slackness (= a Lagrange multiplier can only be non-zero if the constraint is tight). We also have the feasibility constraints

$$A Q^* = B, \quad (25)$$

$$x_e^* \geq 0 \text{ and } x_e^* \geq \|Q_e\|_2 \quad \text{for all } e. \quad (26)$$

Separating the two terms in (22), squaring and summing over i , and using (22) and (21), we obtain

$$\|(A^T)_e P\|_2^2 = \sum_i ((P^i)^T A_{e,i})^2 = 4\alpha_e^2 \|Q_e^*\|_2^2 = 4\alpha_e^2 (x_e^*)^2 = (c_e - \beta_e)^2 \leq c_e^2,$$

where the last inequality uses $\beta_e = 0$ if $x_e^* > 0$ by (24) and $\beta_e = c_e$ if $x_e^* = 0$ by (21).

If $Q_e^* \neq 0$, then $x_e^* \neq 0$ and hence $\beta_e = 0$ and $c_e = 2\alpha_e x_e^*$ or $2\alpha_e = c_e/x_e^*$. In particular, $\alpha_e \neq 0$ and hence (22) implies

$$Q_e^* = \frac{1}{2\alpha_e} \sum_v P_v A_{v,e} = \frac{x_e^*}{c_e} A_e^T P. \quad (27)$$

This equation also holds if $Q_e^* = 0$ and hence $x_e^* = 0$. Multiplying by $(A^T)_e$ from the left and summing over e yields

$$B = A Q^* = A X C^{-1} A^T P. \quad (28)$$

Thus P^i is a potential for the i -th problem with respect to x^* and, by (27) $(Q^*)^i$ is the corresponding electrical flow. Thus $x^* \in \mathcal{F}_g$ by Lemma 2. Moreover,

$$\begin{aligned} \sum_i (P^i)^T b^i &= \sum_i (P^i)^T A (Q^*)^i \\ &= \sum_{i,v,e} P_{v,i} A_{v,e} Q_{e,i}^* \\ &= \sum_{i,v,e, Q_e^* \neq 0} P_{v,i} A_{v,e} Q_{e,i}^* \\ &= \sum_{i,e, Q_e^* \neq 0} c_e \frac{Q_{e,i}^* Q_{e,i}^*}{\|Q_e^*\|_2} \\ &= \sum_e c_e \|Q_e^*\|_2. \end{aligned}$$

Here the fourth equality comes from (27) and $x_e^* \neq 0$ if $Q_e^* \neq 0$; note that

$$\sum_v A_{ve} P_{vi} = (A^T)_e P_i = \frac{c_e}{x_e^*} Q_{e,i}^* = c_e \frac{Q_{e,i}^*}{\|Q_e^*\|_2}.$$

We conclude that P is a feasible solution to (19). Thus $MaxP \geq MinQ$.

Since $x^* \in \mathcal{F}_g$, $\mathcal{L}(x^*) = c^T x^* = \text{Tr}[P^T L(x^*) P]$. Also, $x_e^* = \|Q_e^*\|_2$ by definition of x^* . Thus

$$\sum_e c_e \|Q_e^*\|_2 = c^T x^* = \mathcal{L}(x^*)$$

and hence $MinL \leq MinQ$. \square

Lemma 12. $MaxP \leq MinL$.

Proof. The constraint $\|(A^T)_e P\|_2 \leq c_e$ in (19) can be equivalently written as

$$\frac{c_e}{2} \left(\left\| \frac{1}{c_e} (A^T)_e P \right\|_2^2 - 1 \right) \leq 0.$$

Then the Lagrange dual with non-negative multipliers x_e is an upper bound for $MaxP$, i.e.,

$$MaxP \leq \inf_{x \geq 0} \sup_P \sum_i (b^i)^T P^i - \sum_e \frac{x_e c_e}{2} \left(\left\| \frac{1}{c_e} (A^T)_e P \right\|_2^2 - 1 \right).$$

The inner supremum can be reformulated as

$$\sup_P \sum_i (b^i)^T P^i - \frac{1}{2} \sum_i (P^i)^T L(x) P^i + \frac{1}{2} c^T x, \tag{29}$$

since $(x_e/c_e) \sum_i ((A^T)_e P^i)^2 = \sum_i (P^i)^T A^e (x_e/c_e) (A^T)_e P^i$. Only the first two terms in (29) depend on P . We want to determine the maximizer³ $P(x)$. Taking partial derivatives with respect to the vectors P^i leads to the system

$$A^T C^{-1} A \cdot P^i(x) = b^i \quad \text{for all } i,$$

i.e. $P^i(x)$ is a solution to $L(x)P^i(x) = b^i$ for each i . Since

$$\sum_i (b^i)^T P^i(x) = \text{Tr}[B^T P(x)] = \text{Tr}[P(x)^T L(x) P(x)] = \sum_i (P^i(x))^T L(x) P^i(x)$$

substituting into (29) yields

$$\sup_P \sum_i (b^i)^T P^i - \frac{1}{2} \sum_i (P^i)^T L(x) P^i + \frac{1}{2} c^T x = \frac{1}{2} \left(\text{Tr}[B^T P(x)] + c^T x \right) = \mathcal{L}(x). \quad \square$$

Lemma 13. Let $x^* \in \mathbb{R}_{\geq 0}^m$ be a minimizer of $\mathcal{L}(x)$. Then $x^* \in \mathcal{F}_g$. Let P be a solution to $L(x^*)P = B$ and let $Q = X^* C^{-1} A^T P$. Then $\sum_e c_e \|Q_e\|_2 = \mathcal{L}(x^*)$ and hence $MinQ \leq MinL$.

Proof. Since $\mathcal{L}(x(t))$ is a Lyapunov function of the generalized Physarum dynamics we have $x^* \in \mathcal{V}$. Since $\mathcal{V} = \mathcal{F}_g$, x^* is a fixed point and hence for all e , either $x_e^* = 0$ or $\|A_e\|_2 = 1$. Since x^* is a fixed point, we have $\mathcal{L}(x^*) = c^T x^* = \text{Tr}[P^T L(x^*) P]$ and $x_e^* = \|Q_e\|_2$ for all e . Thus

$$\sum_e c_e \|Q_e\|_2 = c^T x^* = \mathcal{L}(x^*)$$

and hence $MinQ \leq MinL$. \square

11. Convergence to the Lyapunov minimizer

We show that the dynamics converges to the minimizer x^* of the Lyapunov function under the assumption that the set of fixed points of the dynamics is a discrete set.

Assumption 1 (Discrete Set of Fixed Points). \mathcal{F}_g is a finite set of points. For any two points in \mathcal{F}_g , the values of \mathcal{L} are distinct.

Theorem 7. Let $x^* = \text{argmin}_{x \geq 0} \mathcal{L}(x)$. Under the additional Assumption 1, the generalized Physarum dynamics $x(t)$ converges to x^* .

³ In the proof of Lemma 5, we have seen that $L(x) = AD^{1/2}D^{1/2}A^T$ and hence $b_i^T P^i - (P^i)^T L(x) P^i = b_i^T P^i - \|D^{1/2}A^T P^i\|_2^2$. Thus the maximizer is a finite point.

Proof. Since $\mathcal{L}(x(t))$ is non-increasing and non-negative, the dynamics $x(t)$ converges to the set \mathcal{V} . By Theorem 5, $\mathcal{V} = \mathcal{F}_g$. Since \mathcal{F}_g is assumed to be a finite set and any two fixed points have distinct values of \mathcal{L} , there is a fixed point $\hat{x} = \lim_{t \rightarrow \infty} x(t)$. Assume for the sake of a contradiction, $\mathcal{L}(\hat{x}) > \mathcal{L}(x^*)$. Let $P(t)$ be the node potential corresponding to $x(t)$ and let \hat{P} be the potential corresponding to \hat{x} ; recall that node potentials are unique. Since $P(t)$ is a continuous function of $x(t)$, $P(t) \rightarrow \hat{P}$ as $t \rightarrow \infty$. Let $\hat{E} = \left\{ e : \|A_e^T \hat{P}\|_2 \leq c_e \right\} \subseteq E$ and consider the following chain of inequalities:

$$\begin{aligned} \max_P \left\{ \text{Tr}[B^T P] : \|A_e P\|_2 \leq c_e \text{ for all } e \in \hat{E} \right\} &\geq \text{Tr}[B^T \hat{P}] \\ &= \mathcal{L}(\hat{x}) \\ &> \mathcal{L}(x^*) \\ &= \max_P \left\{ \text{Tr}[B^T P] : \|A_e P\|_2 \leq c_e \text{ for all } e \in E \right\}, \end{aligned}$$

where the first inequality follows by the definition of \hat{E} , the first equality follows from Lemma 2, the strict inequality holds by assumption and the last equality follows from Theorem 6. We conclude that \hat{E} is a proper subset of E .

Let $e \in E \setminus \hat{E}$ be arbitrary. Then $\|(A^T)_e \hat{P}\|_2 > c_e$ and hence there are $t_0 > 0$ and $\varepsilon > 0$ such that for every $t \geq t_0$ we have

$$\|\Lambda_e(t)\|_2 = \frac{\|(A^T)_e P(t)\|_2}{c_e} > 1 + \varepsilon.$$

Since g_e is an increasing function with $g_e(1) = 1$, there is an $\alpha > 0$ such that for all $t \geq t_0$

$$g_e(\|\Lambda_e(t)\|_2) \geq g_e(1 + \varepsilon) = 1 + \alpha.$$

Then, for the generalized dynamics we have

$$\dot{x}_e(t) = x_e(t) \cdot (g_e(\|\Lambda_e(t)\|_2) - 1) \geq x_e(t) \cdot (g_e(1 + \varepsilon) - 1) \geq \alpha x_e(t).$$

Further, by Gronwall's Lemma, it follows that

$$x_e(t) \geq x_e(t_0) \cdot e^{\alpha t},$$

and thus

$$\hat{x}_e = \lim_{t \rightarrow \infty} x_e(t) \geq x_e(t_0) \cdot \lim_{t \rightarrow \infty} e^{\alpha t} = +\infty.$$

This is a contradiction to the fact that \hat{x}_e is bounded.

Finally, if $\mathcal{L}(x(t))$ converges to $\min_{x \geq 0} \mathcal{L}(x)$ and the minimizer x^* of \mathcal{L} is unique, then $x(t)$ must converge to x^* . \square

We conjecture that $x(t)$ always converges to some minimizer of \mathcal{L} . If there are several minimizers of \mathcal{L} , the limit depends on the initial configuration and the function g . Consider the following simple example. We have a network with two nodes connected by two links of the same cost, $k = 1$ and the goal is to send one unit between the two nodes. Let x_1 and x_2 be the capacities of the two links, respectively. For $g(z) = z$, any combination (x_1, x_2) with $x_1 + x_2 = 1$ is a fixed point.

12. A connection to mirror descent

We show that the *mirror descent* dynamics on the Lyapunov function \mathcal{L} is equal to a variant of the non-uniform squared Physarum dynamics.

Lemma 14. *The dynamics*

$$\frac{d}{dt} x_e(t) = \frac{c_e}{2} x_e(t) \left(\|\Lambda_e\|_2^2 - 1 \right)$$

is equivalent to the mirror descent dynamics on the Lyapunov function \mathcal{L} .

Proof. By Lemma 10, we have for every index $e \in E$ that

$$\frac{\partial}{\partial x_e} \mathcal{L}(x) = \frac{c_e}{2} (1 - \|\Lambda_e\|_2^2). \tag{30}$$

On the other hand, the mirror descent dynamics on the Lyapunov function \mathcal{L} is given by

$$\frac{d}{dt} x_e(t) = -x_e(t) \frac{\partial}{\partial x_e} \mathcal{L}(x_e(t)) \stackrel{(30)}{=} \frac{c_e}{2} \cdot x_e(t) (\|\Lambda_e\|_2^2 - 1). \quad \square$$

As is [14], we can use the connection to mirror descent to estimate the speed of convergence of the Physarum dynamics to the Lyapunov minimum; [14] builds up on [3,44].

For a differentiable function f in m variables, the Bregman divergence D_f is a function in $2m$ variables defined by the equation

$$D_f(x, y) = f(x) - f(y) - \langle \nabla f(y), x - y \rangle,$$

i.e., as the difference of the function value at x and the value at x of the tangent plane to f at y . Clearly, if f is convex, D_f is non-negative.

Lemma 15. Let $h : \mathbb{R}_{\geq 0}^m \rightarrow \mathbb{R}$ be defined by

$$h(x) = \sum_e x_e \ln x_e - \sum_e x_e.$$

Then h is convex on $\mathbb{R}_{\geq 0}^m$, D_h is non-negative, and

$$D_h(x, y) = \sum_e x_e \ln x_e - \sum_e x_e \ln y_e - \sum_e x_e + \sum_e y_e.$$

Proof. The function h is convex in x_e (partial derivative $\ln x_e$ and second partial derivative $1/x_e$). For its Bregman divergence D_h , we compute

$$\begin{aligned} D_h(x, y) &= h(x) - h(y) - \langle \nabla h(y), x - y \rangle \\ &= \sum_e x_e \ln x_e - \sum_e x_e - \left(\sum_e y_e \ln y_e - \sum_e y_e \right) - \sum_e (x_e - y_e) \ln y_e \\ &= \sum_e x_e \ln x_e - \sum_e x_e \ln y_e - \sum_e x_e + \sum_e y_e. \end{aligned}$$

So D_h is the relative entropy function. \square

Fact 1. [14, Lemma 2.2] \mathcal{L} is convex.

Theorem 8. Let x^* be the global minimizer of $\mathcal{L}(x)$. For the dynamics $\dot{x}_e = (c_e/2) \cdot x_e (\|\Lambda_e\|_2^2 - 1)$, we have

$$\mathcal{L}(x(t)) \leq \mathcal{L}(x^*) + \frac{1}{t} D_h(x^*, x(0)),$$

for all $t \geq 0$. In particular,

$$\lim_{t \rightarrow \infty} \mathcal{L}(x(t)) = \mathcal{L}(x^*).$$

Proof. According to (30) we have

$$\frac{\partial}{\partial x_e} \mathcal{L}(x) = \frac{c_e}{2} (1 - \|\Lambda_e\|_2^2) \text{ and } \dot{x}_e = x_e (g_e(\|\Lambda_e\|_2^2) - 1).$$

The time derivative of $D_h(x^*, x(t))$ is given by

$$\begin{aligned} \frac{d}{dt} D_h(x^*, x) &= \frac{d}{dt} \sum_{e=1}^m x_e^* \ln x_e^* - \frac{d}{dt} \sum_e x_e^* \ln x_e - \frac{d}{dt} \sum_e x_e^* + \frac{d}{dt} \sum_e x_e \\ &= \sum_{e=1}^m x_e^* \left(-\frac{1}{x_e} \cdot \frac{d}{dt} x_e \right) + \sum_e \frac{d}{dt} x_e \\ &= \sum_e (x_e - x_e^*) \frac{c_e}{2} (\|\Lambda_e\|_2^2 - 1) \\ &= -\langle (x - x^*), \nabla \mathcal{L}(x(t)) \rangle. \end{aligned}$$

We now consider the function

$$\mathcal{H}(t) = D_h(x^*, x(t)) + t [\mathcal{L}(x(t)) - \mathcal{L}(x^*)]$$

Since $\frac{d}{dt}\mathcal{L}(x) \leq 0$, by Lemma 5, and $D_{\mathcal{L}}(x^*, x) \geq 0$ for all x , we obtain

$$\begin{aligned} \frac{d}{dt}\mathcal{H}(t) &= -\langle \nabla \mathcal{L}(x(t)), x(t) - x^* \rangle + \mathcal{L}(x(t)) - \mathcal{L}(x^*) + t \cdot \frac{d}{dt}\mathcal{L}(x(t)) \\ &\leq -[\mathcal{L}(x^*) - \mathcal{L}(x(t)) - \langle \nabla \mathcal{L}(x), x^* - x(t) \rangle] \\ &= -D_{\mathcal{L}}(x^*, x(t)) \\ &\leq 0. \end{aligned}$$

Hence $\mathcal{H}(t) \leq \mathcal{H}(0)$ for all $t \geq 0$ and therefore

$$D_h(x^*, x(t)) + t[\mathcal{L}(x(t)) - \mathcal{L}(x^*)] \leq D_h(x^*, x(0)) + 0[\mathcal{L}(x(0)) - \mathcal{L}(x^*)],$$

and further (using $D_h(x^*, x(t)) \geq 0$)

$$\mathcal{L}(x(t)) \leq \mathcal{L}(x^*) + \frac{1}{t}D_h(x^*, x(0)). \quad \square$$

13. Conclusions

We proposed a variant of the Physarum dynamics suitable for network design. We exhibited a Lyapunov function for the dynamics, proved convergence of the dynamics, and gave alternative characterizations for the minimum of the Lyapunov function. In the experimental part, we showed that the dynamics captures the positive effect of sharing links and is able to construct *nice* networks.

Many questions remain open. We do not claim any biological plausibility for our proposal and we have studied one particular form of the dynamics, namely $\dot{x}_e = |q_e| - x_e$. Other dynamics have been studied for the shortest path problem, e.g., $\dot{x}_e = |q_e|^\mu - x_e$ with $\mu > 1$ or $\dot{x}_e = \frac{|q_e|}{a_e + |q_e|} - x_e$ [32,29]. The latter paper also studies the influence of noise on the dynamics. An extension to network design would be interesting.

The papers mentioned in the preceding paragraph are theory papers that investigate variants of the basic dynamics (1). A different line of research aims at a deeper understanding of the inner workings of Physarum polycephalum, for example, how global synchronization can result from random peristaltics [1], how information can be transported and a memory can exist in an organism without a nervous system [2,25], and whether tubes of the mold can transfer electricity [43]. There seems to be little connection between these lines of research.

We used an Euler discretization of the dynamics for the experiments in Section 5. The resulting algorithm is quite slow. The Lyapunov function \mathcal{L} is a convex function and hence the tool box of convex optimization is available for computing its minimum. Does this lead to a practical algorithm for network design? [45] also uses an Euler discretization of the dynamics for their computer experiments. They speed-up the computation by considering only a random subset of the demands instead of all demands in each iteration. If the random subset is not too small, the dynamics seem to converge to the same solution. Is this true generally?

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- [1] K. Alim, G. Amselem, F. Peaudecerf, M.P. Brenner, A. Pringle, Random network peristalsis in Physarum polycephalum organizes fluid flows across an individual, *Proc. Natl. Acad. Sci.* 110 (2013) 13306–13311.
- [2] K. Alim, N. Andrew, A. Pringle, M.P. Brenner, Mechanism of signal propagation in Physarum polycephalum, *Proc. Natl. Acad. Sci.* 114 (2017) 5136–5141.
- [3] F. Alvarez, J. Bolte, O. Brachic, Hessian Riemannian gradient flows in convex programming, *SIAM J. Control Optim.* 43 (2) (2004) 477–501.
- [4] Andrew Adamatzky, Bernard De Baets, Wesley Van Dessel, Slime mould imitation of Belgian transport networks: redundancy, bio-essential motorways, and dissolution, *Int. J. Unconv. Comput.* (2012).
- [5] A. Adamatzky, Developing proximity graphs by Physarum polycephalum: does the plasmodium follow the Toussaint hierarchy?, in: *Parallel Process. Lett.*, 2009, pp. 105–127.
- [6] Andrew Adamatzky, *Physarum Machines: Computers from Slime Mold*, World Scientific Publishing, 2010.
- [7] Andrew Adamatzky, Thirty eight things to do with live slime mould, CoRR, arXiv:1512.08230 [abs], 2015.
- [8] Charu C. Aggarwal, *Neural Networks and Deep Learning*, Springer, 2018.
- [9] Andrew Adamatzky, Genaro Juárez Martínez, Sergio V. Chapa Vergara, René Asomoza-Palacio, Christopher R. Stephens, Approximating Mexican highways with slime mould, *Nat. Comput.* 10 (3) (2011) 1195–1214.
- [10] Ruben Becker, Vincenzo Bonifaci, Andreas Karrenbauer, Pavel Kolev, Kurt Mehlhorn, Two results on slime mold computations, *Theor. Comput. Sci.* 773 (2019) 79–106.
- [11] S.L. Baldauf, W.F. Doolittle, Origin and evolution of the slime molds (Mycetozoa), *Proc. Natl. Acad. Sci. USA* (1997) 12007–12012.
- [12] Vincenzo Bonifaci, Kurt Mehlhorn, Girish Varma, Physarum can compute shortest paths, *J. Theor. Biol.* 309 (2012) 121–133, A preliminary version of this paper appeared at SODA 2012 (pages 233–240).

- [13] Bonifaci Vincenzo, A revised model of fluid transport optimization in Physarum polycephalum, *J. Math. Biol.* 74 (2016) 567–581.
- [14] Bonifaci Vincenzo, A Laplacian approach to ℓ_1 -norm minimization, *Comput. Optim. Appl.* 79 (2021) 441–469.
- [15] Olivier Bournez, Amaury Pouly, A survey on analog models of computation, *CoRR*, arXiv:1805.05729 [abs], 2018.
- [16] Stephen Boyd, Lieven Vandenbergh, *Convex Optimization*, Cambridge University Press, 2004.
- [17] Marco Dorigo, Thomas Stützle, *Ant Colony Optimization*, MIT Press, 2004.
- [18] E. Facca, F. Cardin, M. Putti, Physarum dynamics and optimal transport for basis pursuit, arXiv:1812.11782 [math.NA], December 2018.
- [19] Enrico Facca, Sara Daneri, Franco Cardin, Mario Putti, Numerical solution of Monge-Kantorovich equations via a dynamic formulation, *J. Sci. Comput.* 82 (3) (2020) 68.
- [20] P. Hartman, *Ordinary Differential Equations*, second edition, SIAM, 2002.
- [21] Normal Haughness, *Grandest Illusion: The Seductive Myth of Free-Will*, Echo Park Press, 2006.
- [22] C. Horsman, S. Stepney, R.C. Wagner, V. Kendon, When does a physical system compute?, *Proc. R. Soc. A* 470 (2014).
- [23] Kentaro Ito, Anders Johansson, Toshiyuki Nakagaki, Atsushi Tero, Convergence properties for the Physarum solver, arXiv:1101.5249v1, January 2011.
- [24] Jeff Jones, Andrew Adamatzky, Towards Physarum binary adders, *Biosystems* 101 (2010) 51–58.
- [25] M. Kramar, K. Alim, Encoding memory in tube diameter hierarchy of living flow network, *Proc. Natl. Acad. Sci.* (2021).
- [26] S. Kirkpatrick, C.D. Gelatt Jr, M.P. Vecchi, Optimization by simulated annealing, *Science* 220 (1983) 671–680.
- [27] Andreas Karrenbauer, Pavel Kolev, Kurt Mehlhorn, Convergence of the non-uniform Physarum dynamics, *Theor. Comput. Sci.* 816 (2020) 260–269.
- [28] Alessandro Lonardi, Enrico Facca, Mario Putti, Caterina De Bacco, Optimal transport for multi-commodity routing on networks, arXiv:2010.14377, 2020.
- [29] B. Meyer, C. Ansorge, T. Nakagaki, The role of noise in self-organized decision making by the true slime mold Physarum polycephalum, *PLoS ONE* 12 (e0172933) (2017).
- [30] R. Mayne, When the path is never shortest: a reality check on shortest path biocomputation, in: A. Adamatzky (Ed.), *Shortest Path Solvers*. From Software to Wetware, Springer, 2018, pp. 379–399.
- [31] T. Miyaji, Isamu Ohnishi, Physarum can solve the shortest path problem on Riemannian surface mathematically rigorously, *Int. J. Pure Appl. Math.* 47 (2008) 353–369.
- [32] T. Miyaji, I. Ohnishi, A. Tero, T. Nakagaki, Failure to the shortest path decision of an adaptive transport network with double edges in Plasmodium system, *Int. J. Dyn. Syst. Differ. Equ.* 1 (3) (2008) 210–219.
- [33] Saket Navlakha, Ziv Bar-Joseph, Algorithms in nature: the convergence of systems biology and computational thinking, *Mol. Syst. Biol.* 7 (546) (2011).
- [34] Saket Navlakha, Ziv Bar-Joseph, Distributed information processing in biological and computational systems, *Commun. ACM* 58 (1) (2015) 94–102.
- [35] T. Nakagaki, M. Iima, T. Ueda, Y. Nishiura, T. Saigusa, A. Tero, R. Kobayashi, K. Showalter, Minimum-risk path finding by an adaptive amoebal network, *Phys. Rev. Lett.* 99 (068104) (2007) 4.
- [36] T. Nakagaki, H. Yamada, Á. Tóth, Maze-solving by an amoeboid organism, *Nature* 407 (2000) 470.
- [37] Tomohiro Shirakawa, Andrew Adamatzky, Yukio-Pegio Gunji, Yoshihiro Miyake, On simultaneous construction of Voronoi diagram and Delaunay triangulation by Physarum polycephalum, *Int. J. Bifurc. Chaos* 19 (2009) 3109–3117.
- [38] S. Stepney, The neglected pillar of material computation, *Phys. D, Nonlinear Phenom.* 237 (2008) 1157–1164.
- [39] Damian Straszak, Nisheeth K. Vishnoi, IRLS and slime mold: equivalence and convergence, *CoRR*, arXiv:1601.02712 [abs], 2016.
- [40] Damian Straszak, Nisheeth K. Vishnoi, On a natural dynamics for linear programming, in: *ITCS*, ACM, New York, NY, USA, 2016, p. 291.
- [41] A. Tero, R. Kobayashi, T. Nakagaki, A mathematical model for adaptive transport network in path finding by true slime mold, *J. Theor. Biol.* (2007) 553–564.
- [42] A. Tero, S. Takagi, T. Saigusa, K. Ito, D. Bebbler, M. Fricker, K. Yumiki, R. Kobayashi, T. Nakagaki, Rules for biologically inspired adaptive network design, *Science* 327 (2010) 439–442.
- [43] James G.H. Whiting, Ben P.J. de Lacy Costello, Andrew Adamatzky, Transfer function of protoplasmic tubes of Physarum polycephalum, *Biosystems* 128 (2015) 48–51.
- [44] A. Wilson, Lyapunov arguments in optimization, PhD thesis, University of California at Berkeley, 2018.
- [45] Shin Watanabe, Atsushi Tero, Atsuko Takamatsu, Toshiyuki Nakagaki, Traffic optimization in railroad networks using an algorithm mimicking an amoeba-like organism, *Physarum plasmodium*, *Biosystems* 105 (3) (2011) 225–232.
- [46] Hanchao Yang, Richard Mayne, Yong Deng, A bio-inspired network design method for intelligent transportation, *Int. J. Unconv. Comput.* 14 (3–4) (2019) 199–215.
- [47] Xiaoge Zhang, Andrew Adamatzky, Felix T.S. Chan, Sankaran Mahadevan, Yong Deng, Physarum solver: a bio-inspired method for sustainable supply chain network design problem, *Ann. Oper. Res.* 254 (1–2) (2017) 533–552.

CHAPTER III

NOISE-INDUCED NETWORK TOPOLOGIES

Noise-induced network topologies

Phys. Rev. Lett. 130, 267401 – Published 30 June 2023

© 2023 American Physical Society

DOI: 10.1103/PhysRevLett.130.267401

Authors: Frederic Folz¹, Kurt Mehlhorn², Giovanna Morigi¹

¹ *Theoretische Physik, Universität des Saarlandes, 66123 Saarbrücken, Germany*

² *Algorithms and Complexity Group, Max-Planck-Institut für Informatik, Saarland Informatics Campus, 66123 Saarbrücken, Germany*

Abstract:

We analyze transport on a graph with multiple constraints and where the weight of the edges connecting the nodes is a dynamical variable. The network dynamics results from the interplay between a nonlinear function of the flow, dissipation, and Gaussian, additive noise. For a given set of parameters and finite noise amplitudes, the network self-organizes into one of several metastable configurations, according to a probability distribution that depends on the noise amplitude α . At a finite value α , we find a resonantlike behavior for which one network topology is the most probable stationary state. This specific topology maximizes the robustness and transport efficiency, it is reached with the maximal convergence rate, and it is not found by the noiseless dynamics. We argue that this behavior is a manifestation of noise-induced resonances in network self-organization. Our findings show that stochastic dynamics can boost transport on a nonlinear network and, further, suggest a change of paradigm about the role of noise in optimization algorithms.

Noise-Induced Network Topologies

Frederic Folz¹, Kurt Mehlhorn², and Giovanna Morigi¹

¹*Theoretische Physik, Universität des Saarlandes, 66123 Saarbrücken, Germany*

²*Algorithms and Complexity Group, Max-Planck-Institut für Informatik, Saarland Informatics Campus, 66123 Saarbrücken, Germany*

(Received 8 August 2022; revised 19 April 2023; accepted 23 May 2023; published 30 June 2023)

We analyze transport on a graph with multiple constraints and where the weight of the edges connecting the nodes is a dynamical variable. The network dynamics results from the interplay between a nonlinear function of the flow, dissipation, and Gaussian, additive noise. For a given set of parameters and finite noise amplitudes, the network self-organizes into one of several metastable configurations, according to a probability distribution that depends on the noise amplitude α . At a finite value α , we find a resonantlike behavior for which one network topology is the most probable stationary state. This specific topology maximizes the robustness and transport efficiency, it is reached with the maximal convergence rate, and it is not found by the noiseless dynamics. We argue that this behavior is a manifestation of noise-induced resonances in network self-organization. Our findings show that stochastic dynamics can boost transport on a nonlinear network and, further, suggest a change of paradigm about the role of noise in optimization algorithms.

DOI: [10.1103/PhysRevLett.130.267401](https://doi.org/10.1103/PhysRevLett.130.267401)

The ability to extract information from large databases has become essential to modern science and technologies. This quest is central to foundational studies, such as in astronomy, for shedding light on the constitution of our Universe [1], and in particle physics, for efficiently identifying relevant events in high-energy physics experiments [2], as well as to applications, such as the design of efficient power grids [3] and the sustainable exploitation of water supplies [4]. A question lying at the core of these efforts is, what are the key ingredients and dynamics at the basis of an efficient search in a generic database? This question encompasses a large number of physically relevant situations, including the determination of the ground state of a quantum many-body problem [5–7], the transport of excitons [8,9] and cells [10,11], and the search for food by living organisms [12,13]. The latter is a precious source of insights because of organisms' capability to extract information from and adapt to a dynamically changing environment [12,14]. One example is the food search of *Physarum polycephalum* and of ant colonies, that have inspired optimization algorithms successfully applied to real-world optimization problems [12,13,15–17].

One relevant aspect of biological systems is the capability to efficiently extract relevant information for their survival in a noisy environment, where parameters fluctuate and the amount and location of food sources can change over time. For instance, models simulating excitable systems, such as forest fires [18] and neurons [19], show that noise can lead to qualitatively different effects. These include phenomena such as stochastic and coherence resonance [19–21], synchronization [22,23], and

noise-induced phase transitions [24,25]. A systematic understanding of the role of noise in a search problem would shed light on its role in cooperative dynamics, including neural networks, and might initiate novel applications to optimization problems.

In this work, we analyze the self-organization dynamics of a network in the presence of additive noise and with multiple constraints to be satisfied. The constraints are two pairs of source and sink nodes, as illustrated in Fig. 1(a), at which a constant flow is injected and extracted, respectively. In computer science, it is a multicommodity problem: each pair of source and sink is a *demand* to be satisfied and the path satisfying the demand is a *flow of commodity* [26,27]. Examples are a city transport network, where each commodity is the passengers traveling between two stations, or an electrical circuit, where the commodity is the electrical current satisfying a given potential difference between two nodes. The optimal path is a network topology obtained by integrating a set of equations for the graph's nodes and edges, where the strength of the edges, determining the edge capacity [26], is a dynamical variable subject to the competition between dissipation and an activation force depending on the total flow across the edge [12,15,26,27]. In the absence of noise, the dynamics tends to identify the optimal path satisfying the constraints according to a rule that promotes transport along shared routes and instead inhibits it when the flow along one edge is below a chosen threshold. Differing from the typical settings, in this work, we assume that the edge capacity can also fluctuate due to a Langevin force [28]. We show that the introduction of stochasticity has a dramatic impact on

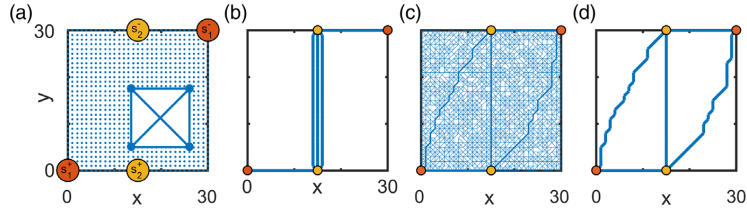


FIG. 1. (a) Network self-organization is simulated on a grid of 31×31 nodes with two demands. The demands are indicated by the pairs of red and yellow nodes, the sources are labeled by s_1^+ , the sinks by s_1^- ; the inset shows that the nodes are connected by horizontal, vertical, and diagonal edges. The network design results from the dynamics of the edges, which are modeled by time-varying conductivity $D_{u,v}$ on an electrical network and in the presence of additive noise according to Eqs. (1) and (2). Panels (b) and (c) display the networks reached after a sufficiently long integration time in the noiseless case ($\alpha = 0$) and for $\alpha = 0.002$, respectively. The widths of the edges are proportional to the corresponding amplitude of $D_{u,v}$. Panel (d) displays the multiscale backbone extracted from (c) using a filtering procedure (see text). See Fig. 2 for details on the numerical simulations.

the convergence to the optimal path. Among several noteworthy features, the solutions follow a multistable distribution that undergoes discontinuous transitions as a function of the noise amplitude. Remarkably, the distribution exhibits a resonant type of behavior as a function of the noise strength. In fact, for a finite range of noise amplitudes the network self-organizes into a topology that maximizes its robustness and that is not found by the noiseless dynamics.

Model.—In the following, we will refer to the multicommodity problem in terms of currents in an electrical circuit, keeping in mind that this is just one possible example. The edge capacity is determined by the conductivity, which is a dynamical variable. The circuit consists of a spatial grid composed of 31×31 nodes. Each node, labeled u , can connect to a number of nearest and next-nearest neighbors, described by the set E_u [see the inset of Fig. 1(a)]. The emerging networks need to serve two demands $i = 1, 2$, each represented by a source node s_1^+ and a sink node s_1^- , where a current is injected ($+I_i$) and extracted ($-I_i$), respectively. Each demand generates a flow across the network: The flow of the demand i is composed of the contributions $Q_{u,v}^i$ at the edge connecting nodes (u, v) . The flow of each demand is conserved at each node u , $\sum_{v \in E_u} Q_{u,v}^i = 0$ (Kirchhoff's law) except for the source and sink where $\sum_{v \in E_{s_1^\pm}} Q_{s_1^\pm, v}^i = \pm I_i$. The flow of the demand i along the edge (u, v) is proportional to the edge conductivity $D_{u,v}(t)$ and to the difference between the potentials of the two nodes $p_u^i(t)$ and $p_v^i(t)$:

$$Q_{u,v}^i(t) = \frac{D_{u,v}(t)}{L_{u,v}} [p_u^i(t) - p_v^i(t)], \quad (1)$$

where $L_{u,v}$ is the edge length and is constant. The edge dynamics is described by the coupled dynamical variables p_u^i and $D_{u,v}$. The potential p_u^i is determined for each demand i as a function of $D_{u,v}(t)$ by solving the linear set of equations in Eq. (1) with Kirchhoff's law, as detailed in

Ref. [26] and in the Supplemental Material (SM) [29]. The conductivity $D_{u,v}(t)$ obeys the stochastic nonlinear equation [30,31]:

$$\partial_t D_{u,v} = f(Q_{u,v}) - \gamma D_{u,v} + \sqrt{2\gamma\alpha} \xi_{u,v}(t). \quad (2)$$

Here, $f(x)$ is the activation function with sigmoidal form: $f(x) = x^n / (\kappa^n + x^n)$ with $n > 0$ (in what follows we choose $n = 1.2$), the argument is the total flow along the edge, $Q_{u,v} = \sum_i |Q_{u,v}^i|$, and f saturates when $Q_{u,v}$ exceeds the threshold κ . Hence, $f(x)$ gives rise to an effective interaction between demands that favors the sharing of transport routes between commodities. The activation is counteracted by dissipation at rate γ . Fluctuations in the conductivity are simulated by the stochastic force $\xi(t)$, whose amplitude is scaled by the parameter α . The force is statistically defined by the average over an ensemble of trajectories: it has no net drift, $\langle \xi_{u,v}(t) \rangle = 0$, and simulates Gaussian white noise, $\langle \xi_{u,v}(t) \xi_{u',v'}(t') \rangle = \delta_{u,u'} \delta_{v,v'} \delta(t-t')$ [28,32].

Our model shares analogies with resistor networks [33] but is essentially different in that the edge conductivities (the metric) are dynamical variables. Equations (1) and (2), in the absence of noise, were used in Ref. [15] for modeling the structures built by a unicellular organism for food search in a maze [34] and on a graph simulating the Tokyo railroad system [35]. These equations set the basis for optimization algorithms [12] and have been applied to multicommodity problems [26,27] using other classes of activation functions than the sigmoidal functions. The studies of Refs. [26,27] showed that the dynamics converges toward networks optimizing between the sharing of transport routes, favored by the activation function, and the total cost of the network (here given by the total length of the edges of the closed paths) that is controlled by dissipation. In Refs. [30,31], stochastic forces were added to the model for one single demand connected by two paths of the same length but different, periodically varying, dissipation rates. In Ref. [31], the resulting flow was analyzed as a function of the frequency of the dissipation

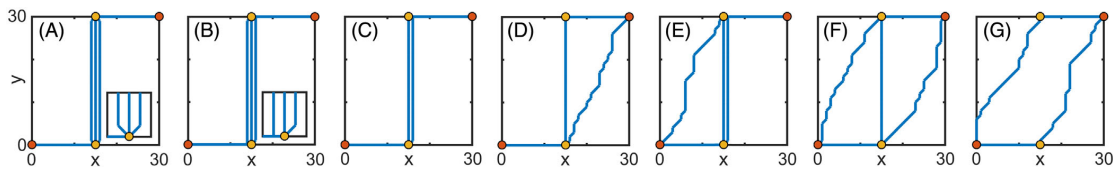


FIG. 2. Network topologies for increasing values of the noise amplitude α (from $\alpha = 0$ to $\alpha = 0.005$). (A) is the noiseless case, (B)–(G) are the typical backbones for $\alpha > 0$; the probability of their occurrence depends on α and is shown in Fig. 3 [for (D) and (E) we report one of the two symmetric configurations]. The networks are the result of the time evolution of Eqs. (1) and (2) for a time $t = 250\gamma^{-1}$ imposing $I_1 = I_2 = 0.45$ and $\kappa = 1$. Initially, we set $D_{u,v} = 0.5$ on all edges. The integration of Eq. (2) is performed using the Euler-Maruyama scheme [37] with step size $\Delta t = 0.1\gamma^{-1}$. In the SM movies are reported which show how the dynamics at different noise amplitudes leads to each of the topologies [29].

rates and amplitude of the noise, manifesting the characteristic features of stochastic resonance and noise-induced limit cycles. In this work, we analyze, for the first time, a multicommodity problem in the presence of noise. The relatively simple geometry of our problem allows us to single out the essential features and visualize the manifold of topologies as a function of the noise amplitude.

Results.—We integrate Eqs. (1) and (2) with the static boundary conditions of Fig. 1(a) after initializing the conductivities on all edges to the same value (see also SM [29]). The system evolution thus initially consists of redirecting the flow along edges by modifying the conductivities. For $\alpha = 0$, the dynamics is noiseless and converges to the configuration of Fig. 1(b): the flow satisfying both demands is routed along the vertical connection. The system tends to generate parallel routes. In fact, the transport along one edge is bound to a maximal value due to the saturation of the sigmoidal function. For $\alpha > 0$, we integrate stochastic differential equations. Figure 1(c) displays a network configuration obtained by integrating the stochastic dynamics for one trajectory and after a sufficiently long simulation time. It is evident that noise leads to a fluctuating distribution of weak connections. In order to be able to perform a classification, we apply a filter mechanism to each trajectory as follows. We level out the fluctuations by taking the time average of the configurations in the regime where the simulation has converged. We then account for the statistical relevance of the links by means of the disparity filter of Ref. [36] (see SM [29]). Figure 1(d) displays the network topology extracted from Fig. 1(c) after applying the disparity filter to the time-averaged configuration. For each value of α , we evaluate 5000 trajectories.

Figure 2 shows the typical network topologies ordered by increasing noise amplitude, starting from the noiseless case (A). Each is unique in terms of connectivity of the hubs and is characterized by a different set of values of the measures we apply, as we detail later. The networks (B) and (C) are found for small $\alpha > 0$ and are similar to the noiseless case with the tendency to decrease the shared routes. In addition, (C) decreases the number of connections. Configurations (B)–(E) are multistable and generally break the point

symmetry of the configuration. For larger values of α , the topologies converge to one of the two configurations (F) and (G), with a bistable region about $\alpha \sim 3 \times 10^{-3}$. Topologies (F) and (G) are point symmetric but qualitatively different from (A). Note that (A)–(G) are fixed points of the noiseless dynamics. Noise dramatically modifies the respective basin of attraction as visible by analyzing the network measures as a function of α .

The network measures are determined on the backbone of each trajectory. (i) The robustness r provides information on the quality of the connections: it increases by adding paths connecting two nodes, which in turn makes the network more robust against edge failures. It is defined by $r = 1/(\sum_{i=1}^2 R_i/2)$, with $R_i = (p_{s_+^i}^i - p_{s_-^i}^i)/I_i$ as the effective resistance between the source node s_+^i and the sink node s_-^i of each demand i ; see Ref. [38] and the SM [29]. (ii) The transport efficiency σ is given by $1/\sigma = \sum_{i=1}^2 d_i/2$, where d_i is the length of the shortest path connecting s_+^i and s_-^i [35]. (iii) Finally, the cost of the network c is the total length, found by summing over the ensemble \mathcal{E} of segments $L_{u,v}$ of the backbone where the conductivity is nonzero [35], $c = \sum_{(u,v) \in \mathcal{E}} L_{u,v}$. The measures (r, σ, c) are displayed in Figs. 3(a)–3(c) as a function of the noise amplitude α . The white lines indicate their mean values. The slope of the mean robustness and cost at $\alpha = 0$ is negative, showing that—on average—for small noise amplitudes the dynamics converges to topologies with worse robustness and lower cost than for the noiseless case. After this transient, they all reach a maximum for an interval of noise amplitudes centered about $\alpha \sim 2 \times 10^{-3}$ that is qualitatively above the noiseless value. For each value of α the distribution of $x = r, \sigma, c$ about the mean is encoded in the color scale. The distribution is clustered about the topologies of Fig. 2 with probabilities depending on α . One striking feature is that (A) disappears for $\alpha > 0$, indicating that it is unstable against fluctuations. As α is increased, the system jumps to different configurations, undergoing discontinuous, noise-induced transitions. The topologies (B)–(E) occur at low, nonvanishing values of α and are generally multistable. Remarkably, for a nonzero interval of values α (in the range 0.001–0.003) the

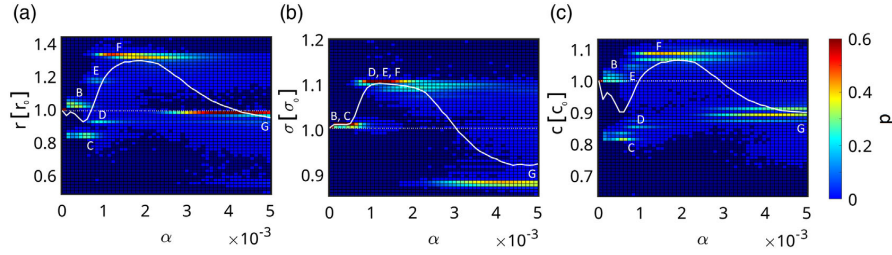


FIG. 3. Network measures as a function of the noise amplitude α : (a) robustness r , (b) transport efficiency σ , and (c) network cost c . Each measure is in units of the respective value r_0 , σ_0 , and c_0 for $\alpha = 0$ (dashed line in the plot). The white solid line is the mean value taken over 5000 trajectories at each value of α , the color scale gives the fraction of trajectories for each value of r , σ , c : dark blue is statistically irrelevant, dark red corresponds to 60%. The distribution clusters about a set of the topologies (the labels follow the legend of Fig. 2) and undergoes discontinuous transitions as α is varied. For $\alpha \in [0.001-0.003]$, it narrows about a single topology (F) with optimal robustness and transport efficiency.

distribution narrows and becomes single peaked and the dynamics converges to (F). This topology optimizes both robustness and transport efficiency, with a qualitative improvement over (A). At even larger amplitudes α , first (F) coexists with (G), then (G) becomes the most probable configuration. Network (G) has the same robustness as (A).

Its worse transport efficiency and lower cost are due to noise: the number of statistically relevant edges decreases with α . The distribution about (G) is broader according to the common expectation that noise increases the variance. Instead, the narrowing at $\alpha \in [0.001-0.003]$ about the topology (F) contradicts this intuition.

The trajectories converge relatively fast toward one of the topologies of Fig. 2. Figure 4(a) displays the average convergence rates to a stationary value of r , σ , c as a function of α . The rates are not monotonous functions of α and exhibit a local maximum corresponding to the network topology (F). In this regime the corresponding variances, Fig. 4(b), are minimal. This behavior provides further evidence that noise substantially modifies the basin of attraction of the individual topologies. The faster convergence rate to the topology (F) at $\alpha \in [0.001-0.003]$, together with the corresponding narrowing of the distribution of trajectories visible in Fig. 3, supports the conjecture that network self-organization into the topology (F) is a noise-induced resonance [19]. We have verified that this behavior also occurs (i) for a relatively wide range of the input and output flows, (ii) for different exponents n of the activation function, and (iii) for a substantially larger number of demands. In general, increasing the flow leads to a larger number of redundant connections. Instead, increasing the value of the exponent n in the activation function f enforces the use of shortest-path connections. Interestingly, we find noise-induced phenomena for all considered values of these parameters. This also holds true when analyzing larger networks, with respect to both the grid size and the pairs of source and sink nodes, i.e., of demands (see SM [29]). An extensive characterization will be reported in Ref. [39].

Discussion.—The noiseless equations at the basis of this study were developed in Ref. [15] for describing the food search of a slime mold [40,41]. From the biological point of view, this model is oversimplified (it discards key features such as the oscillatory flow through the tubes [42,43]), yet it qualitatively reproduces the patterns observed in

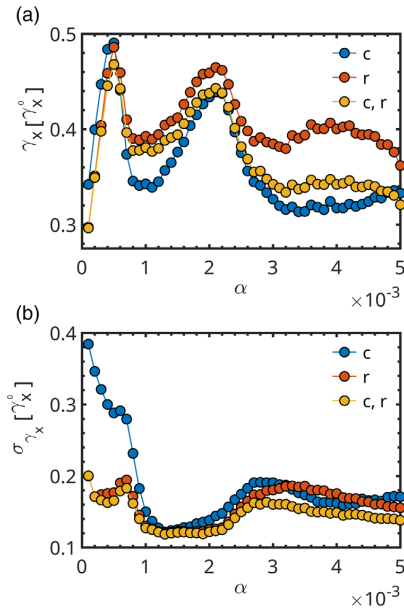


FIG. 4. (a) Average convergence rate γ_x as a function of the noise amplitude α (in units of the respective value γ_x^0 for the noiseless case). γ_x is the inverse of the time that a trajectory needs to reach a stationary value of the cost (blue), of the robustness (red), and of joint cost and robustness (yellow); see SM for the definition [29]. Panel (b) displays the corresponding variance σ_{γ_x} (in units of γ_x^0). The averages are taken over an ensemble of 5000 trajectories. About $\alpha \sim 0.002$, the dynamics converges to (F).

Refs. [34,35]. Moreover, it provides a powerful framework for network design and optimization algorithms [12,16]. Our work shows that the addition of noise to this model provides a qualitative improvement of the algorithmic efficiency by means of noise-induced resonances. This is a change of paradigm with regard to simulated annealing and randomized algorithms [44,45] and calls for a theoretical framework for stochastic nonlinear network dynamics [46,47].

The authors are grateful to Malte Henkel and Reza Shaebani for inspiring discussions and to Ginestra Bianconi for helpful comments. G. M. and F. F. acknowledge support from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) Project-ID No. 429529648, TRR 306 QuCoLiMa (Quantum Cooperativity of Light and Matter), and from the Bundesministerium für Bildung und Forschung (BMBF, German Ministry of Education and Research) under the grant “NiQ: Noise in Quantum Algorithms.” Financial support was also provided by the DFG Priority Program No. 1929 “GiRyd”.

-
- [1] S. Sen, S. Agarwal, P. Chakraborty, and K. P. Singh, *Exp. Astron.* **53**, 1 (2022).
- [2] C. N. Coelho, A. Kuusela, S. Li, H. Zhuang, J. Ngadiuba, T. K. Aarrestad, V. Loncar, M. Pierini, A. A. Pol, and S. Summers, *Nat. Mach. Intell.* **3**, 675 (2021).
- [3] M. Kezunovic, P. Pinson, Z. Obradovic, S. Grijalva, T. Hong, and R. Bessa, *Electric Power Systems Research* **189**, 106788 (2020).
- [4] J. Gohil, J. Patel, J. Chopra, K. Chhaya, J. Taravia, and M. Shah, *Environmental science and pollution research international* **28**, 64084 (2021).
- [5] G. E. Santoro, R. Martoňák, E. Tosatti, and R. Car, *Science* **295**, 2427 (2002).
- [6] G. Carleo and M. Troyer, *Science* **355**, 602 (2017).
- [7] J. Tilly, H. Chen, S. Cao, D. Picozzi, K. Setia, Y. Li, E. Grant, L. Wossnig, I. Rungger, G. H. Booth, and J. Tennyson, *Phys. Rep.* **986**, 1 (2022).
- [8] A. Mattioni, F. Caycedo-Soler, S. F. Huelga, and M. B. Plenio, *Phys. Rev. X* **11**, 041003 (2021).
- [9] F. Mattiotti, M. Sarovar, G. G. Giusteri, F. Borgonovi, and G. L. Celardo, *New J. Phys.* **24**, 013027 (2022).
- [10] M. R. Shaebani, R. Jose, L. Santen, L. Stankevics, and F. Lautenschläger, *Phys. Rev. Lett.* **125**, 268102 (2020).
- [11] H. Meyer and H. Rieger, *Phys. Rev. Lett.* **127**, 070601 (2021).
- [12] C. Gao, C. Liu, D. Schenz, X. Li, Z. Zhang, M. Jusup, Z. Wang, M. Beekman, and T. Nakagaki, *Phys. Life Rev.* **29**, 1 (2019).
- [13] B. Meyer, *Swarm Intell.* **11**, 131 (2017).
- [14] X.-S. Yang, *J. Comput. Sci.* **46**, 101104 (2020).
- [15] A. Tero, R. Kobayashi, and T. Nakagaki, *J. Theor. Biol.* **244**, 553 (2007).
- [16] S. Li, H. Chen, M. Wang, A. A. Heidari, and S. Mirjalili, *Future Gener. Comput. Syst.* **111**, 300 (2020).
- [17] B. N. Örnek, S. B. Aydemir, T. Düzenli, and B. Özak, *Math. Comput. Simul.* **198**, 253 (2022).
- [18] E. Meron, *Phys. Rep.* **218**, 1 (1992).
- [19] B. Lindner, J. García-Ojalvo, A. Neiman, and L. Schimansky-Geier, *Phys. Rep.* **392**, 321 (2004).
- [20] L. Gammaitoni, P. Hänggi, P. Jung, and F. Marchesoni, *Rev. Mod. Phys.* **70**, 223 (1998).
- [21] M. c. v. Perc, *Phys. Rev. E* **72**, 016207 (2005).
- [22] H. Nakao, K. Arai, and Y. Kawamura, *Phys. Rev. Lett.* **98**, 184101 (2007).
- [23] S. Boccaletti, J. Kurths, G. Osipov, D. Valladares, and C. Zhou, *Phys. Rep.* **366**, 1 (2002).
- [24] C. Van den Broeck, J. M. R. Parrondo, and R. Toral, *Phys. Rev. Lett.* **73**, 3395 (1994).
- [25] F. Sagués, J. M. Sancho, and J. García-Ojalvo, *Rev. Mod. Phys.* **79**, 829 (2007).
- [26] V. Bonifaci, E. Facca, F. Folz, A. Karrenbauer, P. Kolev, K. Mehlhorn, G. Morigi, G. Shahkarami, and Q. Vermande, *Theor. Comput. Sci.* **920**, 1 (2022).
- [27] A. Lonardi, M. Putti, and C. D. Bacco, *Sci. Rep.* **12** (2022).
- [28] *Stochastic Processes in Physics and Chemistry*, 3rd ed., edited by N. Van Kampen, North-Holland Personal Library (Elsevier, Amsterdam, 2007).
- [29] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevLett.130.267401> for (1) the parameter values that we used, details on numerical simulations, and details on calculating the potential p_i^u at node u , (2) the steady state and the convergence rate, (3) the disparity filter, (4) the robustness of the network, (5) the dependence on the injection current and for a larger number of demands, and (6) movies of the dynamics leading to the networks (A)–(G) in Fig. 2.
- [30] B. Meyer, C. Ansoorge, and T. Nakagaki, *PLoS One* **12**, 1 (2017).
- [31] F. Folz, K. Mehlhorn, and G. Morigi, *Phys. Rev. E* **104**, 054215 (2021).
- [32] We note that the variable α here is physically equivalent to the temperature T of an external bath according to the relation $T \propto \alpha^2$ [28].
- [33] F. Kaiser and D. Witthaut, *Phys. Rev. Res.* **3**, 023161 (2021).
- [34] T. Nakagaki, H. Yamada, and Á. Tóth, *Nature (London)* **407**, 470 (2000).
- [35] A. Tero, S. Takagi, T. Saigusa, K. Ito, D. P. Bebbler, M. D. Fricker, K. Yumiki, R. Kobayashi, and T. Nakagaki, *Science* **327**, 439 (2010).
- [36] M. A. Serrano, M. Boguna, and A. Vespignani, *Proc. Natl. Acad. Sci. U.S.A.* **106**, 6483 (2009).
- [37] P. E. Kloeden and E. Platen, *Numerical Solution of Stochastic Differential Equations* (Springer, Berlin, 1992).
- [38] W. Ellens and R. Kooij, [arXiv:1311.5064](https://arxiv.org/abs/1311.5064).
- [39] F. Folz, K. Mehlhorn, and G. Morigi (unpublished).
- [40] T. Nakagaki, H. Yamada, and M. Hara, *Biophys. Chem.* **107**, 1 (2004).
- [41] C. Oettmeier, T. Nakagaki, and H.-G. Döbereiner, *J. Phys. D* **53**, 310201 (2020).
- [42] K. Alim, G. Amselem, F. Peaudecerf, M. P. Brenner, and A. Pringle, *Proc. Natl. Acad. Sci. U.S.A.* **110**, 13306 (2013).
- [43] P. Stewart and B. T. Stewart, *Exp. Cell Res.* **18**, 374 (1959).

- [44] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, *Science* **220**, 671 (1983).
[45] R. Motwani and P. Raghavan, *Randomized Algorithms* (Cambridge University Press, Cambridge, England, 1995).
[46] T. D. Frank, *Nonlinear Fokker-Planck Equations* (Springer-Verlag, Berlin, 2005).
[47] Y.-Y. Liu and A.-L. Barabási, *Rev. Mod. Phys.* **88**, 035006 (2016).

Supplemental Material: noise-induced network topologies

Frederic Folz¹, Kurt Mehlhorn², and Giovanna Morigi¹

¹*Theoretische Physik, Universität des Saarlandes, 66123 Saarbrücken, Germany*

²*Algorithms and Complexity Group, Max-Planck-Institut für Informatik, Saarland Informatics Campus, 66123 Saarbrücken, Germany*

(Dated: June 20, 2023)

I. PARAMETERS AND NUMERICAL SIMULATIONS

In the model we set the length $L_{u,v}$ of the edge u, v equal to unity when the nodes are nearest neighbors and equal to $\sqrt{2}$ when they are connected by a diagonal. We take the exponent of the activation function $n = 1.2$ and set $\kappa = \gamma = 1$. We impose $I_i = 0.45$ for all demands i . For reference, we set the potential at the node u neighboring the source node of demand 1 to the right to $p_u = 0$. The initial state of the simulations has the conductivities of all edges equal to the value $D_{u,v}^0 = 0.5$. The conductivities are calculated by numerically integrating Eq. (2), together with Eq. (1). The calculation of the potential p_u^i at node u and for the demand i is performed by solving a set of linear equations. Let ℓ be the number of network nodes and b^i be the vector determining the constraint, such that it has a value of I_i at the source node s_+^i , $-I_i$ at the sink s_-^i , and 0 otherwise. The vector $p^i = (p_1^i, \dots, p_\ell^i)$ containing all node potentials associated to demand i is found by solving the linear system of equations [1, 2]

$$Mp^i = b^i, \quad (1)$$

where M is a $\ell \times \ell$ matrix, $M = ADL^{-1}A^T$ with: (i) $D = \text{diag}(D_{e_1}, \dots, D_{e_m})$ the $m \times m$ diagonal matrix whose diagonal elements are the edge conductivities, (ii) $L = \text{diag}(L_{e_1}, \dots, L_{e_m})$ the $m \times m$ diagonal matrix whose eigenvalues are the edge length, and (iii) A the $\ell \times m$ node-arc incidence matrix of the network. In particular, the column $(A^T)_e$ has a value of 1 in position u and a value of -1 in position v for all edges $e = (u, v)$.

The integration of Eq. (2) is performed using stochastic differential equations that are implemented using the Euler-Maruyama scheme with a step size of $\Delta t = 0.1\gamma^{-1}$, as outlined in Ref. [3]. The evolution time $t_{\text{end}} = 250/\gamma$ is chosen after testing that each trajectory, namely, each individual evolution of the network, has reached a (meta)-stable configuration. Figures S1(a) and (b) display few trajectories at a fixed value of α .

II. STEADY STATE AND CONVERGENCE RATE

The reported topologies are obtained for fixed initial conditions and finite integration times, and thus, we cannot claim that they are the steady state. The steady state, in fact, is the solution of a multi-dimensional and nonlinear Fokker-Planck equation, which is not amenable

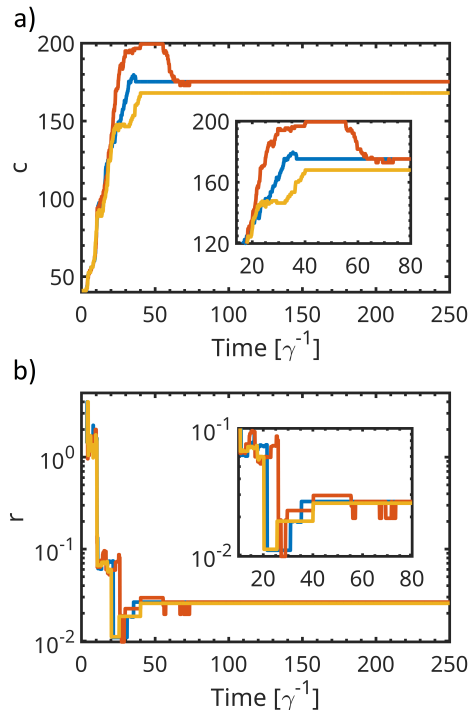


FIG. S1. The time evolution of the cost (a) and robustness (b) for three trajectories, represented by different colors. The trajectories have been determined using $\alpha = 0.002$. The insets show a zoom in the dynamics around the convergence time t_δ^x , see text.

of analytical treatment [4]. Nevertheless, over the considered time, the integrated trajectories converge relatively fast towards one of the topologies of Fig. 2 in the main text. In order to decide whether the dynamics has converged to a (meta)stable configuration, we extend the definition of robustness and cost to a time dependent variable, which we determine on the instantaneous network's backbone, and monitor their dynamics. We identify the steady state as the configuration of the system for which the mean value is constant and the fluctuations are given by the variance set by the noise. To quantify the convergence speed, we introduce the quantity γ_x for $x = c, r$, which has the dimensions of a rate and is defined

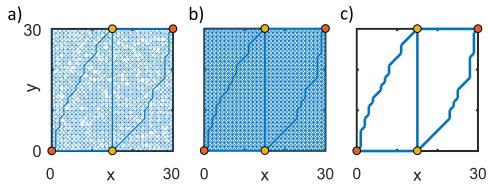


FIG. S2. Subplot (a) displays an example of a network reached after a sufficiently long integration time in the presence of noise ($\alpha = 0.002$), subplot (b) is the corresponding time average over the time interval $[249 : 250]/\gamma$ at the end of the simulation. The widths of the edges are graphically scaled proportionally to the corresponding conductivities. Subplot (c) displays the multi-scale backbone extracted from (b) using a filtering procedure (see text). Details on the numerical simulations are reported in the caption of Fig. 2 in the main text.

as

$$\gamma_x^{-1} = \langle t_\delta^x \rangle, \quad (2)$$

with $\langle \cdot \rangle$ the ensemble average over the convergence time t_δ^x . The latter is defined as $t_\delta^x = \max(\{t; |x(t') - x(t_{end})| < \delta \cdot x(t_{end}) \forall t' \in [t, t_{end}]\})$ with $\delta > 0$. In the following, we set $\delta = 0.05$, unless otherwise stated. The rates γ_x quantify the average convergence rate of the costs and the robustness to the steady state. Furthermore, we define the quantity $\gamma_{c,r}^{-1} = \langle \max(t_\delta^c, t_\delta^r) \rangle$, which accounts for the combined convergence time of the costs and the robustness. We calculate γ_x for different values of the noise strength α by numerically solving the model given by Eqs. (1)-(2). Hereby, we average over 5000 simulation runs for each value of α . In Figs. 4(a) and (b) in the main text the average convergence rates γ_x and the standard deviations σ_{γ_x} are shown as a function of the noise strength α .

III. DISPARITY FILTER

Extracting the network topology requires filtering the connection above a certain threshold. A possible ansatz consists of choosing a constant threshold for all edges. However, this approach does not account for the statistical importance that certain links of a node have over others: even if all values of the conductivities might be below threshold, some links can be statistically relevant.

In order to avoid this problem, we apply the following procedure. For each realization (trajectory) we first average the conductivities $D_{u,v}$ over the time interval $[t_{end} - d_t, t_{end}]$, with $d_t = 1/\gamma$ corresponding to 10 time steps. The time d_t is fixed by requiring that over this time the average distribution solely due to noise is stationary and is verified integrating Eq. (2) after setting $f = 0$. After the averaging, the effect of fluctuations is leveled to a background value as visible by comparing Fig. S2(a)

with Fig. S2(b), where the time averaging was performed. We then introduce a global offset $D_{u,v} \rightarrow D_{u,v} + 0.05$ for all edges (u, v) and apply the disparity filter of Ref. [5] using the significance level $\beta = 0.3$. The procedure of Ref. [5] is implemented as follows. We determine the strength of each node u : $s_u = \sum_{v \in E_u} D_{u,v}$ and then normalize the conductivities of the edges that connect a node with its nearest neighbors by $p_v = D_{u,v}/s_u$ such that $\sum_{v \in E_u} p_v = 1$. We remove all edges whose conductivities are not statistically significant, i.e. are purely random. As a null hypothesis, it is assumed that the edge conductivities of a certain node of degree k (which can be either 8, or 5, or 3, here depending on the node location within the grid) are produced by a random assignment from a uniform distribution. In order to find the null hypothesis we use the method of induction. For $k = 2$ edges we have $p_1 + p_2 = 1$ and $p_1 = x$ where x is a random number in the interval $[0, 1]$. We divide the interval into infinitesimal steps dx and introduce the probability density $\rho(x)$ such that $p_1 = \rho(x)dx$. For $k = 2$, then $\rho(x) = 1$. For $k > 2$, we find $\rho(x)$ by solving the nested integral $\rho(x)dx = dxk \int_0^{1-x} dx_1 \dots \int_0^{1-x_{k-2}} dx_{k-3}$, which gives [5]

$$\rho(x)dx = (k-1)(1-x)^{k-2}dx. \quad (3)$$

The probability $\beta_{u,v}$ that the edge (u, v) is compatible with the null hypothesis is given by

$$\beta_{u,v} = 1 - (k-1) \int_0^{\tilde{D}_{u,v}} (1-x)^{k-2}dx. \quad (4)$$

The disparity filter removes all edges for which it holds $\beta_{u,v} \geq \beta$ with a significance level $\beta \in [0, 1]$ as these edges are not statistically relevant. We note that the filter fails at sufficiently large values of the noise amplitude α , which we do not consider here. These large values correspond to the physical situation where disorder prevails over the order imposed by the nonlinear force.

IV. ROBUSTNESS OF THE NETWORK

In order to determine the robustness, we count the number of links of the filtered network. For this purpose, we assign the same conductivity to all edges of the network's backbone. We remark that various approaches to define a measure of robustness are discussed in literature. In the work of Ref. [6], the fault-tolerance of a network was measured by counting the number of edges that can be removed without separating the network into two parts. Here, we chose the inverse of the total effective resistance of the network as the measure of robustness, see [7]. This approach takes into account both the number of different paths that can be used to fulfill a demand and the paths length. Before calculating the total effective resistance, we normalize all edge conductivities $D_{u,v} > 0$ as we intend to focus on the length as the quality criterion for a path for simplicity. Extending the

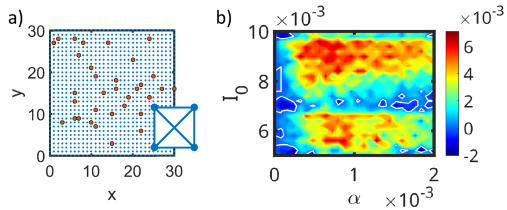


FIG. S3. (a) Network self-organization is simulated on a grid of 31×31 nodes, whereby source and sink nodes are placed such that they represent the relative location of major cities around Tokyo and are labelled red. We assume that there is a demand between each pair of cities which gives rise to a total of 528 demands. The network design results from the dynamics of the edges, which are modelled by time-varying conductivity on an electrical network and in the presence of additive noise according to Eq. (1) and (2). The difference in robustness between the stochastic case and the deterministic case as a function of the noise and the flow I_0 which is the same for all demands is shown in (b). The difference in robustness is given as a color code: red means that the robustness in the stochastic case is larger than in the deterministic case and blue means the opposite. The solid white lines indicate that the robustness in the stochastic case is the same as in the deterministic case. The difference in robustness was calculated by averaging over 50 simulation runs for each pair of I_0 and α .

analysis to a measure of robustness that also takes into account the amplitude of the edge conductivities could be an interesting future consideration.

V. DEPENDENCE ON THE INJECTION CURRENT AND FOR A LARGER NUMBER OF DEMANDS

We consider a grid with a larger number of demands in the following. In Fig. S3(a), a grid of 31×31 nodes is shown, whereby source and sink nodes are placed such that they represent the relative location of major cities around Tokyo [6] and are labelled red. We assume that there is a demand between each pair of cities, which gives rise to a total of 528 demands. Figure S3(b) displays the difference in robustness between the noiseless case and the stochastic case as a function of the noise amplitude α and of the flow I_0 (which is the same for all demands) for $n = 1.6$. The difference in robustness is given as a color code: red means that the robustness in the stochastic case is larger than in the deterministic case, blue means the opposite. The solid white lines indicate that the robustness in the stochastic case is the same as in the deterministic case. The difference in robustness was calculated by averaging over 50 trajectories for each pair of I_0 and α .

In general, the noise-induced resonances appear for all values of the injection current we considered. We note

that they also occur when considering different values of I_i for different demands. We note that the dependence on the injected current introduces additional features, which are due to discontinuous transitions and which we will discuss elsewhere.

VI. MOVIES OF THE DYNAMICS

As part of the Supplemental Material, we provide movies of the dynamics leading to the networks (A)-(G) shown in Fig. 2 of the main text:

- Network (A): https://www.uni-saarland.de/fileadmin/upload/lehrstuhl/morigi/noise-induced-network-topologies/network_A.gif
- Network (B): https://www.uni-saarland.de/fileadmin/upload/lehrstuhl/morigi/noise-induced-network-topologies/network_B.gif
- Network (C): https://www.uni-saarland.de/fileadmin/upload/lehrstuhl/morigi/noise-induced-network-topologies/network_C.gif
- Network (D): https://www.uni-saarland.de/fileadmin/upload/lehrstuhl/morigi/noise-induced-network-topologies/network_D.gif
- Network (E): https://www.uni-saarland.de/fileadmin/upload/lehrstuhl/morigi/noise-induced-network-topologies/network_E.gif
- Network (F): https://www.uni-saarland.de/fileadmin/upload/lehrstuhl/morigi/noise-induced-network-topologies/network_F.gif
- Network (G): https://www.uni-saarland.de/fileadmin/upload/lehrstuhl/morigi/noise-induced-network-topologies/network_G.gif

The movies cover the first 20% of the total simulation time. For the networks (B, C), we used a noise amplitude of $\alpha = 0.0001$, for the network (D), we used $\alpha = 0.0008$, for the networks (E, F), we used $\alpha = 0.001$ and for the network (G), we used $\alpha = 0.005$. All other parameter values are chosen the same as described in the caption of Fig. 2 of the main text.

-
- [1] V. Bonifaci, E. Facca, F. Folz, A. Karrenbauer, P. Kolev, K. Mehlhorn, G. Morigi, G. Shahkarami, and Q. Vermande, *Theoretical Computer Science* **920**, 1 (2022).
- [2] A. Lonardi, M. Putti, and C. D. Bacco, *Scientific Reports* **12** (2022), <https://doi.org/10.1038/s41598-022-11348-9>.
- [3] P. E. Kloeden and E. Platen, *Numerical Solution of Stochastic Differential Equations* (Springer Berlin Heidelberg, 1992).
- [4] T. D. Frank, *Nonlinear Fokker-Planck Equations* (Springer-Verlag (Berlin, Heidelberg, New York), 2005).
- [5] M. A. Serrano, M. Boguna, and A. Vespignani, *Proceedings of the National Academy of Sciences* **106**, 6483 (2009), <https://www.pnas.org/doi/pdf/10.1073/pnas.0808904106>.
- [6] A. Tero, S. Takagi, T. Saigusa, K. Ito, D. P. Beber, M. D. Fricker, K. Yumiki, R. Kobayashi, and T. Nakagaki, *Science* **327**, 439 (2010), <https://www.science.org/doi/pdf/10.1126/science.1177894>.
- [7] W. Ellens and R. Kooij, arXiv , arXiv:1311.5064 (2013).

CHAPTER IV

PART 1: SELF-ORGANIZED TRANSPORT IN NOISY DYNAMIC NETWORKS

Self-organized transport in noisy dynamic networks

Phys. Rev. E 110, 044310 – Published 21 October 2024

© 2024 American Physical Society

DOI: 10.1103/PhysRevE.110.044310

Authors: Frederic Folz¹, Kurt Mehlhorn², Giovanna Morigi¹




¹ *Theoretische Physik, Universität des Saarlandes, 66123 Saarbrücken, Germany*

² *Algorithms and Complexity Group, Max-Planck-Institut für Informatik, Saarland Informatics Campus, 66123 Saarbrücken, Germany*

Abstract:


We present a numerical study of multicommodity transport in a noisy, nonlinear network. The nonlinearity determines the dynamics of the edge capacities, which can be amplified or suppressed depending on the local current flowing across an edge. We consider network self-organization for three different nonlinear functions: For all three we identify parameter regimes where noise leads to self-organization into more robust topologies, that are not found by the sole noiseless dynamics. Moreover, the interplay between noise and specific functional behavior of the nonlinearity gives rise to different features, such as (i) continuous or discontinuous responses to the demand strength and (ii) either single or multistable solutions. Our study shows the crucial role of the activation function on noise-assisted phenomena.

Self-organized transport in noisy dynamic networks

Frederic Folz ¹, Kurt Mehlhorn ², and Giovanna Morigi ¹

¹*Theoretische Physik, Universität des Saarlandes, 66123 Saarbrücken, Germany*

²*Algorithms and Complexity Group, Max-Planck-Institut für Informatik, Saarland Informatics Campus, 66123 Saarbrücken, Germany*

 (Received 19 June 2024; accepted 2 October 2024; published 21 October 2024)

We present a numerical study of multicommodity transport in a noisy, nonlinear network. The nonlinearity determines the dynamics of the edge capacities, which can be amplified or suppressed depending on the local current flowing across an edge. We consider network self-organization for three different nonlinear functions: For all three we identify parameter regimes where noise leads to self-organization into more robust topologies, that are not found by the sole noiseless dynamics. Moreover, the interplay between noise and specific functional behavior of the nonlinearity gives rise to different features, such as (i) continuous or discontinuous responses to the demand strength and (ii) either single or multistable solutions. Our study shows the crucial role of the activation function on noise-assisted phenomena.

DOI: [10.1103/PhysRevE.110.044310](https://doi.org/10.1103/PhysRevE.110.044310)

I. INTRODUCTION

Networks are a commonly used concept in many disciplines and powerful models for transport. Efficient routing of commodities such as water, power, or information from sources to sinks can be described as a problem of connecting nodes on a graph for given constraints and requirements [1–3]. There are different approaches to network design. Some consist of solving differential equations, which are derived from an appropriately identified cost function [4,5]. The chosen rule determines the dynamics of the edges connecting the nodes, whose stable fixed point is a target topology. Extensive studies show how, for a given class of power-law functions determining the equations of motion, a variation of the exponent can give rise to phase transitions in the network structure, from spanning trees that minimize the cost to loops that maximize the robustness [5–7].

Another approach is based on bioinspired algorithms. Prominent examples are algorithms inspired by the structures formed by ant colonies [8] or by the filaments of *Physarum polycephalum*, a single-celled organism that is also known as true slime mold [9]. Despite its lack of any form of a nervous system, *Physarum polycephalum* is able to find good solutions to small instances of popular optimization problems, such as finding the shortest path through a maze [10,11], creating efficient and fault-tolerant networks [12], and solving the traveling salesman problem [13]. Algorithms inspired by *Physarum polycephalum* have been implemented for various optimization problems, including network design for multicommodity flows [14–16].

Bioinspired algorithms find successful applications for multicommodity flow problems [6,12,17,18]. A prominent example is a city transportation network: Each commodity models the passengers traveling between two given stations. *Physarum*-inspired algorithms for solving multicommodity flow problems typically tend to identify a network satisfying the constraints according to a rule that promotes transport along shared routes and inhibits it when the flow along one

edge is below a chosen threshold [17]. In Ref. [7], such an algorithm was used to find the optimal routing of passenger flows through the network of the Paris metro. This study includes a comparative assessment of different activation functions governing the edge dynamics: When the activation functions are power laws of the flow, the exponent of the power law determines the topological properties of the emerging networks, and the network topologies undergo a phase transition between treelike and loopy topologies [5].

In the present study, we analyze the influence of noise on multicommodity flow problems, where the edge capacity is a variable that depends on the current flowing across the edge through a nonlinear (activation) function [19]. We assume that the edge capacity can undergo stochastic fluctuations and examine the emerging network topologies for three different functional behaviors of the activation function. In particular, we determine the characteristics of the emerging networks as a function of the noise amplitude and analyze their robustness, transport efficiency, and cost. For our analysis, we choose the same case study as in Refs. [12,17], where the graph has a number of demands geometrically arranged to mimic the relative locations of the major cities around Tokyo. The graph and the demands are illustrated in Fig. 1(a). Examples of the emerging networks are provided in Figs. 1(b) and 1(c).

Our study extends recent work on optimization of transport in simple systems, consisting of either a single or two commodities, and in the presence of Gaussian white noise [18,20,21]. These works showed that, for a finite range of noise amplitudes, the interplay between noise and a sigmoidal activation function can lead to the most robust solution in a relatively short convergence time. The phenomenology is reminiscent of noise-induced coherent effects, found in models simulating forest fires [22] and neurons [23], and include phenomena such as stochastic and coherence resonance [23–25], synchronization [26,27], and noise-induced phase transitions [28,29].

Our research question, the role of the specific activation function on network self-organization, is also motivated by

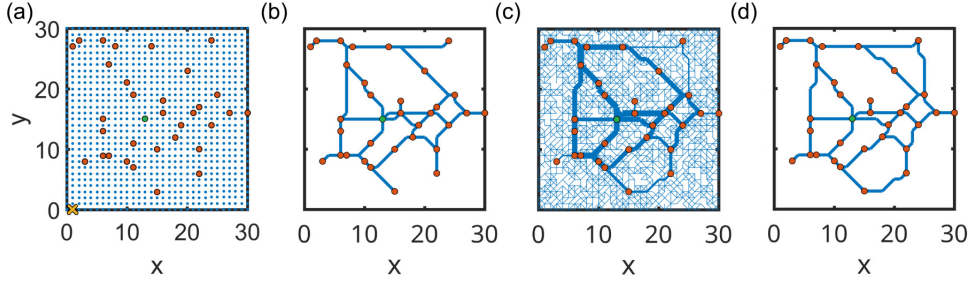


FIG. 1. Network self-organization for a multicommodity flow problem in the presence of stochastic fluctuations. The sources and sinks simulate the metro stations of Tokyo railroad on a grid of 31×31 nodes with 528 demands, as illustrated in (a). The red dots represent the relative locations of the major cities around Tokyo; Tokyo itself is indicated by the green dot. The network design results from the dynamics of the edges, which are modeled by time-varying edge capacities [Eq. (10)]. The emerging network topology is shown in (b) for the noiseless case ($\alpha = 0$). Subplot (c) displays a single trajectory in the presence of noise ($\alpha = 0.002$) and for a specific choice of the activation function (sigmoidal with exponent $n = 1.6$). The widths of the edges are proportional to the corresponding capacities. The backbone of the network in (c) is extracted using a filtering procedure, the filtered network is shown in (d). In all simulations, we initially set all edge capacities equal to the value $D_{u,v}^0 = 0.5$. Furthermore, for each commodity i , we fix the potential $p_x^i = 0$ at the node indicated by the yellow cross as a reference. Further details on the numerical simulations are reported in the main text.

the observation that, in deep learning, the choice of the activation function determines the network expressivity [30,31]. Within a different, yet connected, framework, here we study how noise-induced network self-organization depends on the choice of one of three representative functionals, which have been considered in the literature of multicommodity transport.

This paper is organized as follows. In Sec. II, we introduce the model and define the measures used to evaluate the emerging network topologies for a graph with demands mimicking the Tokyo railroad transport problem. We then determine the network measures as a function of the demand strength for the three different activation functions with no noise. Section III is devoted to the numerical methods and the algorithms used to extract the network topology from the stochastic dynamics. The network measures as a function of the noise amplitude are presented in Sec. IV. Conclusions and outlook are drawn in Sec. V.

II. NOISELESS TRANSPORT ON A GRID

In this section, we introduce the model, define the network measures used for evaluating the topologies and apply the multicommodity flow problem to routing transport demands with the geometry of the greater area of Tokyo. We focus on noiseless transport and study the network topologies for different activation functions.

A. The model

The multicommodity flow problem is represented by a set of k transport demands on a graph that discretizes the space and is composed of N nodes. Each node, labeled $u = 1, \dots, N$, is connected to a number of other nodes, described by the set E_u . We denote the edge-connecting nodes u and the neighbor $v \in E_u$ by the pair (u, v) . The edge length is given by $L_{u,v}$.

Each demand is formed by a pair of a source node s_+^i , where a flow is injected ($+I^i$), and a sink node (s_-^i), where

it is extracted ($-I^i$). The demand i is realized by a flow I^i across the edges of the network. The flow through the edge-connecting nodes u and v is denoted as $Q_{u,v}^i$. At the sources and sinks, it obeys the constraints

$$\sum_{v \in E_{s_{\pm}^i}} Q_{s_{\pm}^i, v}^i = \pm I^i. \quad (1)$$

At the other nodes, instead, the flow is conserved:

$$\sum_{v \in E_u} Q_{u,v}^i = 0. \quad (2)$$

For an electrical circuit, this is Kirchhoff's law and Q is the electrical current. In the dynamics of network self-organization, these nodes are decision points. The flow of commodity i is directed along the edges (u, v) with nonvanishing edge capacity $D_{u,v}(t)$ and obeys the equation

$$Q_{u,v}^i(t) = \frac{D_{u,v}(t)}{L_{u,v}} (p_u^i(t) - p_v^i(t)), \quad (3)$$

where p_u^i is a potential (or pressure) at the node u for commodity i . In an electrical circuit, which is the example we will refer to throughout this paper, Eq. (3) is Ohm's law and $D_{u,v}(t)$ is a dynamical conductivity. The dynamics of the variables p_u^i and $D_{u,v}$ determine the resulting network.

B. The potential

The potential p_u^i is determined for each demand i as a function of $D_{u,v}(t)$ by solving the linear set of equations in Eq. (3), subject to the boundary conditions at the source and sink nodes given by Eq. (1), and Kirchhoff's law at any other node, Eq. (2). We fix the potential $p_x^i = 0$ at the node indicated by the yellow cross in Fig. 1(a) as a reference. Let N be the number of network nodes and $b^i \in \mathbb{R}^N$ the vector representing the constraints, such that it has a value of $+I^i$ at the source node s_+^i and $-I^i$ at the sink s_-^i . Otherwise, the entry is zero. Then, the vector $p^i = (p_1^i, \dots, p_N^i)$ containing all node

TABLE I. Three classes of activation functions. Here, $Q_{u,v}$ denotes the total flow along the edge, which is determined as $Q_{u,v} = \sum_i |Q_{u,v}^{(i)}|$ (one norm). For the Hill function, we consider $n = 1.6$ and $n = 2$. In the definition of the ReLU function, $\theta(x)$ is Heaviside's function [$\theta(x) = 1$ for $x > 0$ and $\theta(x) = 0$ otherwise].

Activation function	Functional form	Threshold
Two-norm	$f(\{Q_{u,v}^{(i)}\}) = \sqrt{(Q_{u,v}^{(1)})^2 + \dots + (Q_{u,v}^{(k)})^2}$	-
Hill/sigmoidal	$f(\{Q_{u,v}^{(i)}\}) = (Q_{u,v})^n / (\kappa^n + (Q_{u,v})^n)$	$\kappa = 1$
ReLU	$f(\{Q_{u,v}^{(i)}\}) = (Q_{u,v} - \kappa)\theta(Q_{u,v} - \kappa)$	$\kappa = 0.01$

potentials associated to demand i is determined by the linear system of equations [7,17]

$$Mp^i = b^i, \quad (4)$$

where M is a $N \times N$ matrix,

$$M = ADL^{-1}A^T, \quad (5)$$

with: (i) $D = \text{diag}(D_{e_1}, \dots, D_{e_m})$ the $m \times m$ diagonal matrix whose diagonal elements are the edge capacities (conductivities), (ii) $L = \text{diag}(L_{e_1}, \dots, L_{e_m})$ the $m \times m$ diagonal matrix whose eigenvalues are the edge lengths, and (iii) A the $N \times m$ node-arc incidence matrix of the network, i.e., the column $(A^T)_e$ has a value of 1 in position u and a value of -1 in position v for all edges $e = (u, v)$.

C. The edge capacity (conductivity)

The edge capacity (conductivity) $D_{u,v}(t)$ obeys the nonlinear differential equation [20,21]

$$\partial_t D_{u,v} = f(\{Q_{u,v}^{(i)}\}) - \gamma D_{u,v}, \quad (6)$$

where $f(\{Q_{u,v}^{(i)}\})$ is the activation function modeling the feedback of the flows associated to the commodities $j = 1, \dots, k$ on the evolution of the edge conductivity. The edge capacities are damped with rate γ and hence edges with little flow on them (in particular, flow below a fixed threshold) are suppressed.

The network topology is determined by the resulting edge capacities $D_{u,v}$. They are found by integrating Eq. (6) using Eq. (3), where the potential is found from Eqs. (4) and (5).

D. The activation function

We consider three different classes of activation functions: the two-norm, the Hill (or sigmoidal) function, and the ReLU function. Each of them depends on the total flow across an edge, as shown in Table I. As a consequence, sharing of edges between different commodities is rewarded. The different functional forms of the activation functions leads to different behavior as we discuss next.

Of the three activation functions, only the two-norm function is a homogeneous function of the flow. Specifically, it is a homogeneous function with degree 1. As a consequence, the dynamics is independent of the demand strength I_0 . In fact, equations (6) and (3) are invariant after rescaling the conductivity and the flow by I_0 .

The ReLU function is not a homogeneous function (even though it tends to behave as if it were a homogeneous function for demand strengths much larger than the threshold, $I_0 \gg \kappa$).

In this work, we consider values of $I_0 < \kappa$, i.e., smaller than the threshold, where the response depends on the choice of I_0 .

The sigmoidal function saturates: This limits the flow that can be transported through the edge. As the flow between two nodes approaches the maximum capacity of the edge, the dynamics tends to construct multiple connections between them. In the following, we consider two different powers n for the Hill function and analyze the regime where the dynamics is sensitive to small gradients as the demand strength I_0 increases above threshold.

The dynamics governed by the two-norm function has been extensively studied in Ref. [17]. In that work, a Lyapunov function was determined, the limit of the dynamics was formally characterized, and it was shown that the limit optimizes a mixture of transport efficiency and network cost. No Lyapunov functions for the Hill and the ReLU function are known. Moreover, since the sigmoidal and the ReLU function are nonlinear and nonhomogeneous functions of the flow, it is generally difficult to systematically choose the parameters in order to compare the network topologies for different activation functions. For this reason, we fix the parameters using a phenomenological approach that takes as a reference the robustness of the network when the activation function is the two-norm function. We then choose the threshold κ of the Hill and the ReLU function and identify the interval of values of the demand strength I_0 that give a comparable network robustness.

E. Performance measures of networks

In what follows, we characterize the resulting network as a function of I_0 . For this purpose, we introduce measures for (i) the robustness, (ii) the transport efficiency, and (iii) the cost of the network.

The robustness is measured by the inverse of the average effective resistance of the network [32]:

$$r = \frac{k}{\sum_{i=1}^k R^i}, \quad (7)$$

where

$$R^i = (p_{s_+^i}^i - p_{s_-^i}^i) / I^i$$

is the effective resistance between the source node s_+^i and the sink node s_-^i of demand i . The quantity R^i takes into account both the number of connections between s_+^i and s_-^i as well as their lengths. The measure r is an indicator of robustness against edge failure. In fact, adding edges and/or reducing the length of connections increases the measure [33]. This is a different, nevertheless similar approach to existing

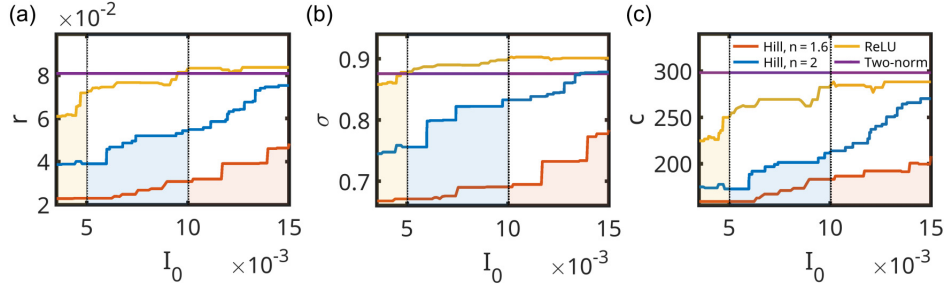


FIG. 2. Network measures for the noiseless dynamics and as a function of the demand strength I_0 . The subplot display (a) the robustness r , (b) the transport efficiency σ , and (c) the cost c . They are extracted from the network obtained by integrating the coupled equations (3)–(6) for a time of $1000\gamma^{-1}$ using a different activation function in each case, see legend and Table I. The measures of the networks are calculated after applying the disparity filter, which we detail in Sec. III B. In the rest of this work, we will focus on the range of I_0 indicated by the shaded regions.

measures, where the fault-tolerance of a network is measured by counting the number of edges that can be removed without separating the network into two parts, for example [12]. Note that we set all nonvanishing edge conductivities to the same value (here, $D_{u,v} = 0.5$) before calculating r , corresponding to a fixed resistance per length.

The transport efficiency σ of the network is defined as the average over all demands of the length of the shortest path d^i connecting the source and the sink node s_+^i and s_-^i . It reads

$$\sigma = \frac{k}{\sum_{i=1}^k d^i}. \quad (8)$$

The cost of the network c is the total length, found by summing over the segments $L_{u,v}$ where the conductivity is nonzero:

$$c = \sum_{(u,v) \in E} L_{u,v}, \quad (9)$$

with E the set of all edges with $D_{u,v} > 0$.

These quantities are displayed in Fig. 2 as a function of the demand strength I_0 ; the colors represent the different activation functions from Table I.

For the two-norm function, the performance of the networks is, as expected, independent of I_0 . For the sigmoidal and the ReLU functions, the measures increase in a step-like fashion, indicating a discontinuous transition to different topologies of increasing robustness, transport efficiency, and cost. Each topology exists for a certain range of the demand strength I_0 . For the ReLU function, the measures reach a constant value for $I_0 \gg \kappa$, whereas for the sigmoidal function, they keep increasing with I_0 due to the saturating character of the activation function that limits the maximum edge capacity. The measures can locally decrease for increasing I_0 , which may be an artifact of our filtering procedure (see next section and Appendix).

F. Application to the Tokyo transport problem

We apply the model introduced in this section to the multicommodity flow problem represented by the distribution of sources and sinks illustrated by the red circles in Fig. 1(a). We use the same setup as in Refs. [12,17], where the sources and

sinks are geometrically arranged to mimic the relative locations of the major cities around Tokyo. A transport demand between each pair of cities is assumed, which means that the emerging network has to satisfy $k = 528$ demands.

We consider a grid of $N = 31 \times 31$ nodes, of which 33 are the cities. Each node, labeled $u = 1, \dots, N$, is connected to a number of nearest and next-nearest neighbors. The edge length $L_{u,v}$ is set to unity when the nodes u and v are nearest neighbors and to $\sqrt{2}$ when the nodes are connected by a diagonal.

We choose $I^i = I_0$ for all demands i except for those that involve the node at the relative location of Tokyo: For these demands, as in Ref. [12], we set $I^i = 7I_0$ to reflect the importance of Tokyo as the center of the region.

III. MODEL AND NUMERICAL METHODS FOR THE STOCHASTIC DYNAMICS

We finally come to the core of the paper, the study of the networks emerging from the interplay of stochastic dynamics and the nonlinear activation functions. The conductivity now evolves according to the equation

$$\partial_t D_{u,v} = f(\{Q_{u,v}^{(i)}\}) - \gamma D_{u,v} + \sqrt{\gamma} \alpha \xi_{u,v}(t), \quad (10)$$

with the stochastic force $\xi_{u,v}(t)$, whose amplitude is scaled by the parameter α . The force is statistically defined by the average over an ensemble of trajectories: it has no net drift, $\langle \xi_{u,v}(t) \rangle = 0$, and simulates Gaussian white noise with no spatial correlations, $\langle \xi_{u,v}(t) \xi_{u',v'}(t') \rangle = \delta_{u,u'} \delta_{v,v'} \delta(t-t')$ [34]. The network dynamics results from integrating the coupled equations (3), (4), (5), and (10).

In what follows, we first introduce the numerical methods used to integrate the stochastic differential equations. Since the Langevin force $\xi_{u,v}$ gives rise to fluctuations of the edge conductivities, the stationary values of the measures have a finite variance. In order to eliminate the background noise, we apply the filter of Ref. [18] that allows us to identify the statistically relevant edges. In the Appendix, we benchmark it with other filtering procedures.

A. Numerical methods

The integration of Eq. (10) is performed using the Euler-Maruyama scheme, as outlined in Ref. [35]. We have analyzed convergence for different step sizes: For the parameter values considered here, we set the step size $\Delta t = 0.05 \gamma^{-1}$. We then set the evolution time as $t_{\text{end}} = 1000 \gamma^{-1}$. In fact, extensive tests over a statistically relevant ensemble of trajectories (namely, individual network evolutions) show that each trajectory has reached a (meta-)stable configuration. In all simulations, the conductivities of all edges are initially equal to the value $D_{u,v}^0 = 0.5$. For each commodity i , we fix the potential $p_x^i = 0$ at the node indicated by the yellow cross in Fig. 1(a) as a reference.

Figure 1(c) displays a typical network obtained at time t_{end} . Edges with a nonzero conductivity $D_{u,v} > 0$ are drawn with blue lines, whose width is proportional to $D_{u,v}$. Noise leads to a fluctuating distribution of weak connections as well as to statistically relevant links, that are otherwise absent in the noiseless dynamics.

We remark that, as the reported topologies are obtained for finite integration times, we cannot claim that they are the steady state. The dynamics, in fact, is described by a multidimensional and nonlinear Fokker-Planck equation [36], and it is not even guaranteed that a steady state exists. Nevertheless, over the considered time, the integrated trajectories converge relatively fast towards a certain cost, robustness, and transport efficiency and then remain stably trapped about these values, performing fluctuations of the order of the noise amplitude. To these configurations, we then apply the filtering procedure, which we describe in what follows.

B. Filtering procedures

For each simulation run, we apply a filtering procedure to the final conductivities $D_{u,v}$ to retrieve the backbone of the network. In order to identify the statistically relevant edges, we use the filter of Ref. [18]. This filter extends the method of Ref. [37] to noisy networks. We discuss alternative filters in the Appendix. These other filters lead to similar results giving evidence to the appropriateness of our filtering approach.

At the end of our simulations, we average the conductivity over the time interval $\mathcal{I}_t = [t_{\text{end}} - d_t, t_{\text{end}}]$, where t_{end} is the total integration time and $d_t = 1/\gamma$. The averaging levels out the fluctuations. We verified that the interval is long enough for averaging; in particular, changing d_t by a factor of 5 does not change the results. We denote the time-averaged conductivities by $\langle D_{u,v} \rangle$:

$$\langle D_{u,v} \rangle = \frac{1}{d_t} \int_{\mathcal{I}} d\tau D_{u,v}(\tau). \quad (11)$$

We then define $s_u = \sum_{v \in E_u} \langle D_{u,v} \rangle$ as the strength of node u and normalize the conductivities of the incident edges as $\mathcal{P}_v = \langle D_{u,v} \rangle / s_u$ such that $\sum_{v \in E_u} \mathcal{P}_v = 1$. Hereby, E_u denotes the set of edges that are connected to node u , and v denotes the neighboring nodes.

All edges whose conductivities are purely random are statistically not significant. Assume that $\langle D_{u,v} \rangle$ of an edge incident to node u of degree k is sampled from a random uniform distribution. To find the corresponding probability density $\rho(x)$, we use the method of induction. For $k = 2$

edges, it holds $\mathcal{P}_1 + \mathcal{P}_2 = 1$. We set $\mathcal{P}_1 = x$ with a random number x in the interval $[0, 1]$ and define the probability density $\rho(x)$ such that $\mathcal{P}_1 = \rho(x)dx$. It follows that $\rho(x) = 1$. For $k > 2$, it holds

$$\rho(x)dx = dx(k-1) \int_0^{1-x} dx_1 \dots \int_0^{1-x_{k-2}} dx_{k-3},$$

yielding [37]

$$\rho(x)dx = (k-1)(1-x)^{k-2}dx. \quad (12)$$

The probability that the edge (u, v) is compatible with the null hypothesis, i.e., is purely random, is given by [37]

$$\beta_{u,v} = 1 - (k-1) \int_0^{D_{u,v}} (1-x)^{k-2}dx. \quad (13)$$

This value shall be compared with a threshold β , the significance level, which we choose in the interval $\beta \in [0, 1]$. For $\beta_{u,v} \geq \beta$, the edge is statistically not significant and filtered out.

The choice of the significance level has a certain arbitrariness (see, e.g., the discussion in Ref. [37]). We reduce this arbitrariness by benchmarking our results with the results of other filtering procedures, as shown in Appendix. In the following, we use $\beta = 0.12$ for the two-norm function and $\beta = 0.3$ for all other activation functions, yielding comparable values of the network measures.

The significance level limits the range of noise amplitudes for which the filter is useful. In fact, for sufficiently large values of α , the stochastic dynamics becomes dominant so that it fails to extract the backbone of the network. We thus limit our analysis to noise amplitudes for which the filter can be successfully applied.

IV. NETWORK TOPOLOGIES IN THE PRESENCE OF NOISE

In this section, we present the network topologies emerging from the interplay of noise and nonlinear dynamics for each of the three classes of activation functions. We investigate the (filtered) networks as a function of the demand strength I_0 and the noise amplitude α . For each fixed pair of α and I_0 , we evaluate an ensemble of 250 trajectories and determine the robustness, the transport efficiency, and the cost of the backbone of each individual trajectory as defined in Sec. II E.

A. Two-norm function

We first discuss the networks generated using the two-norm function for Gaussian white noise. The white solid line in Fig. 3 displays (a) the mean robustness, (b) the mean transport efficiency, and (c) the mean cost of the network as a function of $I_0 \in [1.5 \times 10^{-2}, 5.5 \times 10^{-2}]$ and for a fixed noise amplitude $\alpha = 0.001$. The mean is taken over 250 trajectories and shows that noise tends to spread out the flow across the network. The trajectory distribution for each I_0 is indicated by the color scale: the measures of each trajectory cluster about the mean value.

We note that the topology depends on I_0 . This is in striking contrast with the noiseless behavior (magenta curves), where all measures are independent of the demand strength I_0 . The

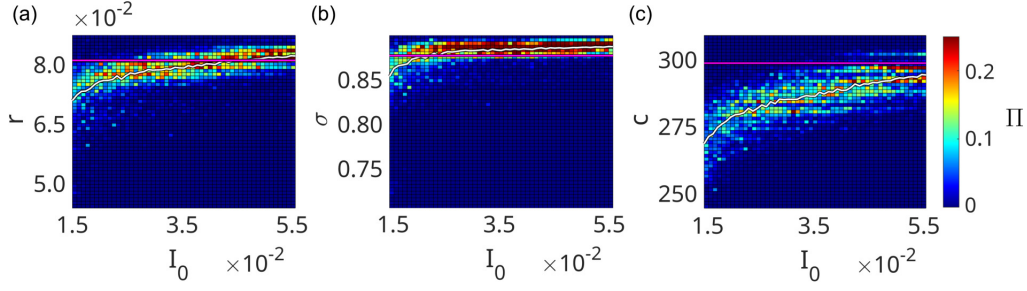


FIG. 3. Network measures as a function of the demand strength I_0 for the two-norm function and in the presence of Gaussian noise ($\alpha = 0.001$). The panel displays (a) the robustness r , (b) the transport efficiency σ , and (c) the cost c . The color scale at a given value of I_0 indicates the percentage Π of 250 trajectories at the corresponding value of the measure (dark blue is statistically irrelevant, dark red corresponds to more than 20%). The solid white line indicates the average for each value of I_0 . The solid purple line shows the deterministic value ($\alpha = 0$) for comparison. The network measures are calculated after applying the disparity filter (see text). Details on the numerical simulations are reported in Sec. III A.

average robustness, transport efficiency, and cost are generally increasing as a function for I_0 . For small I_0 ($I_0 \sim 1.5 \times 10^{-2}$), the mean values of the robustness, transport efficiency, and cost are lower than in the noiseless case. For larger values of I_0 ($I_0 \sim 5.5 \times 10^{-2}$), the mean value of robustness and transport efficiency slightly exceed the corresponding noiseless measures [see Figs. 3(a) and 3(b)], whereas the average cost is smaller. At values of I_0 outside the interval, these measures all saturate to their values in the noiseless case. Remarkably, there is a regime where the dynamics converges to more efficient network designs, namely, networks that are more robust and efficient, and at the same time less costly.

In order to identify the regimes where noise leads to more efficient topologies, we analyze the difference between the measures with and without noise over a broad range of demand strengths I_0 and noise amplitudes α . Figure 4 displays the differences $\Delta y = \langle y \rangle - y_0$ of the mean values over the noisy trajectories $\langle y \rangle$ from the noiseless measure y_0 for (a) robustness $y = r$, (b) transport efficiency $y = \sigma$, and (c) cost $y = c$. The differences are displayed as color plots in the I_0 - α plane. We first note that the contours of the equipotential lines follow an underlying linear behavior $I_0 \propto \alpha$. In fact, in the

presence of noise, Eq. (10) regains the invariance by I_0 when rescaling $\alpha \rightarrow \alpha I_0$. The plot indicates that all measures tend to the noiseless case for $I_0 \gg \alpha$ when the stochastic term becomes an infinitesimally small perturbation to the dynamics governed by Eq. (10).

It is tempting to assume that noise will lead to inferior networks. This is true for large enough noise. However, small noise can be beneficial. In the dark red part of the diagram the noisy solutions are more robust and efficient than the noiseless networks and the network topologies reach the largest values of r and σ . A cut of the plot at a fixed I_0 , Fig. 5(a), indicates a resonancelike behavior as a function of the noise amplitude α . Remarkably, in this region, the standard deviation of the robustness decreases, see Fig. 5(b). This behavior challenges the conventional expectation that the size of the fluctuations increase with α .

Following the criterion introduced in Ref. [18], we denote the region of increased robustness and efficiency as a noise-induced resonance. This terminology is borrowed from nonlinear dynamics, where noise-induced coherent effects have been reported in models simulating forest fires [22] and neurons [23]. As in stochastic resonance [21,23–25], we

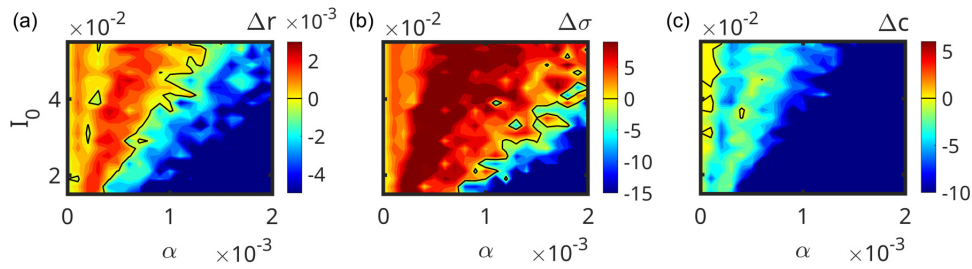


FIG. 4. Color plot of the difference between the measures of the stochastic and the noiseless networks as a function of I_0 and α for the two-norm activation function. The measures of the networks are evaluated after applying the disparity filter. Hot colors mean that the measure takes a larger value in the presence of noise, cold colors the opposite. Subplot (a) displays the difference $\Delta r = \langle r \rangle - r_0$ between the ensemble average $\langle r \rangle$ of the robustness and the value r_0 of the noiseless dynamics, (b) $\Delta \sigma = \langle \sigma \rangle - \sigma_0$ for the transport efficiency, and (c) $\Delta c = \langle c \rangle - c_0$ for the cost. The mean value of the stochastic dynamics is calculated over 50 trajectories for each value of I_0 and α . The solid black lines indicate the value where the difference exactly vanishes, the fluctuations are attributed to the relatively small number of trajectories.

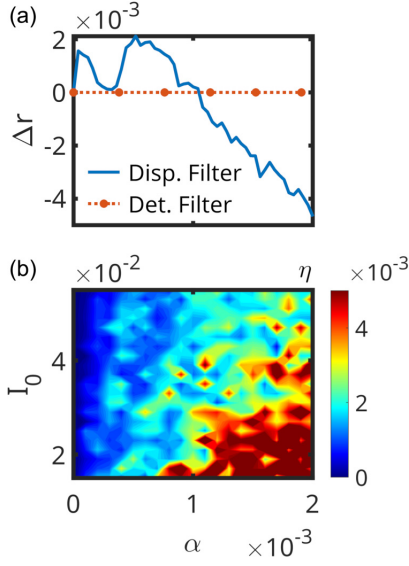


FIG. 5. (a) The solid line indicates the difference $\Delta r = \langle r \rangle - r_0$ between the ensemble average ($\langle r \rangle$) of the robustness and the value r_0 of the noiseless dynamics as a function of the noise amplitude α for $I_0 = 4.5 \times 10^{-2}$. Each data point represents the average value over 250 simulation runs. The red dots indicate the difference Δr after applying the deterministic filter to the noisy networks for a time $t = 2000 \gamma^{-1}$. (b) The standard deviation of the robustness distribution shown in Fig. 4 as a function of the demand strength I_0 and the noise amplitude α .

observe an optimal range of noise amplitudes at which noise increases the efficiency of the algorithm leading to more robust network topology than in the noiseless case. We note that these solutions are not fixed points of the noiseless dynamics, as we verify by taking them as initial conditions and determine the evolution according to Eq. (6). This procedure, which we refer to as the deterministic filter and detail in Appendix, shows that for sufficiently long integration times, $\Delta r \rightarrow 0$, as demonstrated by the red dots in Fig. 5(a). This suggests that the topologies obtained using the noisy dynamics are novel solutions, which are absent in the noiseless case.

B. Sigmoidal and ReLU functions

We now turn to the sigmoidal (Hill) and the ReLU function. Figure 6 displays the network measures as a function of I_0 and for fixed, nonvanishing noise amplitude α for three cases: in row (a) we report the network measures for the sigmoidal function with exponent $n = 1.6$, in row (b) for the sigmoidal function with exponent $n = 2$, and in row (c) for the ReLU function.

The behavior of the mean value as a function of the demand strength (white solid line) indicates that the measures generally increase with I_0 . The comparison with the noiseless measures (magenta curves) indicates the existence of multiple intervals of I_0 where noise leads to more robust networks. These networks simultaneously optimize transport, at the

expense of larger cost. At each I_0 , the trajectories are clustered about the fixed points of the noiseless case: The effect of noise is to increase the range of stability of certain topologies, extending the plateaus to lower values of I_0 (compare with Fig. 2).

The contour plots in Fig. 7 show the regions in the I_0 - α plane where noise leads to network topologies with larger robustness, transport efficiency, and cost (see dark red colored regions). For the sigmoidal function with $n = 2$ and the ReLU function, the size of these regions increases monotonously with α . Figure 8 illustrates the standard deviation of the corresponding robustness distribution: the variance does not seem to depend on α . A more detailed analysis actually shows that the standard deviation narrows in correspondence of the regions where noise leads to more robust networks. This indicates a resonancelike response to noise. We verified that the topologies found by the noisy dynamics do not depend on the filtering procedure we apply, including the deterministic filter. This shows that they are also solutions of the noiseless dynamics. Noise, in this case, modifies the landscape of local minima by favoring the more robust solutions.

In summary, we find that also for the sigmoidal and ReLU functions, noise can be used for optimizing robustness and efficiency of the emerging network topologies. This underscores the generality of our results, indicating that noise can enhance optimization algorithms for different classes of nonlinearities.

V. CONCLUSIONS

We have investigated the network topologies for a multicommodity flow problem and compared the robustness, transport efficiency, and cost of the emerging network topologies in the presence and in the absence of stochastic fluctuations. The networks have been calculated on a graph with a geometry of demands representing the relative locations of the major cities around Tokyo. The equations used for determining the topology consist of evolving the edge capacity according to Eq. (10). The growth of the edge capacity depends nonlinearly on the flow along the edge according to an activation function. Moreover, we have assumed that the edge capacity can additionally experience stochastic fluctuations.

We have analyzed different nonlinear functions of the flow and determined the behavior of the measures of the resulting self-organized networks as a function of the demand strength and the noise amplitude. The response to noise is different, yet for all the considered activation functions, we can identify parameter regimes where the interplay of noise and the nonlinear activation function gives rise to a resonancelike convergence to a more robust network topology. For the specific case of the two-norm function, the noisy dynamics converges to solutions that are otherwise unstable in the absence of noise. For the sigmoidal and the ReLU, instead, noise favors the most robust solutions of the deterministic dynamics. These results support and complement previous insights obtained with a simplified multicommodity flow problem consisting of two demands [18] and demonstrates their generality. It indicates that noise can be a resource for optimization algorithms.

Interestingly, the features of the noise-induced resonances can change dramatically depending on the activation function.

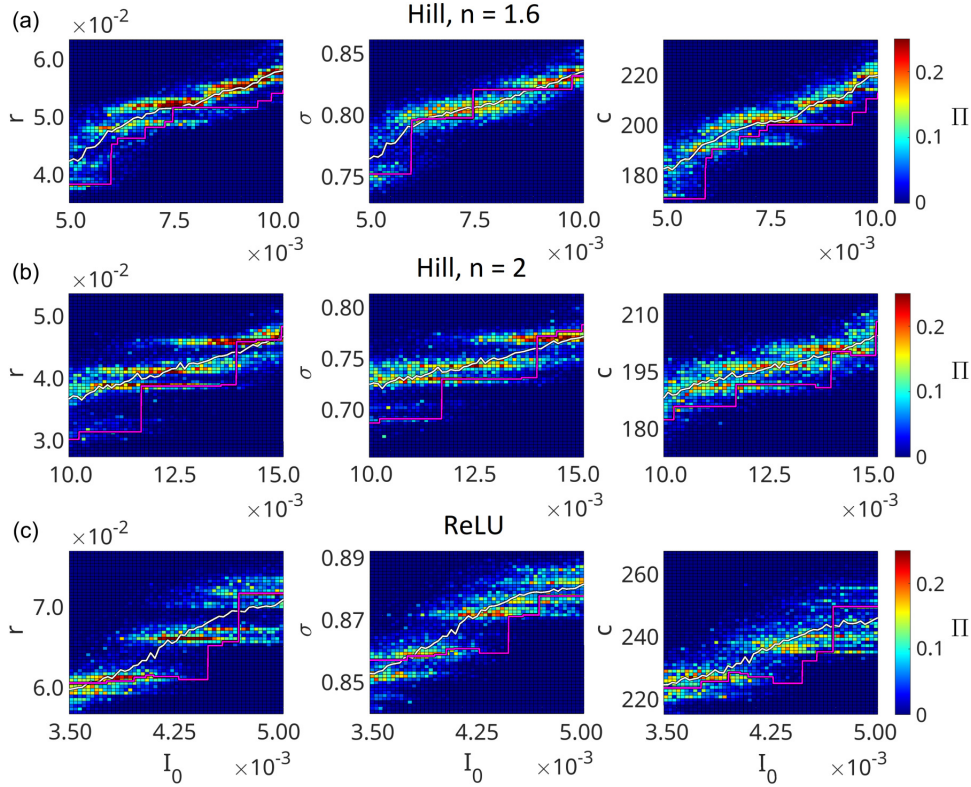


FIG. 6. Network measures as a function of the demand strength I_0 and for $\alpha = 0.001$. The panels in row (a) are obtained using the sigmoidal function with $n = 1.6$, in row (b) using the sigmoidal function with $n = 2$, and in row (c) using the ReLU function. The panels display (left) the robustness r , (central) the transport efficiency σ , and (right) the cost c . The color scale at a given value of I_0 indicates the percentage Π of 250 trajectories at the corresponding value of r (dark blue is statistically irrelevant, dark red corresponds to more than 20%). The solid white line indicates the average for each value of I_0 . The solid purple line shows the noiseless value and is plotted for comparison. The results of the noisy dynamics are invariant under the filtering procedure we apply. Details on the numerical simulations are reported in Sec. III A.

This observation suggests a different strategy for optimization, such that for a given noise and demand strength, one could identify a class of activation functions that maximizes the robustness and the transport efficiency of the network. Remarkably, the activation functions of this dynamics play a similar role in neural networks, where it is key for the expressivity [31]. In general, this study puts forward the need of identifying an adequate functional that permits one to systematically shed light on the interplay of nonlinearity and stochastic dynamics for optimization.

ACKNOWLEDGMENTS

The authors thank Philipp Hövel for stimulating discussions and acknowledge support from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) Project No. 429529648, TRR 306 QuCoLiMa (Quantum Cooperativity of Light and Matter), from the Bundesministerium für Bildung und Forschung (BMBF, German Ministry of Education and Research) under the Grant “NiQ: Noise in

Quantum Algorithms”, and in part by the National Science Foundation under Grants No. NSF PHY-1748958 and No. PHY-2309135.

APPENDIX: TESTING THE DISPARITY FILTER

We benchmark the disparity filter using two different procedures. In all cases, we first divide the time-averaged conductivity by the maximal edge conductivity $D_{\max} = \max_{(u,v) \in E_0} \langle D_{u,v} \rangle$, where E_0 is the set of all edges:

$$D_{u,v} \rightarrow D'_{u,v} \equiv \langle D_{u,v} \rangle / D_{\max}.$$

Disparity filter. We first apply a constant offset δ_1 , such that

$$D'_{u,v} \rightarrow D''_{u,v} + \delta_1.$$

We then apply the disparity filter as described in Sec. III B. The offset is introduced in order to level out fluctuations left after the time average. It is necessary to avoid that the filtering procedure artificially amplifies fluctuations due to noise. For

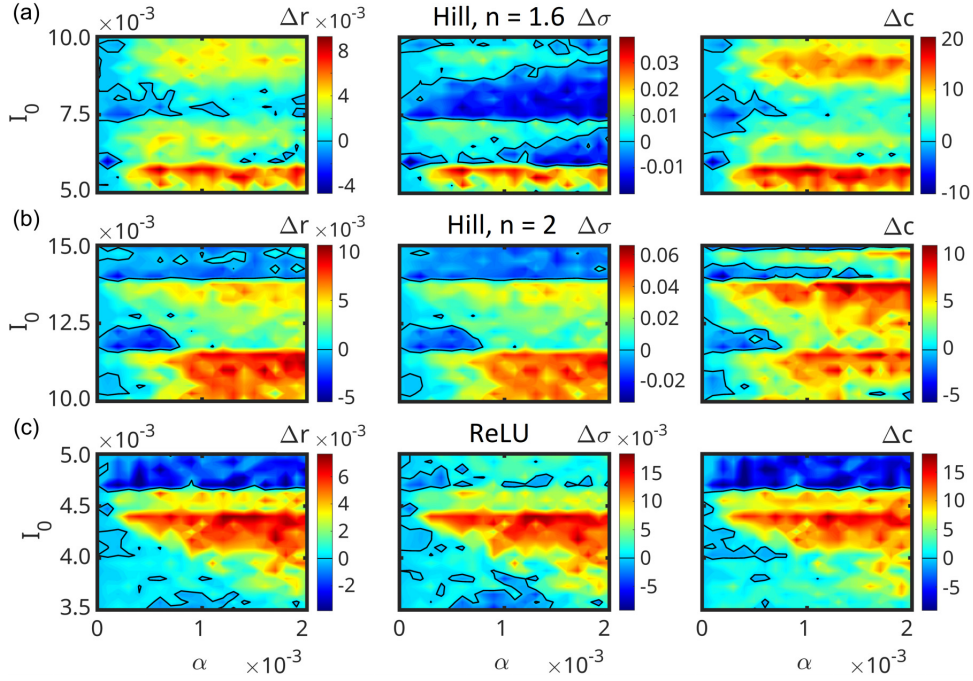


FIG. 7. Comparison between the measures of the stochastic and the noiseless networks as a function of I_0 and α when the activation function is [row (a)] the Hill function with $n = 1.6$, [row (b)] the Hill function with $n = 2$, and [row (c)] the ReLU function. The panels display (left) the robustness difference $\Delta r = \langle r \rangle - r_0$, (central) the transport efficiency difference $\Delta \sigma = \langle \sigma \rangle - \sigma_0$, and (right) the cost difference $\Delta c = \langle c \rangle - c_0$ between the corresponding ensemble averages $\langle r \rangle$, $\langle \sigma \rangle$, $\langle c \rangle$ of the stochastic case and the values r_0 , σ_0 , c_0 of the deterministic case. The mean value of the stochastic dynamics is calculated over 50 trajectories for each value of I_0 and α . The solid black lines indicate the value where the difference exactly vanishes. The fluctuations in these lines are attributed to the relatively small number of trajectories.

the given choice of d_i and for the considered values of α , we take $\delta_1 = 0.05$.

Cutoff filter. We consider the threshold $\delta_2 = 0.05$ and set to zero all conductivities with $D'_{u,v} < \delta_2$. The resulting network is composed of all nonvanishing edges.

“Deterministic” filter. We take $D'_{u,v}$ as initial conditions of a numerical integration according to the noiseless evolution of Eq. (6) and then verify whether it converges to the filtered network over the same integration time t_{end} .

In all cases, we remove any dead ends after the filter was applied. We show the comparative assessment for the robustness using the Hill function with $n = 1.6$. Figure 9 displays the robustness r as a function of the demand strength I_0 for the three different filtering procedures. We note that the qualitative behavior is the same for all filtering procedures, indicating that our results are independent of the specific choice of the filtering procedure.

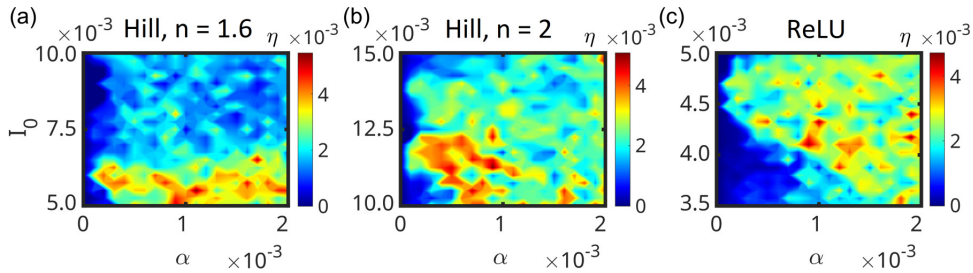


FIG. 8. The standard deviations of the robustness distributions shown in Fig. 7 as a function of the demand strength I_0 and the noise amplitude α . The subplots correspond to the different activation functions: (a) the sigmoidal with $n = 1.6$, (b) the sigmoidal with $n = 2$, and (c) the ReLU.

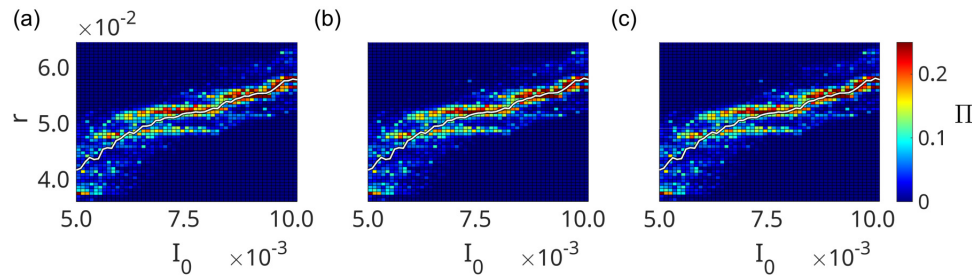


FIG. 9. The robustness r as a function of the demand strength I_0 for the case of using the filtering procedure based on (a) the disparity filter, (b) a cutoff filter with threshold $d = 0.05$, and (c) the deterministic filter, applied for a time $t = 1000 \gamma^{-1}$. For each value of I_0 , a total of 250 simulation runs were performed. The solid white line represents the ensemble average, and the color code indicates the probability distribution: dark blue is statistically irrelevant, and dark red corresponds to more than 20% of the simulation runs, yielding a network topology with the respective robustness. For all simulations, we chose $\alpha = 10^{-3}$. Details on the filtering procedures are given in the main text.

- [1] H. Yang and M. G. H. Bell, *Transport Rev.* **18**, 257 (1998).
- [2] C. Feremans, M. Labbé, and G. Laporte, *Eur. J. Oper. Res.* **148**, 1 (2003).
- [3] R. Z. Farahani, E. Miandoabchi, W. Szeto, and H. Rashidi, *Eur. J. Oper. Res.* **229**, 281 (2013).
- [4] C. H. Yeung and K. Y. M. Wong, *Phys. Rev. E* **80**, 021102 (2009).
- [5] J. R. Banavar, F. Colaiori, A. Flammini, A. Maritan, and A. Rinaldo, *Phys. Rev. Lett.* **84**, 4745 (2000).
- [6] A. Lonardi, E. Facca, M. Putti, and C. De Bacco, *Phys. Rev. Res.* **3**, 043010 (2021).
- [7] A. Lonardi, M. Putti, and C. De Bacco, *Sci. Rep.* **12**, 7474 (2022).
- [8] M. Dorigo, M. Birattari, and T. Stutzle, *IEEE Comput. Intell. Mag.* **1**, 28 (2006).
- [9] C. Oettmeier, T. Nakagaki, and H.-G. Döbereiner, *J. Phys. D* **53**, 310201 (2020).
- [10] T. Nakagaki, H. Yamada, and Á. Tóth, *Nature (London)* **407**, 470 (2000).
- [11] T. Nakagaki, H. Yamada, and M. Hara, *Biophys. Chem.* **107**, 1 (2004).
- [12] A. Tero, S. Takagi, T. Saigusa, K. Ito, D. P. Bebbler, M. D. Fricker, K. Yumiki, R. Kobayashi, and T. Nakagaki, *Science* **327**, 439 (2010).
- [13] L. Zhu, S.-J. Kim, M. Hara, and M. Aono, *R. Soc. Open Sci.* **5**, 180396 (2018).
- [14] C. Gao, C. Liu, D. Schenz, X. Li, Z. Zhang, M. Jusup, Z. Wang, M. Beekman, and T. Nakagaki, *Phys. Life Rev.* **29**, 1 (2019).
- [15] S. Li, H. Chen, M. Wang, A. A. Heidari, and S. Mirjalili, *Future Gener. Comput. Syst.* **111**, 300 (2020).
- [16] B. N. Örnek, S. B. Aydemir, T. Düzenli, and B. Özak, *Math. Comput. Simul.* **198**, 253 (2022).
- [17] V. Bonifaci, E. Facca, F. Folz, A. Karrenbauer, P. Kolev, K. Mehlhorn, G. Morigi, G. Shahkarami, and Q. Vermande, *Theor. Comput. Sci.* **920**, 1 (2022).
- [18] F. Folz, K. Mehlhorn, and G. Morigi, *Phys. Rev. Lett.* **130**, 267401 (2023).
- [19] A. Tero, R. Kobayashi, and T. Nakagaki, *J. Theor. Biol.* **244**, 553 (2007).
- [20] B. Meyer, C. Ansoorge, and T. Nakagaki, *PLoS One* **12**, 1 (2017).
- [21] F. Folz, K. Mehlhorn, and G. Morigi, *Phys. Rev. E* **104**, 054215 (2021).
- [22] E. Meron, *Phys. Rep.* **218**, 1 (1992).
- [23] B. Lindner, J. García-Ojalvo, A. Neiman, and L. Schimansky-Geier, *Phys. Rep.* **392**, 321 (2004).
- [24] L. Gammaitoni, P. Hänggi, P. Jung, and F. Marchesoni, *Rev. Mod. Phys.* **70**, 223 (1998).
- [25] M. c. v. Perc, *Phys. Rev. E* **72**, 016207 (2005).
- [26] H. Nakao, K. Arai, and Y. Kawamura, *Phys. Rev. Lett.* **98**, 184101 (2007).
- [27] S. Boccaletti, J. Kurths, G. Osipov, D. Valladares, and C. Zhou, *Phys. Rep.* **366**, 1 (2002).
- [28] C. Van den Broeck, J. M. R. Parrondo, and R. Toral, *Phys. Rev. Lett.* **73**, 3395 (1994).
- [29] F. Sagués, J. M. Sancho, and J. García-Ojalvo, *Rev. Mod. Phys.* **79**, 829 (2007).
- [30] Y. Bahri, J. Kadmon, J. Pennington, S. S. Schoenholz, J. Sohl-Dickstein, and S. Ganguli, *Annu. Rev. Condens. Matter Phys.* **11**, 501 (2020).
- [31] I. Gühring, M. Raslan, and G. Kutyniok, Expressivity of deep neural networks, [arXiv:2007.04759](https://arxiv.org/abs/2007.04759).
- [32] W. Ellens and R. Kooij, [arXiv:1311.5064](https://arxiv.org/abs/1311.5064).
- [33] W. Ellens, F. Spieksma, P. Van Mieghem, A. Jamakovic, and R. Kooij, *Linear Algebra Appl.* **435**, 2491 (2011).
- [34] N. VAN KAMPEN (ed.), *Stochastic Processes in Physics and Chemistry*, 3rd ed., North-Holland Personal Library (Elsevier, Amsterdam, 2007).
- [35] P. E. Kloeden and E. Platen, *Numerical Solution of Stochastic Differential Equations* (Springer-Verlag Berlin, Heidelberg, 1992).
- [36] T. D. Frank, *Nonlinear Fokker-Planck Equations* (Springer-Verlag, Berlin, Heidelberg, New York, 2005).
- [37] M. A. Serrano, M. Boguna, and A. Vespignani, *Proc. Natl. Acad. Sci. USA* **106**, 6483 (2009).

PART 2: UNPUBLISHED MATERIAL

In Part 1 of this chapter, we applied the Physarum-inspired model defined by Eqs. (4.4), (4.5) and (4.10) to the transport problem of routing multi-commodity flows through the greater region of Tokyo. We analyzed how the interplay of nonlinear dynamics and stochastic forces influences the resulting network topologies and studied their robustness, efficiency, and cost. However, key topological properties—such as the occurrence of loops, which are strongly linked to the robustness, efficiency and cost—remain to be investigated. Specifically, an essential question arises: *How do the noise amplitude α and the demand strength I_0 influence the formation and characteristics of loops, and how are these changes related to the trends in robustness, efficiency, and cost?* To address this, based on the same model and activation functions introduced in Part 1, we next examine the number of loops n_c , their length l_c , and the cycle edge fraction e_c , as defined in Sec. I.1.3.

4.1 Two-norm function

Let us first consider the two-norm as the activation function. Figure 10 shows (a) the number of loops n_c , (b) their length l_c , and (c) the cycle edge fraction e_c as a function of the demand strength I_0 . In the noiseless case, all three measures remain constant. This is a consequence of the fact that the two-norm is a homogeneous function of the flow with degree 1. In striking contrast, for the stochastic case with a fixed noise amplitude $\alpha = 0.001$, the measures change as a function of I_0 : on average, the number of loops n_c and the cycle edge fraction e_c increase with I_0 whereas the length of the loops l_c decreases. These trends are consistent with the analysis in Part 1, which showed that robustness and efficiency increase with I_0 . As discussed in Sec. I.1.3, loops inherently enhance network robustness and efficiency as they provide redundant paths between nodes. Therefore, the observed increase in the number of loops is directly related to the rise in robustness and efficiency.

Interestingly, in the range of I_0 that we consider, the average values of the number of loops n_c and the fraction e_c are below the corresponding noiseless case. Conversely, the average value of the length of the loops l_c surpasses the deterministic one, indicating that noise tends to establish fewer but larger loops. However, for sufficiently large I_0 , the averages of all measures converge to the noiseless values, demonstrating that the deterministic behavior dominates the effect of noise in the limit $I_0/\alpha \rightarrow \infty$. This behavior occurs because the two-norm is a non-saturating function: in Eq. (4.10), which models the dynamics of the edge capacities, the magnitudes of the deterministic terms grow without bound as I_0 increases, while the noise term remains independent of I_0 . These results also explain the lowered network cost in the stochastic case reported in Part 1, particularly for small I_0 . This cost reduction is mainly caused by the significant decrease in the number of loops, which corresponds to a removal of redundant connections that is only marginally compensated by the slightly enhanced loop length. A similar argument holds for the fraction e_c , which is almost identical for the stochastic and the deterministic case but exhibits slightly larger noiseless values for small demand strengths.

To shed light on the role of the noise amplitude, we analyze the combined influence

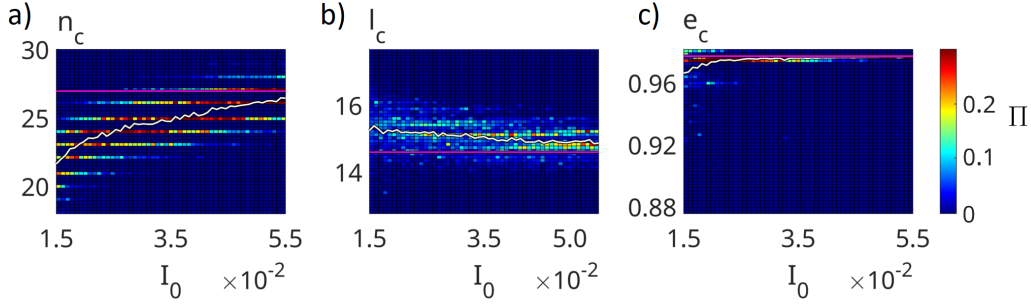


Figure 10: Loops of noisy networks generated with the two-norm activation function for $\alpha = 0.001$, displayed as a function of the demand strength I_0 . The subplots display (a) the number of loops n_c , (b) their length l_c and (c) the fraction of edges e_c that are part of at least one loop. The color scale at a given value of I_0 indicates the fraction Π of trajectories at the corresponding value: dark blue is statistically irrelevant, and dark red corresponds to more than 20% (see colorbar). The solid white line indicates the average over 250 simulation runs for each value of I_0 and the solid purple line indicates the deterministic values.

of I_0 and α on the loop-based network measures. Specifically, we quantify the difference $\Delta y = \langle y \rangle - y_0$ between the average $\langle y \rangle$ over the noisy trajectories and the deterministic value y_0 , as introduced in Part 1. Figure 11 shows Δy as a function of I_0 and α for the number of loops n_c , their length l_c and the fraction e_c . The previously observed trends for a fixed value of $\alpha = 0.001$ are confirmed across the whole range of noise amplitudes: n_c and e_c increase with I_0 , whereas l_c decreases. Further, the number of loops n_c and the fraction e_c decrease with the noise amplitude α , while the loop length l_c increases. As the relative change in n_c is larger compared to l_c , these results suggest that stronger stochastic forces reduce the number of redundant edges in the network, consistent with the drop in robustness, efficiency, and cost for large α discussed in Part 1.

Interestingly, the equipotential lines in Fig. 11, represented by contours of equal color, follow the linear relation $I_0 \propto \alpha$, mirroring the scaling of the robustness, transport efficiency, and cost discussed in Part 1. Further, for $I_0/\alpha \gg 1$, which corresponds to equipotential lines with a large slope, the differences between the stochastic and the deterministic values disappear, as the fluctuations become less and less relevant. Analogous to the case of a fixed noise amplitude $\alpha = 0.001$, this behavior arises from the deterministic terms in Eq. (4.10) dominating the stochastic term.

For small noise amplitudes, the number of loops n_c and their length l_c exhibit several less pronounced local maxima and minima, respectively, which resemble noise-induced resonances. These resonances indicate that certain noise amplitudes locally optimize specific topological properties of the network. For moderate to large α , the presence of stochastic forces consistently tends to establish fewer but larger loops compared to the noiseless case, which is consistent with the behavior observed for $\alpha = 0.001$.

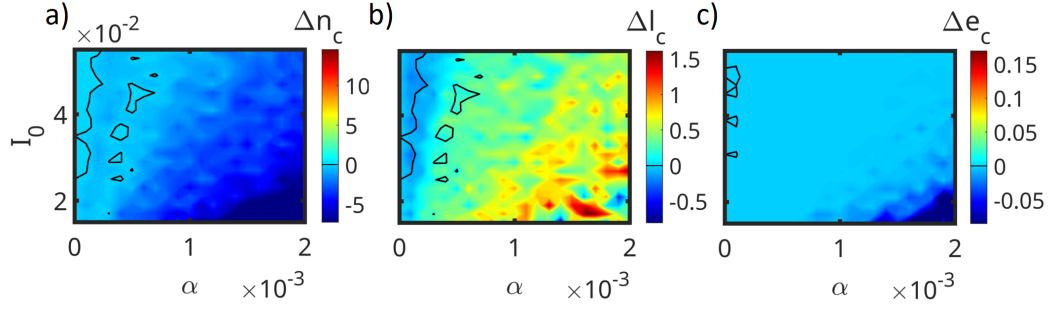


Figure 11: (a) The difference $\Delta y = \langle y \rangle - y_0$ between the average $\langle y \rangle$ over the noisy trajectories and the deterministic value y_0 as a function of the demand strength I_0 and the noise amplitude α for (a) the number of loops n_c , (b) their length l_c and (c) the fraction e_c . For the stochastic case, we considered 50 simulation runs for each value of I_0 and α . We use the two-norm as the activation function and display Δy as a color code. Hereby, dark blue corresponds to a reduction of the measures due to the presence of noise, whereas dark red indicates a significant increase (see colorbar). The solid black lines represent $\Delta y = 0$, indicating that the average of the stochastic case and the deterministic value are equivalent. The fluctuations of the solid black lines are attributed to the relatively small number of trajectories.

4.2 Sigmoidal and ReLU functions

We now consider the dynamics when the activation function is the sigmoidal (Hill) with $n = 1.6$ or $n = 2$ or the ReLU. Figure 12 shows the number of loops n_c , their length l_c and the cycle edge fraction e_c as a function of the demand strength I_0 . In the noiseless case, all measures exhibit a steplike behavior, unlike the two-norm, where the measures are independent of I_0 . This behavior arises from the non-homogeneity of the sigmoidal and the ReLU functions. In general, the number of loops n_c and the fraction e_c tend to increase with I_0 . The loop length l_c , however, shows a more intricate behavior as a function of I_0 , which strongly depends on the activation function: for the sigmoidal, it tends to decrease when $n = 1.6$, while it increases when $n = 2$. For the ReLU, instead, l_c reaches a global minimum in the center of the considered interval of I_0 . In the stochastic case with a fixed noise amplitude $\alpha = 0.001$, the steplike behavior of the noiseless case disappears due to averaging over an ensemble of trajectories with varying topological properties. Despite this, general trends for the number of loops n_c and the fraction e_c persist, which continue to increase with I_0 . However, the behavior of the loop length l_c differs for certain activation functions compared to the noiseless case. For the sigmoidal function with $n = 1.6$, it decreases with I_0 , whereas for $n = 2$, it stays approximately constant. For the ReLU function, loops tend to be shorter for smaller I_0 and longer for larger I_0 , although the loop length l_c varies only slightly overall.

Remarkably, for all measures, there are regions where the averages over the noisy trajectories exceed the corresponding noiseless values. This behavior arises from stochastic forces broadening the stability range of fixed points of the noiseless dynamics, which is also reported for the robustness, efficiency, and cost in Part 1. This effect is particularly evident for the measure e_c , where the plateaus observed in the deterministic case are considerably extended towards smaller values of I_0 .

To examine the influence of the noise amplitude, Fig. 13 displays the differences Δy

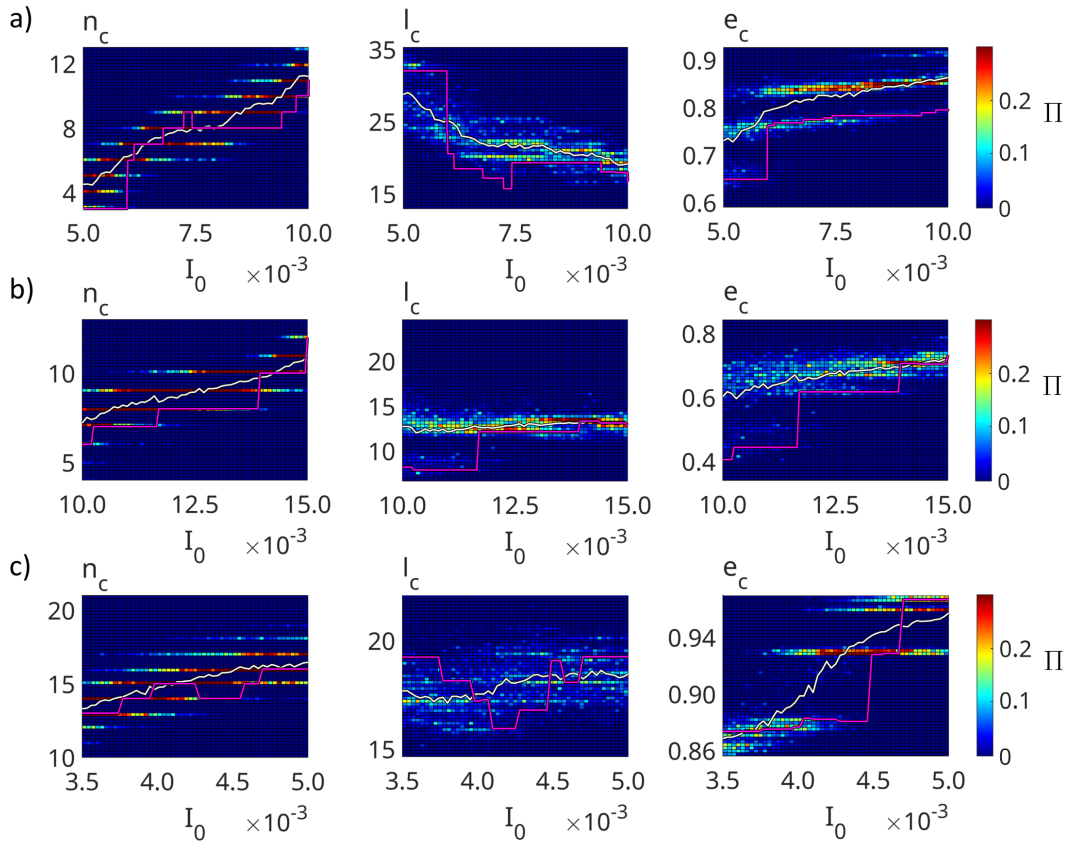


Figure 12: The number of loops n_c (left), their length l_c (central), and the cycle edge fraction e_c (right) as a function of the demand strength I_0 for $\alpha = 0.001$ and (a) the Hill function with $n = 1.6$, (b) the Hill function with $n = 2$, and (c) the ReLU function as the activation function. The color scale at a given value of I_0 indicates the fraction Π of trajectories at the corresponding value of r : dark blue is statistically irrelevant, and dark red corresponds to more than 20% (see colorbar). The solid white line indicates the average over 250 simulation runs for each value of I_0 and the solid purple line the deterministic values.

as a function of I_0 and α , where y represents the number of loops n_c , their length l_c and the fraction e_c . These differences exhibit a steplike behavior as a function of I_0 , which is inherited from the noiseless dynamics. In the red regions, the averages over the noisy trajectories exceed the corresponding noiseless values. These regions extend across a wide range of noise amplitudes, starting at small α , and arise from the broadening of the deterministic plateaus towards lower I_0 . Specifically, for the sigmoidal function with $n = 2$, the number of loops n_c displays pronounced red regions that widen as the noise amplitude α increases. Similarly, for the ReLU function, the fraction e_c exhibits an increasingly large range of I_0 , where the stochastic averages exceed the noiseless values. These regimes take the form of noise-induced resonances, demonstrating that finite noise amplitudes can maximize loop number and length, for example.

Remarkably, the relationship between the number of loops and their length varies significantly depending on the activation function and the demand strength I_0 . For the sigmoidal function with $n = 1.6$, regions with more loops are associated with

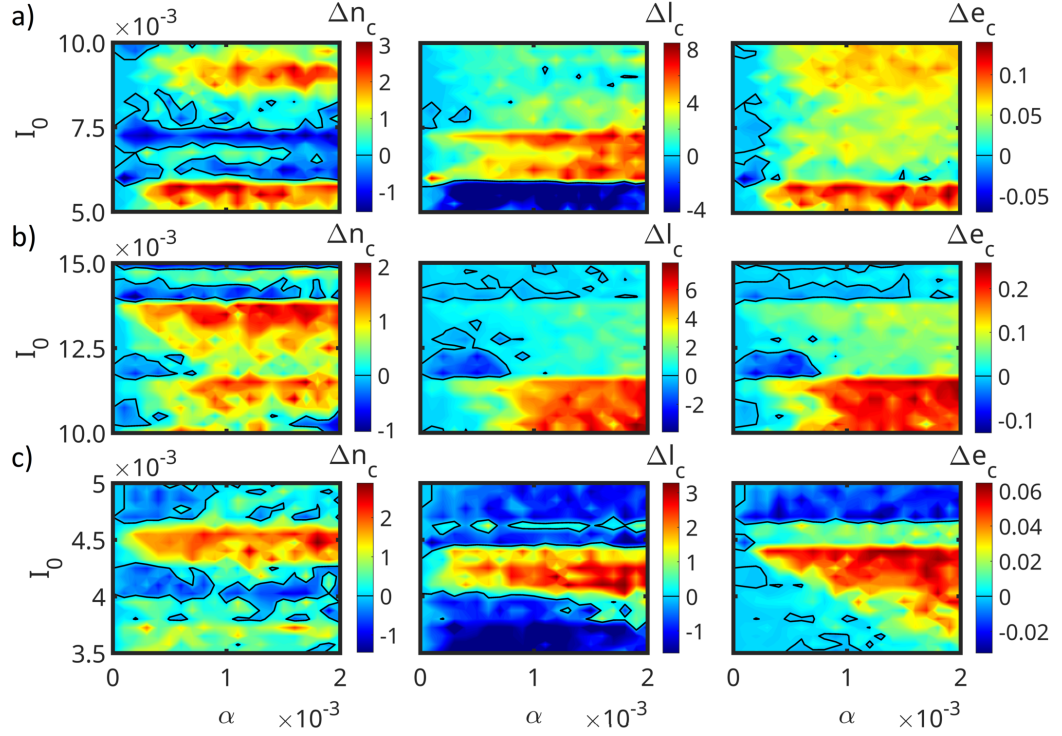


Figure 13: The difference $\Delta y = \langle y \rangle - y_0$ between the average $\langle y \rangle$ over the noisy trajectories and the deterministic value y_0 as a function of the demand strength I_0 and the noise amplitude α for (a) the Hill function with $n = 1.6$, (b) the Hill function with $n = 2$ and (c) the ReLU as the activation function. The panels show the difference in the number of loops n_c (left), their length l_c (central), and the cycle edge fraction e_c (right). For the stochastic case, we considered 50 simulation runs for each value of I_0 and α . The differences Δy are displayed as a color code, whereby dark blue corresponds to decreased values and dark red indicates increased values (see colorbar). The solid black lines represent $\Delta y = 0$, indicating that the average of the stochastic case and the deterministic value are equivalent. The fluctuations of the solid black lines are attributed to the relatively small number of trajectories.

shorter loops at small I_0 . In contrast, for large I_0 , regions with more loops exhibit longer loops. The fraction e_c of edges that are involved in at least one loop consistently exceeds the deterministic value, as noise increases either the number of loops or their length. For the sigmoidal function with $n = 2$, regions with a larger number of loops coincide with longer loops across the whole range of I_0 , which also leads to an increased cycle edge fraction e_c . For the ReLU function, regions with more loops generally exhibit shorter loops. Interestingly, the most significant gain in loop length occurs in a region where the number of loops n_c in the stochastic and deterministic cases is almost identical. At large I_0 , noise has a minimal effect on the number of loops n_c , but reduces their length l_c , limiting the increase in the fraction e_c . These results suggest that, under certain conditions, noise can promote the formation of loops and their length selectively. Overall, the qualitative behavior of the differences Δy in the loop-based measures is reminiscent of the trends reported for robustness, efficiency, and cost in Part 1. However, deviations arise because these properties are influenced by the intricate interplay of the number of loops, their

length, and the cycle edge fraction, which may be independent of each other or not, depending on the activation function and the demand strength.

CONCLUSION AND OUTLOOK

C.1 Summary

In this dissertation, we have presented studies on network self-organization. Spanning four interrelated publications, the dissertation progresses from analyzing a biologically-inspired model for finding shortest paths to generalizing it to multi-commodity flow settings. As a key result, we found that moderate noise can dramatically increase the efficiency of the algorithm. Specifically, we reported that the adaptability of the emerging networks to periodically changing environmental conditions is maximized in the presence of noise due to stochastic resonance. Further, we found self-organization into networks with optimal robustness, efficiency, and cost at finite noise amplitudes, where the convergence speed of the algorithm is maximized as well. This peak in system performance at optimal noise levels, accompanied by a narrowing of the distributions of the stochastic trajectories, is characteristic of noise-induced resonances. Moreover, we analyzed how the interplay between nonlinear dynamics and stochastic forces drives the emergence of these resonances. We demonstrated that they strongly depend on the algorithm's nonlinearity while remaining observable across all considered classes of activation functions.

In general, our results challenge the traditional view of noise as being detrimental; instead, they suggest that it can even be harnessed as a resource. This offers new opportunities for real-world applications, especially for the design of transport networks. By leveraging noise-induced effects, one might be able to tailor specific topological properties, such as efficiency and robustness, by promoting shortest-path-like connections or loops, respectively. Consequently, noise emerges as a parameter to optimize network performance. Based on these insights, the following section proposes further research directions for network design.

C.2 Future directions in network design

For instance, future work could introduce edge-specific dissipation rates in the Physarum-inspired model for network design, simulating geographical constraints, for example. As a consequence, the algorithm would reduce the usage of edges with larger dissipation. In addition to that, time-dependent dissipation rates could be explored to model dynamically changing constraints, similar to the setup discussed in Chapter 1 for a single commodity. Another promising direction is to consider time-varying demand strengths to reflect variations in traffic loads—an approach recently explored by Lonardi et al. [97].

These extensions would broaden the applicability of the algorithm considerably and allow for a detailed study of the effect of stochasticity in more complex settings. We expect that noise could also prove beneficial in such scenarios, potentially enhancing the minimization of dissipation and the adaptability to time-dependent constraints. In addition to these promising research directions in the context of Physarum-inspired algorithms, the idea of harnessing noise and nonlinear dynamics extends beyond network design. For this reason, the following section discusses

how stochastic fluctuations might be leveraged in other, well-known optimization problems.

C.3 Adding noise and nonlinear dynamics to established optimization algorithms

The Physarum-inspired approaches considered in this dissertation can be regarded as reference models for optimization algorithms, which play a crucial role in various disciplines. Given that adding noise enhances the emerging networks in the Physarum-inspired model, an intriguing question arises: *Can the combination of stochasticity and nonlinearity also benefit well-known optimization algorithms, such as gradient-based approaches?* These algorithms, widely used in fields like machine learning [98–100], often face challenges such as slow convergence or getting trapped in local minima, particularly for high-dimensional optimization problems. In fact, noise is already used in several optimization strategies to improve performance. For instance, simulated annealing [101, 102] leverages stochasticity to navigate through intricate optimization landscapes. Similarly, Eisert et al. [103] demonstrated that adding noise to the standard gradient descent method significantly improves its convergence in the context of quantum computing. However, while they focus solely on noise, the modifications proposed in this dissertation include both nonlinear dynamics and stochastic forces, which may further enhance the performance of optimization algorithms. To explore this approach, we analyze several use cases based on the standard gradient descent algorithm. The following section provides a summary of our ansatz and preliminary findings.

C.3.1 Standard gradient descent

In this section, we investigate how adding nonlinear dynamics and noise to the standard gradient descent method [104], a commonly used approach to minimize a cost function $C(x_1, \dots, x_k)$, influences its ability to reach the global minimum of C . In its original form, the algorithm starts at the initial values (x_1^0, \dots, x_k^0) and updates each variable x_j according to the rule

$$x_j^{i+1} = x_j^i - \eta \frac{\partial C}{\partial x_j} (x_1^i, \dots, x_k^i), \quad (5.1)$$

with $j \in \{1, \dots, k\}$, the iteration i and the learning rate η . For sufficiently small η , Eq. (5.1) converges to the system of differential equations

$$\frac{\partial x_j}{\partial t} = - \frac{\partial C}{\partial x_j}. \quad (5.2)$$

Its steady state is given by

$$\frac{\partial C}{\partial x_j} = 0, \quad (5.3)$$

which is a necessary condition for x_j to be at a local minimum of the cost function. Since x_j is updated along the negative gradient of C , the algorithm guarantees convergence to a local minimum, but not necessarily to the global minimum.

As an example, let us consider the one-dimensional cost function

$$C(x) = (x - 1)^2 + \sin(10x), \quad (5.4)$$

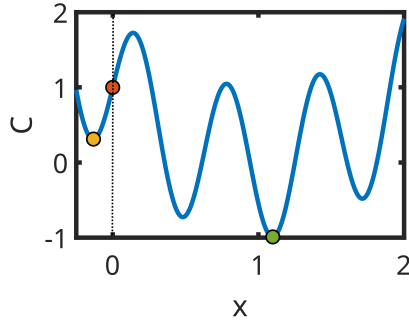


Figure 14: Cost function $C(x)$ given by Eq. (5.4). The red dot marks the initial value $C_0 = C(x_0)$, with the dotted line indicating x_0 . The standard gradient descent method as defined by Eq. (5.1) was applied for a fixed number of $m = 10,000$ iterations using the learning rate $\eta = 0.01$. The algorithm converged to the nearest local minimum, marked by the yellow dot. The global minimum is labeled by the green dot.

which has an intricate landscape with multiple local minima due to the sinusoidal term. Its global minimum is located at $x_{\min} \approx 1.099$, corresponding to a cost of $C_{\min} \approx -0.99$. Starting from an initial value x_0 , the algorithm defined by Eq. (5.1) follows the steepest descent path. However, as illustrated in Fig. 14, it converges to the nearest local minimum rather than the global minimum, demonstrating a common limitation of gradient-based methods.

To address this issue, we propose several variants of the approach given by Eq. (5.1), incorporating nonlinear dynamics and stochastic forces. While we use Eq. (5.4) for benchmarking these modifications, the algorithms are formulated for the general case of a k -dimensional cost function.

C.3.1.1 Nonlinear gradient descent with fixed noise amplitude

In a first attempt, we modify the gradient descent method by adjusting the update rule to

$$x_j^{i+1} = x_j^i - \eta \operatorname{sgn} \left(\frac{\partial C}{\partial x_j} (x_1^i, \dots, x_k^i) \right) f \left(\left| \frac{\partial C}{\partial x_j} (x_1^i, \dots, x_k^i) \right| \right) + \alpha \sqrt{\eta} \xi_j^i. \quad (5.5)$$

Here, $\operatorname{sgn}(x)$ is the sign function, $f(x)$ is a positive, nonlinear activation function, and ξ_j^i is Gaussian white noise, with α capturing the amplitude of the stochastic fluctuations. The sign function guarantees that, in the noiseless case ($\alpha = 0$), the updates to x_j decrease the value of the cost function. However, due to the component-wise application of the activation function $f(x)$, the trajectories do not necessarily follow the steepest descent anymore. The noise enables the system to escape from local minima by deviating from the deterministic descent path. For $\eta \ll 1$, Eq. (5.5) converges to the system of stochastic differential equations

$$\frac{\partial x_j}{\partial t} = - \operatorname{sgn} \left(\frac{\partial C}{\partial x_j} \right) f \left(\left| \frac{\partial C}{\partial x_j} \right| \right) + \alpha \sqrt{\eta} \xi_j(t). \quad (5.6)$$

The steady state of its deterministic part is given by

$$f \left(\left| \frac{\partial C}{\partial x_j} \right| \right) = 0. \quad (5.7)$$

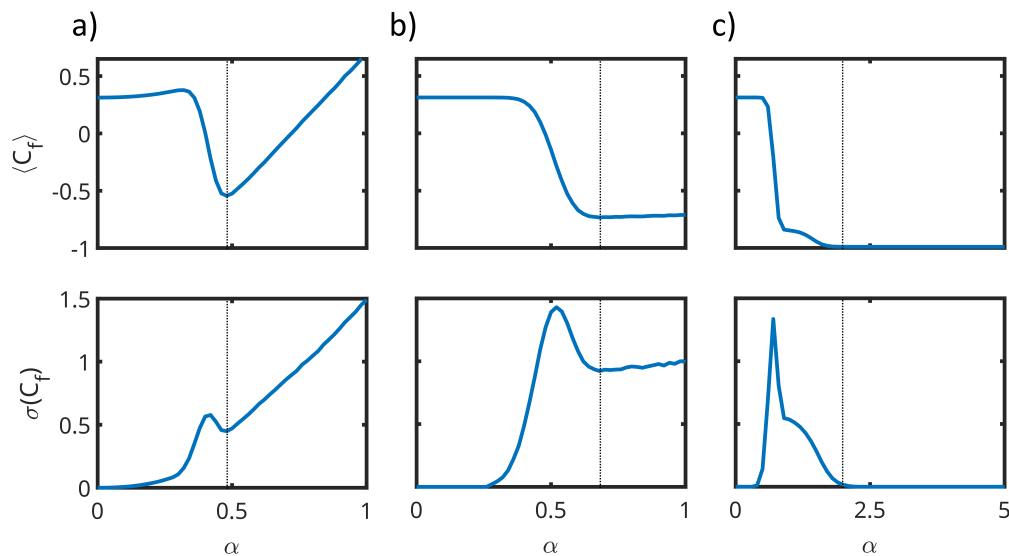


Figure 15: Convergence behavior of different algorithms as a function of the noise amplitude α . The top panel shows the average final cost $\langle C_f \rangle$ and the bottom panel displays its standard deviation $\sigma(C_f)$, both determined for a total of 100,000 simulation runs with a learning rate of $\eta = 0.01$. The subfigures employ different update rules: (a) Eq. (5.5), with fixed noise amplitude, (b) Eq. (5.5), with a noise amplitude that linearly decreases over time, as described by Eq. (5.9), and (c) Eq. (5.13). The dotted lines indicate the smallest noise amplitude at which $\langle C_f \rangle$ reaches its minimum.

To ensure that the necessary condition for a local minimum, Eq. (5.3), is satisfied, we impose the requirement

$$\begin{aligned} f(x) &= 0 \\ \Leftrightarrow x &= 0 \end{aligned} \quad (5.8)$$

on the activation function. In the following, we use $f(x) = \sqrt{x}$, which enhances update steps for small magnitudes of the gradient while reducing them for larger ones (note that $x = |\partial_{x_j} C| \geq 0$). This choice improves the algorithm's ability to navigate through flat landscapes and prevents overshooting. Figure 15(a) displays the convergence of the algorithm for varying noise amplitudes α . For small noise amplitudes, the average final cost $\langle C_f \rangle$ first increases slightly before decreasing sharply. It reaches a minimum value of $\langle C_f \rangle \approx -0.5$ at $\alpha \approx 0.5$, which is marked by the dotted line. Although this minimum is lower than the final cost achieved by the standard gradient descent method, it remains significantly higher than the global minimum of $C_{\min} \approx -0.99$. As the noise amplitude increases further, $\langle C_f \rangle$ rises again, indicating the presence of a noise-induced resonance at which the algorithm's performance is maximized. This notion is further strengthened by the standard deviation $\sigma(C_f)$ reaching a local minimum at this optimal noise level.

To sum up, the approach given by Eq. (5.5) enhances the convergence of the algorithm to the global minimum, but it remains suboptimal. While stochastic fluctuations enable the system to escape from local minima, they also hinder settling to the global minimum. To address this, the following section introduces a dynamic noise amplitude.

C.3.1.2 Nonlinear gradient descent with linearly decreasing noise amplitude

Optimization algorithms inspired by simulated annealing [101, 102] typically employ stochastic fluctuations with an amplitude that decreases over time, representing a lowering of the temperature of the system. While this cooling procedure still facilitates the escape from local minima during the initial iterations, it also enhances convergence to the global minimum as the system stabilizes. Based on this concept, we further modify our approach by considering a noise amplitude that decreases linearly over time. Specifically, in Eq. (5.6), we replace

$$\alpha \rightarrow \alpha^i = (m - i)/m \alpha, \quad (5.9)$$

where m is the total number of iterations. While the actual noise level α^i depends on the current iteration i , we refer to the constant parameter α as the noise amplitude in the following. Figure 15(b) illustrates the results of this modification. In contrast to the approach outlined in Eq. (5.6), both $\langle C_f \rangle$ and $\sigma(C_f)$ remain constant for small noise amplitudes. For larger α , the average final cost decreases, whereas the standard deviation increases. At $\alpha \approx 0.7$, both $\langle C_f \rangle$ and $\sigma(C_f)$ reach a minimum, which extends over a wide range of noise amplitudes. This minimum, $\langle C_f \rangle \approx -0.7$, is significantly lower than the minimum achieved using the update rule defined by Eq. (5.5), which employs a fixed noise level. For values of α larger than the considered range, we expect both $\langle C_f \rangle$ and $\sigma(C_f)$ to rise again.

While a linearly decreasing noise amplitude enables the algorithm to converge to values closer to the global minimum on average, the relatively large variance in the final value C_f indicates that the system frequently gets stuck in local minima. To mitigate this issue, the following section implements additional modifications to the update rule.

C.3.1.3 Nonlinear gradient descent with adaptive noise control

We attempt to further enhance the convergence of the optimization algorithm to the global minimum of the cost function by introducing adaptive noise control. We use an approach inspired by the Metropolis algorithm [105], a widely used Monte Carlo method. Specifically, we modify the algorithm so that it probabilistically switches between deterministic and stochastic update steps based on the local behavior of the cost function. In particular, we explicitly allow intermediate update steps that increase the value of the cost function to facilitate the escape from local minima. To adapt this concept for a gradient descent method, we define a candidate for the updated value of the next step

$$\tilde{x}_j^i(\alpha) = x_j^i - \eta \operatorname{sgn} \left(\frac{\partial C}{\partial x_j} (x_1^i, \dots, x_k^i) \right) f \left(\left| \frac{\partial C}{\partial x_j} (x_1^i, \dots, x_k^i) \right| \right) + \alpha \sqrt{\eta} \xi_j^i \quad (5.10)$$

and the probability

$$p^i = e^{-\Delta C^i / (\alpha^i)^2} \quad (5.11)$$

to realize this step, with the difference in the cost function

$$\Delta C^i = C(\tilde{x}_1^i(\alpha), \dots, \tilde{x}_k^i(\alpha)) - C(x_1^i, \dots, x_k^i). \quad (5.12)$$

Note that in Eq. (5.10), the value of α is fixed, whereas in Eq. (5.11), we consider a linearly decreasing noise amplitude as given by Eq. (5.9). We modify the update

rule given by (5.5) to

$$x_j^{i+1} = \begin{cases} \tilde{x}_j^i(\alpha), & \text{if } q^i < \min(1, p^i) \\ \tilde{x}_j^i(0), & \text{else} \end{cases} \quad (5.13)$$

with a uniformly distributed random number $q^i \in [0, 1]$. Figure 15(c) shows the convergence of this algorithm. For small values of α , both the average final cost $\langle C_f \rangle$ and the standard deviation $\sigma(C_f)$ are constant. For larger noise amplitudes, $\langle C_f \rangle$ sharply decreases to an intermediate plateau before reducing further to a minimum of $\langle C_f \rangle \approx -0.99$, which extends over a wide range of noise amplitudes. Remarkably, this is the global minimum C_{\min} of the cost function, demonstrating the superiority of this algorithm over the previously considered approaches, including the standard gradient descent method. Correspondingly, the standard deviation first increases significantly, reaches a maximum, and then decreases again. For values of α at which $\langle C_f \rangle$ reaches the global minimum of the cost function, the standard deviation vanishes, indicating that all trajectories converge to C_{\min} .

Overall, the approach defined by Eqs. (5.10) - (5.13), which effectively combines deterministic and stochastic updates, enables consistent convergence to the global minimum.

C.3.2 Concluding remarks

To conclude, these preliminary studies suggest that the performance of well-known optimization algorithms, such as the standard gradient descent method, can be enhanced by adding nonlinear dynamics and stochastic forces. In particular, the modifications introduced in Sec. C.3.1.3 enable the algorithm to navigate through intricate landscapes more effectively, avoid getting trapped in local minima, and reach the global minimum of a cost function with greater reliability.

These preliminary findings hold significant potential for fields where optimization is central, such as machine learning and quantum computing. For instance, the Variational Quantum Eigensolver [106, 107] often faces the problem of barren plateaus [108], where the landscape of the cost function becomes extremely flat, posing a major challenge to optimization. While introducing nonlinearity and noise alone may not overcome this issue, these modifications could potentially complement other strategies to mitigate such problems.

Future research could examine how the proposed modifications influence the convergence speed of the algorithms, with the results discussed in previous chapters suggesting an intricate dependency. Overall, this proof of concept demonstrates that the insights gained in this dissertation are not limited to the design of transport networks but hold broader applicability to general optimization challenges.

REFERENCES

- [1] P. W. Anderson, *Science* **177**, 393–396 (1972).
- [2] M. Mitchell, *Complexity* (Oxford University Press, New York, NY, USA, 2011).
- [3] J. H. Holland, *Adaptation in natural and artificial systems* (Bradford Books, Cambridge, MA, USA, 2019).
- [4] A.-L. Barabási and J. Frangos, *Linked* (Perseus Books, Boulder, CO, USA, 2003).
- [5] S. P. Borgatti, A. Mehra, D. J. Brass, and G. Labianca, *Science* **323**, 892–895 (2009).
- [6] J. Portenoy, J. Hullman, and J. D. West, *Frontiers in Research Metrics and Analytics* **2**, 8 (2017).
- [7] G. Luo, Z. Zhang, and S. Diao, *Information Sciences* **608**, 1202–1220 (2022).
- [8] D. Centola, *Science* **329**, 1194–1197 (2010).
- [9] J. A. Dunne, R. J. Williams, and N. D. Martinez, *Marine Ecology Progress Series* **273**, 291–302 (2004).
- [10] J. A. Dunne, R. J. Williams, N. D. Martinez, R. A. Wood, and D. H. Erwin, *PLoS Biology* **6**, e102 (2008).
- [11] S. Allesina and M. Pascual, *Theoretical Ecology* **1**, 55–64 (2007).
- [12] J. A. Dunne, R. J. Williams, and N. D. Martinez, *Proceedings of the National Academy of Sciences* **99**, 12917–12922 (2002).
- [13] P. E. Vértes, A. F. Alexander-Bloch, N. Gogtay, J. N. Giedd, J. L. Rapoport, and E. T. Bullmore, *Proceedings of the National Academy of Sciences* **109**, 5868–5873 (2012).
- [14] M. P. van den Heuvel and O. Sporns, *Trends in Cognitive Sciences* **17**, 683–696 (2013).
- [15] F. V. Farahani, W. Karwowski, and N. R. Lighthall, *Frontiers in Neuroscience* **13**, 00585 (2019).
- [16] D. S. Bassett and E. T. Bullmore, *Current Opinion in Neurology* **22**, 340–347 (2009).
- [17] S. Hong, B. C. Coutinho, A. Dey, A. L. Barabási, M. Vogelsberger, L. Hernquist, and K. Gebhardt, *Monthly Notices of the Royal Astronomical Society* **459**, 2690–2700 (2016).
- [18] R. de Regt, S. Apunevych, C. von Ferber, Y. Holovatch, and B. Novosyadlyj, *Monthly Notices of the Royal Astronomical Society* **477**, 4738–4748 (2018).
- [19] V. Springel, C. S. Frenk, and S. D. M. White, *Nature* **440**, 1137–1144 (2006).

- [20] G. A. Pagani and M. Aiello, *Physica A: Statistical Mechanics and its Applications* **392**, 2688–2700 (2013).
- [21] J. M. Oliver, T. Kipouros, and A. M. Savill, *Procedia Computer Science* **29**, 1948–1958 (2014).
- [22] A.-L. Barabási, *Network science* (Cambridge University Press, Cambridge, England, 2016).
- [23] A.-L. Barabási and R. Albert, *Science* **286**, 509–512 (1999).
- [24] M. E. J. Newman, *SIAM Review* **45**, 167–256 (2003).
- [25] J. Gao, B. Barzel, and A.-L. Barabási, *Nature* **530**, 307–312 (2016).
- [26] F. A. Rodrigues, L. d. F. Costa, and A. L. Barbieri, *Molecular BioSystems* **7**, 1263–1269 (2011).
- [27] J. Stelling, S. Klamt, K. Bettenbrock, S. Schuster, and E. D. Gilles, *Nature* **420**, 190–193 (2002).
- [28] Y. Tang, F. Qian, H. Gao, and J. Kurths, *Annual Reviews in Control* **38**, 184–198 (2014).
- [29] A. Arenas, A. Díaz-Guilera, J. Kurths, Y. Moreno, and C. Zhou, *Physics Reports* **469**, 93–153 (2008).
- [30] B. Ermentrout, *Journal of Mathematical Biology* **29**, 571–585 (1991).
- [31] H. Hamann, T. Schmickl, H. Wörn, and K. Crailsheim, *Neural Computing and Applications* **21**, 207–218 (2010).
- [32] V. In and A. Palacios, *Symmetry in complex network systems* (Springer, Berlin, Germany, 2017).
- [33] F. Mandl and G. Shaw, *Quantum field theory*, 2nd ed. (Wiley-Blackwell, Chichester, England, 2007).
- [34] H. Nakao and A. S. Mikhailov, *Nature Physics* **6**, 544–550 (2010).
- [35] M. Rietkerk and J. van de Koppel, *Trends in Ecology & Evolution* **23**, 169–175 (2008).
- [36] G. Barzon, G. Nicoletti, B. Mariani, M. Formentin, and S. Suweis, *Journal of Physics: Complexity* **3**, 025010 (2022).
- [37] G. Rogers, R. T. Elliott, D. J. Trudnowski, F. Wilches-Bernal, D. Osipov, and J. H. Chow, *Power system oscillations*, 2nd ed. (Springer International Publishing, Cham, Switzerland, 2025).
- [38] M. Barthélemy, *Physics Reports* **499**, 1–101 (2011).
- [39] P. Manser, H. Becker, S. Hörl, and K. W. Axhausen, *Transportation Research Part A: Policy and Practice* **137**, 1–15 (2020).
- [40] J. L. Adler and V. J. Blue, *Transportation Research Part C: Emerging Technologies* **10**, 433–454 (2002).
- [41] M. Gen and R. Cheng, *Genetic algorithms and engineering design* (John Wiley & Sons, Nashville, TN, USA, 1997).
- [42] G. C. Dandy, A. R. Simpson, and L. J. Murphy, *Water Resources Research* **32**, 449–458 (1996).

- [43] M. G. Bell, J.-J. Pan, C. Teye, K.-F. Cheung, and S. Perera, *Transportation Research Part B: Methodological* **132**, 15–28 (2020).
- [44] L. Xiaowei, C. Hong, and M. Juan, in *Proceedings of the 2010 International Conference on Optoelectronics and Image Processing* (2010).
- [45] S. E. Christodoulou, *Journal of Transportation Engineering* **136**, 102–109 (2010).
- [46] S. V. Buldyrev, R. Parshani, G. Paul, H. E. Stanley, and S. Havlin, *Nature* **464**, 1025–1028 (2010).
- [47] D. Ngoduy, *Transportation Research Part B: Methodological* **150**, 260–278 (2021).
- [48] S. Viciani, M. Lima, M. Bellini, and F. Caruso, *Physical Review Letters* **115**, 083601 (2015).
- [49] R. d. J. León-Montiel and J. P. Torres, *Physical Review Letters* **110**, 218101 (2013).
- [50] F. Caruso, A. W. Chin, A. Datta, S. F. Huelga, and M. B. Plenio, *The Journal of Chemical Physics* **131**, 105106 (2009).
- [51] Y. Ilan, *Progress in Biophysics and Molecular Biology* **178**, 83–90 (2023).
- [52] L. S. Tsimring, *Reports on Progress in Physics* **77**, 026601 (2014).
- [53] M. Dorigo, M. Birattari, and T. Stutzle, *IEEE Computational Intelligence Magazine* **1**, 28–39 (2006).
- [54] A. S. Mohaymany and A. Gholami, *Journal of Transportation Engineering* **136**, 323–331 (2010).
- [55] H. Poorzahedy and F. Abulghasemi, *Transportation* **32**, 251–273 (2005).
- [56] A. Tero, R. Kobayashi, and T. Nakagaki, *Journal of Theoretical Biology* **244**, 553–564 (2007).
- [57] A. Tero, S. Takagi, T. Saigusa, K. Ito, D. P. Bebber, M. D. Fricker, K. Yumiki, R. Kobayashi, and T. Nakagaki, *Science* **327**, 439–442 (2010).
- [58] M. Le Verge-Serandour and K. Alim, *Annual Review of Condensed Matter Physics* **15**, 263–289 (2024).
- [59] O. Fiz-Palacios, M. Romeralo, A. Ahmadzadeh, S. Weststrand, P. E. Ahlberg, and S. Baldauf, *PLoS ONE* **8**, e74374 (2013).
- [60] Y. Tu and W.-J. Rappel, *Annual Review of Condensed Matter Physics* **9**, 183–205 (2018).
- [61] R. K. Ahuja, T. L. Magnanti, and J. B. Orlin, *Network flows* (Pearson, Upper Saddle River, NJ, USA, 1993).
- [62] M. Newman, *Networks* (Oxford University Press, London, England, 2010).
- [63] W. Ellens and R. Kooij, arXiv preprint arXiv:1311.5064 (2013).
- [64] P. Y. Sohounou, P. Christidis, A. Christodoulou, L. A. Neves, and D. L. Presti, *International Journal of Critical Infrastructure Protection* **29**, 100353 (2020).
- [65] M. Oehlers and B. Fabian, *Mathematics* **9**, 895 (2021).
- [66] E. W. Dijkstra, *Numerische Mathematik* **1**, 269–271 (1959).

- [67] R. Diestel, *Graph theory*, 6th ed. (Springer, Berlin, Germany, 2024).
- [68] D. B. West, *Introduction to graph theory*, 2nd ed. (Pearson, Upper Saddle River, NJ, USA, 2000).
- [69] R. Albert, H. Jeong, and A.-L. Barabási, *Nature* **406**, 378–382 (2000).
- [70] J. Kleinberg and E. Tardos, *Algorithm design: pearson new international edition* (Pearson Education, London, England, 2013).
- [71] T. H. Cormen and C. E. Leiserson, *Introduction to algorithms*, 4th ed. (MIT Press, London, England, 2022).
- [72] J. B. Kruskal, *Proceedings of the American Mathematical Society* **7**, 48–50 (1956).
- [73] R. C. Prim, *Bell System Technical Journal* **36**, 1389–1401 (1957).
- [74] L. R. Ford and D. R. Fulkerson, *Flows in networks* (Princeton University Press, Princeton, NJ, USA, 2024).
- [75] A. Bondy and U.S.R. Murty, *Graph theory*, 1st ed. (Springer, London, England, 2007).
- [76] K. Salimifard and S. Bigharaz, *Operational Research* **22**, 1–47 (2020).
- [77] W. Gu, C. Archetti, D. Cattaruzza, M. Ogier, F. Semet, and M. G. Speranza, *European Journal of Operational Research* **317**, 1–15 (2024).
- [78] C. Archetti, L. Peirano, and M. G. Speranza, *European Journal of Operational Research* **299**, 1–20 (2022).
- [79] D. Bertsekas, *Network optimization: continuous and discrete models* (Athena Scientific, Belmont, MA, USA, 1998).
- [80] L. Zhu, S.-J. Kim, M. Hara, and M. Aono, *Royal Society Open Science* **5**, 180396 (2018).
- [81] V. Bonifaci, K. Mehlhorn, and G. Varma, *Journal of Theoretical Biology* **309**, 121–133 (2012).
- [82] V. Bonifaci, E. Facca, F. Folz, A. Karrenbauer, P. Kolev, K. Mehlhorn, G. Morigi, G. Shahkarami, and Q. Vermande, *Theoretical Computer Science* **920**, 1–20 (2022).
- [83] F. Folz, K. Mehlhorn, and G. Morigi, *Physical Review Letters* **130**, 267401 (2023).
- [84] A. Lonardi, E. Facca, M. Putti, and C. De Bacco, *Physical Review Research* **3**, 043010 (2021).
- [85] A. Lonardi, M. Putti, and C. De Bacco, *Scientific Reports* **12**, 7474 (2022).
- [86] F. Folz, K. Mehlhorn, and G. Morigi, *Physical Review E* **110**, 044310 (2024).
- [87] N. Kamiya, *CYTOLOGIA* **15**, 183–193 (1950).
- [88] P. Stewart and B. T. Stewart, *Experimental Cell Research* **18**, 374–377 (1959).
- [89] G. Isenberg and K. Wohlfarth-Bottermann, *Cell and Tissue Research* **173**, 495–528 (1976).
- [90] T. Miyaji, I. Ohnishi, A. Tero, and T. Nakagaki, *International Journal of Dynamical Systems and Differential Equations* **1**, 210–219 (2008).

- [91] J. R. Banavar, F. Colaiori, A. Flammini, A. Maritan, and A. Rinaldo, *Physical Review Letters* **84**, 4745–4748 (2000).
- [92] F. Kaiser, H. Ronellenfitsch, and D. Witthaut, *Nature Communications* **11**, 5796 (2020).
- [93] F. Folz, K. Mehlhorn, and G. Morigi, *Phys. Rev. E* **104**, 054215 (2021).
- [94] B. Meyer, C. Ansorge, and T. Nakagaki, *PLOS ONE* **12**, e0172933 (2017).
- [95] A. A. Ibrahim, A. Lonardi, and C. De Bacco, *Algorithms* **14**, 189 (2021).
- [96] A. Lonardi, D. Baptista, and C. De Bacco, *Frontiers in Physics* **11**, 1089114 (2023).
- [97] A. Lonardi, E. Facca, M. Putti, and C. De Bacco, *Physical Review E* **107**, 024302 (2023).
- [98] D. Soydaner, *International Journal of Pattern Recognition and Artificial Intelligence* **34**, 2052013 (2020).
- [99] S. Sun, Z. Cao, H. Zhu, and J. Zhao, *IEEE Transactions on Cybernetics* **50**, 3668–3681 (2020).
- [100] L. Bottou, F. E. Curtis, and J. Nocedal, *SIAM Review* **60**, 223–311 (2018).
- [101] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, *Science* **220**, 671–680 (1983).
- [102] P. J. M. Van Laarhoven and E. H. L. Aarts, *Simulated annealing: theory and applications*, Mathematics and Its Applications (Kluwer Academic, Dordrecht, Netherlands, 1987).
- [103] J. Liu, F. Wilde, A. A. Mele, L. Jiang, and J. Eisert, arXiv preprint arXiv:2210.06723 (2022).
- [104] J. Nocedal and S. Wright, *Numerical optimization*, 2nd ed. (Springer, New York, NY, USA, 2006).
- [105] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, *The Journal of Chemical Physics* **21**, 1087–1092 (1953).
- [106] A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. O’Brien, *Nature Communications* **5**, 4213 (2014).
- [107] J. Tilly et al., *Physics Reports* **986**, 1–128 (2022).
- [108] J. R. McClean, S. Boixo, V. N. Smelyanskiy, R. Babbush, and H. Neven, *Nature Communications* **9**, 4812 (2018).

DANKSAGUNG

Mit dem Abschluss dieser Dissertation geht eine intensive und prägende Phase zu Ende, die ohne die Unterstützung und Begleitung vieler Menschen nicht möglich gewesen wäre. An dieser Stelle möchte ich meinen tiefen Dank aussprechen.

Zunächst danke ich meiner Betreuerin Prof. Giovanna Morigi für die fachliche Unterstützung, die inspirierenden Diskussionen und die wertvollen Anregungen. Ihr Vertrauen in meine Arbeit hat mich stets angespornt und die nicht selbstverständliche Freiheit, die sie mir in unserer gemeinsamen Forschung gewährt hat, hat sehr zu meiner persönlichen Weiterentwicklung beigetragen. Gleiches gilt für Prof. Kurt Mehlhorn, dessen Zusammenarbeit während meiner Promotion ich sehr geschätzt habe. Ich möchte auch Prof. Malte Henkel, Prof. Karl Jansen, Prof. Albrecht Ott und Dr. Philipp Hövel für die anregenden Diskussionen und wertvollen Einschätzungen danken. Ein weiterer Dank gilt Prof. Karsten Kruse, der als Co-Betreuer meiner Bachelor-Arbeit einen prägenden Einfluss auf meine akademische Laufbahn hatte.

Ein besonderer Dank gilt meiner Familie, vor allem meinen Eltern Karin und Hermann sowie meiner Großmutter Helga, die mich in jeder Hinsicht unterstützt, mir den Rücken freigehalten und mich ermutigt haben, meinen Weg zu gehen.

Auch meinen aktiven und ehemaligen Kollegen möchte ich danken, insbesondere, in alphabetischer Reihenfolge, Aleksei Konovalov, Emma King, Gage Harmon, Grecia Guijarro, Marvin Gajewski, Raphael Menu, Rebecca Kraus, Robin Krill, Sayan Roy und Tom Schmit. Ihre Freundschaft, ihr Humor und ihre Ermutigung waren ein wertvoller Ausgleich zu den Herausforderungen des wissenschaftlichen Arbeitens.

Diese Dissertation ist nicht nur das Ergebnis meiner Arbeit, sondern auch ein Zeugnis der Verbundenheit und Unterstützung, die ich in den vergangenen Jahren erfahren durfte. Danke an alle, die mich auf diesem Weg begleitet haben.