Towards a Theory of Modular-Hierarchical Structures in Living Systems

1. Results from Previous NSF Support

Award PHY-0848755: Experimental and Theoretical Analysis of Collective Dynamics in Swarming Systems. Subaward PI: Cristián Huepe; Subaward amount: $194,100 (9/1/09-8/31/13) (main-award PI: Prof. I D Couzin, Princeton U.; total award: $543,472). This project studied collective motion, schooling and swarms experimentally and theoretically from a non-equilibrium statistical physics perspective.

Intellectual Merit: Theory (Huepe): developed detailed and idealized models to capture specific and universal aspects of swarm dynamics based on experimental results and theoretical analysis; introduced an adaptive-network description of collective motion; solved the inverse problem for small groups of fish, deducing the interactions between individuals from experimental data thus unveiling non-additive effects;; extended standard swarming models to include variable speeds, visual interactions, etc.; characterized transition dynamics between states of collective motion; developed new approaches for understanding collective decision making in animal groups; discovered a novel elasticity-based self-organization mechanism for active systems that leads to collective motion even in the absence of explicit alignment. Experiments (Couzin): implemented fully controlled fish swarming experiments, tracking all individuals.

Broader Impacts: Huepe co-trained four graduate students & five postdoctoral researchers; organized the interdisciplinary art/science meetings *Arts, Humanities and Complex Networks* (within NetSci2012) and *Networks and Nonlinearities in the Musical Experience* (at the ZiF Center for Interdisciplinary Research, Germany), which received broad media coverage; participated in public lectures and a roundtable on *Physics of Music* at the Chicago Cultural Center; worked as a scientific consultant for the *eduMedia* educational website; developed control algorithms that were applied in swarm robotics engineering; gave keynote public presentations on music and complex systems in Koblenz (Germany) and Viña del Mar (Chile). Media coverage: Wired Magazine, National Geographic, Focus Online, MSNBC blog, etc**.,** and articles on a broad range of international press. (Main award PI, I. D. Couzin, also contributed to the Broader Impacts of the project by leading CouzinLab, teaching at Princeton U, participating in advisory panels, and giving public lectures and media interviews.)

Publications (by C. Huepe) within this project include work on: experimental studies of collective behavior in schooling fish (1,2), adaptive network models of swarm dynamics (3); decentralized control of swarm robotics and artificial life applications (4,5); continuous theory of collective motion with variable speeds (6), network-based analyses of the opinion-formation models (7), criticality in living systems (8,9), self-organization in swarms and active elastic systems (10,11), and art/science interdisciplinary research on music and complex systems (12). (This project also led to over 25 other publications by Couzin's group –including 2 in Science and 5 in PNAS).

Other products and activities: Data, models, and code are available at CouzinLab's and Huepe's websites. Wrote chapter on 'Flocking and Music' for the book *Controls and Art* (Springer, 2014) (13). Member of the Advanced Study Group: *Statistical Physics of Collective Motion* at the Max Planck Institute for the Physics of Complex Systems (Dresden, Germany, 2011-2013).

2. Introduction

A common feature of many biological systems is that they are organized as a hierarchy of modules that help structure the interactions and processes of their components. Basic building blocks combine into more complex structures, which serve as modules that can in turn organize into structures at yet higher levels (see Figure 1), a process that can repeat multiple times and span many spatiotemporal scales. This multi-scale modular and hierarchical (MH) organization can be observed in biological interaction networks such as genetic networks, metabolic networks, neural networks, or ecological networks.

In recent years, transformative approaches for characterizing living systems have produced substantial quantitative information describing biological processes at the molecular, cellular, organismal and even population levels. Providing an adequate theoretical framework to transform these data into understanding is a grand challenge and opportunity for this century (14). Much of this information can be effectively represented by various biological interaction networks, including genetic regulatory networks,

protein interaction networks, metabolic networks on a microscopic scale, interaction networks of cells in tissues, neural networks on a mesoscopic scale, and the interaction of species on a large scale in ecological networks. Biological interaction networks thus exist across many scales and generically exhibit fractal-like MH architecture. This MH organization impacts the dynamical properties of these systems, since interactions between building blocks are also organized as a hierarchy of processes and subprocesses, often spanning multiple timescales. Here, we refer to systems that follow such organization in either their interaction topologies or their dynamics as MH *systems* or MH *structures*.
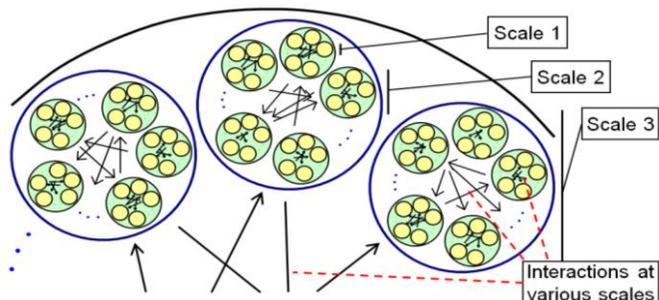
Figure 1: Schematic representation of modular-hierarchical (MH) structure.
Interacting modules organize in a hierarchy of levels or scales. Small yellow discs are modules that interact (through arrows) to produce the dynamics of the embedding green discs (modules at the next scale). These interact in turn within the embedding modules (blue circles), yielding dynamics at the next interaction scale. Only three levels are displayed, but the illustrated hierarchy can be repeated any number of times.

Despite the ubiquity of MH structures in living systems, we still lack formal tools for defining and characterizing them, and have limited understanding of their origin and dynamics. They have not received much attention from the perspective of Physics, where scaling and self-similarity analyses are common but traditional systems do not exhibit the level of emergent organization required to develop complex MH features. This opens the tantalizing opportunity of exploring the properties of MH systems with the type of minimal, general, and hopefully universal descriptions often used in Physics. Given how little this targeted and unifying theoretical perspective has been developed, even a basic understanding of the dynamical and structural properties expected from MH biological systems would help confront the rapidly developing quantitative information described above, and could lay the groundwork for a future theory of MH living systems. Such understanding would also help clarify, from the point of view of Biology, the consequences of these structures on the dynamics of organisms and on evolutionarily-important properties such as robustness and fitness, and the associated causes for their development.

The purpose of this proposal is to lay the theoretical groundwork for understanding the role of MH structures in living systems by combining new models and approaches derived from statistical physics, non-equilibrium dynamics, complex networks, and active matter with insights and data from a broad spectrum of experiments by multiple collaborators.

We will be guided by the following central, overarching and open questions:

- **What types of dynamics are supported by MH structures?**
- **How do MH structures emerge and evolve?**

We will thus directly address the fundamental properties that can lead to, and result from, MH structures. This ambitious, novel and distinct research program will be confronted from two complementary perspectives. On the one hand, we will consider minimal models that reveal the classes of generic and universal features of MH systems; on the other hand, we will develop specific descriptions of MH structures and dynamics using experiments and data from concurrent collaborative investigations with various groups, whose range of expertise span synthetic biology, metabolic and genetic networks, microbial symbioses, evolutionary dynamics, collective behavior, network theory, and active matter.

The type of research program that we are proposing has a long and successful tradition in Physics. Indeed, understanding the generic features expected from classes of systems has allowed us to unify knowledge, unveil universal features, classify systems, formulate null hypotheses, learn through analogies, and develop approximate mathematical descriptions and a qualitative understanding of essential system components. An example is the study of idealized descriptions in Statistical Physics, such as the Ising model and kinetic theory, which has helped us understand the properties of a number of systems that extend far beyond the ferromagnetic materials and ideal gases for which they were originally designed, even including complex social, ecological, and technological ones. As increasingly growing information on MH systems is gathered, a similar approach (albeit using new models and techniques) appears timely and necessary to understand more deeply the origins and dynamics of MS structures.

3. Current Knowledge on MH Structures

We will describe here some of the elements of current knowledge, originating from different fields, that will serve as starting point for our research program. To date, much of the work involving MH systems has consisted of isolated efforts focusing on: (i) identifying MH structures in biological systems (ii) developing network tools to analyze their degree of hierarchy and modularity, (iii) implementing network growth models that produce MH topologies, and (iv) investigating examples of the dynamics supported by them. Because these investigations are disparate in both methodology and context, however, they have not enabled a direct investigation of the type of fundamental questions that are the focus of this proposal.

MH structures are believed to be an essential part of biological organization, related to the many levels and types of functional and structural heterogeneity in living systems (15–17). The existence of MH structures in biological networks is evidenced by a number of recent studies in which data from different biological systems was analyzed to reveal underlying organizational principles. These studies have identified at least three topological and dynamical features that evidence MH structures, namely, the presence of *motifs*, *modules*, and a *hierarchical organization*. For example, genetic regulatory networks contain a small set of recurring patterns, called *network motifs*, with similar regulation functions (18). Specific motifs, classified into families such as feedforward loops, single-input modules, or dense overlapping regulons, have been identified in *E. coli* (19,20), yeast (21,22), and other organisms (23–25). Motifs can be viewed as the lower-level organizational units of a MH structure, with interactions among motifs forming the next hierarchical level. The presence of underlying MH architectures is also supported by the studies of *modularity* (15) in biological systems, which emerged in parallel in evolutionary developmental biology (26,27) and molecular systems biology (28,29). This abstract property is observed in a variety of contexts and is believed to be an essential part of biological organization, related to the many levels and types of functional and structural heterogeneity in living systems (16,17). While specific formal definitions vary, a well-established working definition describes a module as a part of an organism that integrates a set of processes and operates comparatively autonomously with respect to other modules in a system (15,30). Biological networks with modular organization include protein-protein interactions in yeast (31,32), gene regulatory networks in *Drosophila* and in processes of higher organisms such as vertebrate myogenesis (33), co-varying traits of evolving organisms associated to modular morphogenesis (34,35), and RNA secondary structure (36). As in the case of motifs, modularity implies at least two structural levels: one for the internal functioning of the modules and the other for interactions between modules. A natural extension of these analyses is to consider subsequent levels of organization in which structures of structures are identified. This corresponds precisely to the recently considered property of having a *hierarchical* organization. While motifs and modules require only a two-level hierarchy, it is harder to identify multi-level hierarchies. Several approaches for extracting these structural properties from the architecture of natural networks have been developed (37–41). Although the unambiguous identification of modules at different levels of the hierarchy is still an open problem, there is growing consensus that these are recurrent structures in biological networks (42,43). Indeed, topologies displaying hierarchical modularity have been identified in a variety of biological networks (44,45). However, the relationship between topological modules defined based on connectivity and functional units remains unclear, and thus it is not known whether these structures emerge due to natural selection or rather due to other underlying processes (46).

From a theoretical perspective, other studies have considered the question of which network architectures are compatible with the MH structures observed in nature. When modular topology is combined with the approximately power-law, long-tail *degree distribution* (number of links per node (47)) also common in biological networks, an apparent contradiction emerges. Indeed, a characteristic feature of scale-free networks is the presence of a small number of nodes with many interactions. These highly connected nodes, or hubs, tend to integrate the whole system in a way that would seem incompatible with a naive picture of relatively autonomous modules. However, in a study of 43 distinct organisms (48) it was shown that their metabolic networks are organized into a hierarchy of highly connected modules that are combined into successively larger and less connected structures, satisfying simultaneously the conditions of modularity and scale-free connectivity. The resulting architecture, defined as a *hierarchical network*, also reproduces other features of metabolic networks, such as their clustering coefficient (49). Various deterministic hierarchical algorithms for constructing networks that are both modular and scale-free have been introduced (48,50–52). These serve as a starting point for the investigations in **Section 5** below.

The studies described above focus on the network topology, without addressing the evolutionary dynamics that is implied (as either a cause or consequence) by it. Evolution is a strong unifying principle in biology and an essential component for the development of living systems at all scales, from basic genetic processes to complex interactions between species. In spite of its apparent universality and simple conceptual basis, the processes through which random fluctuations and natural selection can lead to the remarkable level of self-organization observed in living systems is far from understood. Notably, the role and mechanisms of natural selection at different levels of a MH structure are still the subject of intense debate (53,54). Many models have been developed to try to reproduce the essential properties of evolving systems. These include a variety of approaches such as the analysis of minimal cellular automata (55) and Boolean network models (56), the use of concepts from game theory (57) and graph theory (58), and the implementation of Genetic Algorithms (59) and Artificial Life programs (60) like *PolyWorld* (61), *Tierra* (62), *Avida* (63), *ECHO* (64), and *AntFarm* (65), to name a few. While interesting lessons have been learned from these studies, they are not designed to explore the evolution of MH structures, since they typically lack the necessary degrees of freedom, sufficient spatiotemporal scales, and the multi-scale environmental pressures required to develop or sustain them. One of the objectives of the proposed project (detailed in **Aim 3** below) is to combine algorithms that generate MH networks with simulations of evolutionary dynamics, in order to study the emergence and stability of MH structures. Our initial steps will be guided by recent work suggesting that modular structures emerge in systems that evolve under alternating fitness criteria (66–68). This allows them to rapidly access different functions by activating different modules when selection pressures change

4. <u>Project Overview</u>

The project aims to advance our understanding of MH structures by investigating the dynamical consequences and evolutionary origins of this ubiquitous feature of living systems. To address the fundamental questions listed in the **Introduction**, we propose to develop the essential modelling and theoretical tools and concepts necessary to address the structural, dynamical, and evolutionary aspects of MH biological systems in close collaboration with experimentalists and data-analysis teams.

<u>Our goals are to</u>: (a) characterize MH systems, formalizing existing definitions, and link them to theory and experiments; (b) explore MH dynamics on network models, spatial dynamical models, and the experimental MH systems studied by our collaborators; and (c) investigate whether evolutionary forces favor MH structures and the effect of these structures on fitness in these models and specific systems.

<u>We will study, in particular</u>: (i) the presence and emergence of MH structures in network-based models and various experimental systems; (ii) the connection between MH interaction topologies and MH dynamics; (iii) properties of dynamical processes supported by MH network models and experimental systems; (iv) hybrid models that combine network-like and diffusive couplings to describe the physical constraints of the interactions at different levels of the hierarchy; (v) spatial models that link the self-organization of physical forces required to produce work at higher scales with the emergence of MH structures; (vi) network evolution and growth algorithms that produce MH architectures and dynamics, and their extension to artificial life simulations; and (vii) the robustness and evolvability of dynamics on MH networks, including experimental MH systems subjected to changing environmental conditions and selective pressures. Finally, we will enhance the broader impact of this work by implementing point v above as a mechanical analogy for a MH biological system that will serve as a tool for teaching nonequilibrium dynamics, self-organization and evolution; together with a means to engage a large audience through an online game-based crowdsourcing experiment, which will result in a data source of MS structures obtained from human design that will be compared to those obtained through evolution.

<u>Our activities</u> will focus on: (i) establishing and analyzing archetypical minimal models that capture the conceptual origin and potential behaviors of the MH organization observed in a diversity of biological systems, while testing their universality; and (ii) providing theoretical analysis, guidance, and insights to the experimental and data-driven collaborators for the interpretation of results and implementation of new experiments or analysis. We expect these activities to stimulate effects somewhat analogous to those resulting from the development of the aforementioned Ising or kinetic theory approaches, which have led to numerous conceptual advances, experimental predictions, simulations, analytical developments, and extensions to other fields. Our models will specifically focus on capturing nonequilibrium emerging features that are evolutionary-based and fundamentally multi-scale.

Our current research work places us in a unique position to pursue this ambitious research program. Indeed, such an abstract approach could easily lose focus or relevance if it was not strongly motivated by experiments and applications. However, we have established a network of fruitful collaborations with experimental, applied, and theoretical groups, which were at the origin of this research program, and will continue to inform it. A quick overview of these efforts is the following. (i) In a different collaborative project with Prof. Joshua Leonard (Northwestern U), we are studying the dynamics, evolvability, and adaptability of MH bacterial toxin-antitoxin networks and other experimental systems that are being built through synthetic biology his lab. (ii) In independent efforts with Prof. Calin Guet (IST, Austria), we are analyzing other synthetic biology systems based on *E-coli* to identify candidate networks and processes with MH features that can be tinkered with in his lab (see **Section 6.1**). (iii) In current work with Prof. Luis Amaral (Northwestern U) we are implementing flux balance analysis on the biomass production network of various organisms to compare their MH structure and function by taking advantage of his previously developed algorithms that identify and help visualize MH structures in networks (see **Section 6.2**). (iv) In interactions with Prof. Mark Mandel (Northwestern U), we are identifying which genetic and metabolic modular features of the host organism must be reflected as modular features of the guest organism in the beneficial colonization of Euprymna scolopes squid by Vibrio fischeri bacteria, which he studies in detail, given the continuous coadaptation that they are subjected to. (v) In conversations with Prof. Sui Huang (Institute for Systems Biology, Seattle), we are discussing specific example model systems that display MH features in tumor biology and, in particular, his data on gene regulatory networks and cell-cell communication networks of tumor cells as they evolve to become drug resistant (see **Section 6.3**). (vi) In collaboration with Prof. Dirk Brockmann (Humboldt University and Robert Koch Institute, Berlin), we developed a new, physically and biologically plausible algorithm that results in fully randomized MH networks and are currently studying their topological and dynamical features, while also studying systems that combine diffusive processes and modular networks (see **Section 5.1**). (vii) In research with Prof. Maximino Aldana (UNAM, Mexico), we are investigating under what conditions MH structures may emerge from dynamical fitness requirements imposed on the attractors of Boolean networks that are the focus of his studies (see **Section 5.2**). (viii) In continuing efforts with Prof. Thilo Gross (U of Bristol), we have developed approaches to model MH dynamics in space and generate MH structures using adaptive networks and will soon start analyzing MH networks using spectral methods (see **Section 5.3**). (ix) In collaboration with Prof. Ali-Emre Turgut and Dr. Eliseo Ferrante (KU-Leuven), we developed the AEM model described in **Section 5.4** below based on a decentralized control algorithm that they implemented on swarming autonomous robots, which we have recently extended to describe MH systems.

The disparate systems studied through these collaborations will help this early stage of development of a theory of MH biological systems. They will help cross-fertilize ideas from different contexts; search for universal organizational principles, dynamical features, and evolutionary origins of MH structures; and find specific experimental realizations and potential applications on concrete systems. While they may not all lead to the discovery of strong manifestations of MH features, it is precisely in their diversity that we maximize our chances of finding systems where MH structures are critical for their understanding. If this happens in any of these, it will likely lead to the opening of a completely new approach in their study.

5. <u>Theoretical Approaches and Activities</u>

We describe below the modeling and theoretical approaches and activities of our research program. The first three consider non-spatial network-based interactions, and can describe a broad range of complex interacting systems. They will allow us to establish parallels with the diverse and rapidly growing datasets on biological MS interaction networks (especially those of our collaborators), and to integrate and extend the current knowledge on MS topologies described in **Section 3**. The last two frameworks consider simple generic spatial interactions, to include the effects of distance between interacting components. Note that these distances are not restricted to physical space, and could also describe abstract spaces such as fitness landscapes or biochemical affinities, thus greatly expanding model applicability. They provide concrete, intuitive mechanical representations of the connectivity, dynamics, and evolution of MH systems. Including spatial dynamics also relates our work to spatially-based MH processes (e.g. morphogenesis) and has technological applications (69–71) (e.g. in swarm robotics). The two modeling frameworks will be developed in parallel, addressing questions like those in **Section 7**.
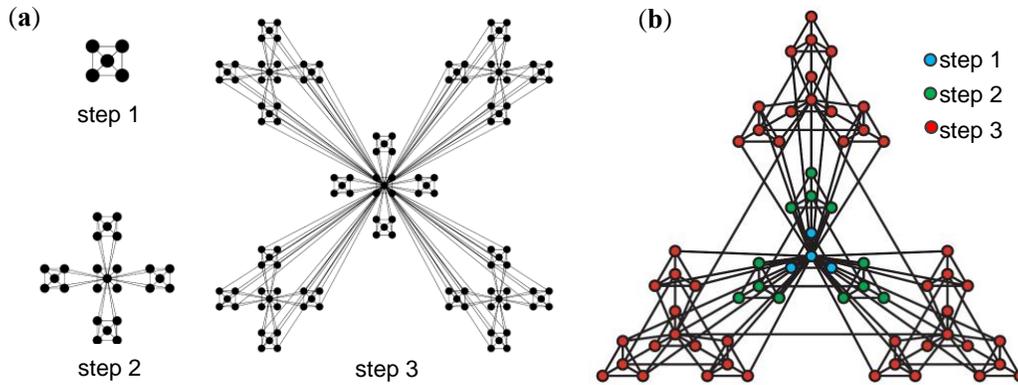
Figure 2: Examples of published duplication-based MH network generating algorithms (48,51). Steps are displayed separately (a) or labeled by color (b). Central node connectivity makes the central module the highest hierarchical level in both cases. Direct same-level interactions between modules are included in (b) but not in (a).

5.1. The MH Complex Networks Approach: Topology and Basic Dynamics

We introduce here an approach for our theoretical investigations based on classes of network models that provide a common framework to study of MH structures in living systems. MH network topologies can be generated by various algorithms recently proposed in the literature. These can be categorized as: duplication based, subdivision based, and adaptive-network based. Duplication based algorithms grow the network by making copies of the current network at every step and wiring these to produce a larger network. This process is repeated until the desired size is reached (48,50–52,72). Subdivision based algorithms, instead, repeatedly replace individual nodes of the current network by sub-networks, following a given protocol (73–75). Adaptive-network based algorithms evolve the connectivity based on the dynamics of node states supported by the network or on emergent properties of the topology (66–68,76–81). Since one of our objectives is to find biological mechanisms that could lead to MH networks, we will focus here on duplication and adaptive-network based algorithms (82–84), which have clear evolutionary interpretation and can be designed to match our experimental systems. We will then explore the connection between MH topologies and MH dynamics on our resulting networks.

In order to structure our exploration of the dynamics and evolution of different MH network systems, we will consider a family of models based on a single Generalized Algorithm (GA). We start by explaining two duplication-based algorithms found in the literature (48,51), which are displayed on Fig. 2. These create first a network of 5 or 4 all-to-all connected nodes (step 1). This module is then copied 4 or 3 times (step 2) and the outer nodes of the new copies connected to the central node of the original module. In panel (b) the central nodes of the new structures are also linked. These steps are repeated until the desired network size is reached. If we interpret these algorithms as representing idealized developmental and evolutionary dynamics, it is natural to extend them by introducing mutations in the module duplication steps. Further evolutionary aspects will be included using adaptive-network algorithms that modify connectivity, for which we will benefit from the expertise of external collaborator Prof. Gross (who will also help classify topologies using spectral analysis –see attached letter). The analogy with evolution can be completed by adding selective pressures to the fitness of the resulting structures, defined through the emergent properties of either the topology or the resulting dynamics of network states (considering these as a proxy for the system's function), as described in **Section 5.3**. Our GA is defined as follows:

1. **Seed topology and dynamics:** Create seed module with N_m nodes connected according to an internal coupling (adjacency) matrix $\mathbf{W}_=$. If Step 4 is required, define a dynamical system on the network by specifying node states and node processes.
2. **Multiply:** Make multiple versions of the current structure either by duplicating and mutating it or by generating an ensemble of variations of it, creating a set of modules of modules after the 2nd iteration.
3. **Connect:** Define node subsets; S_-, $S_=$, S_+; that mediate each module's connectivity to others in lower, equal, or higher hierarchical levels, respectively. Connect new (lower-level) modules among them following $\mathbf{W}_=$, linking only $S_=$ nodes, and to the old (higher-level) module according to a $\mathbf{W}_+$ matrix, linking only lower-level S_+ nodes to higher-level S_- ones.

4. **Select:** Select nodes, links, or structures based on fitness criteria imposed on the dynamical activity of the states of specific nodes and modules. In some setups, prune the network by also discarding nodes or links that are irrelevant for the resulting dynamics of module states.
5. **Repeat:** Go to Step 2 to continue evolving and/or adding new levels to the MH structure.

Examples (a) and (b) in Fig. 3 are special cases of this GA with $N_m = 5$ and $N_m = 4$, respectively, where we skip Step 4, $\mathbf{W}_=$ and $\mathbf{W}_+$ describe all-to-all coupling, $S_=$ and S_- contain only the central node, and S_+ contains all outermost nodes of a module. Note that we will generalize these central and outer node definitions in terms of their connectivity (instead of their location on the diagrams). This GA will provide a common framework for our investigations and also allow us to develop a single computer code with modular subroutines to carry out our studies. We will implement several variations of this GA. In some of its versions, the GA is simply a network generating algorithm that produces different MH structures for our research, in other versions it can be viewed as a simple Artificial Life system that allows us to simulate MH evolution.

In addition to the algorithms found in the literature (which we generalize and extend in our GA above) we are currently studying with Prof. Brockmann a novel network generating algorithm, the Random MH Network (RMHN) algorithm that aims to produce networks that are akin to Erdős–Rényi random graphs (85), but with MH topology. These have the advantage of avoiding the presence of special nodes in the structures, such as the central nodes in Fig. 2(a) and (b). They can also generate different degree distributions that are not necessarily scale-free. The algorithm is simple and natural. It only requires defining a base number B (the number of structures per level, which can be generalized to a level-dependent distribution), the total number of levels N, and a list of probabilities P = $\{P_1, P_2, \cdots, P_N\}$ of linking nodes of different structures at each level. This algorithm is defined as follows:

1. **Create structure** of embedded hierarchies of groups of nodes containing B^N nodes in total, with each node labeled by a series of numbers (each from 1 to N) that specify their module, sub-module, etc. For example, node $\{3,1,4\}$ identifies node #4 in module #1, which is a sub-module of module #3.
2. **Generate links** between different structures at the same level l, using a probability P_l of creating a link between each pair of nodes in different modules at the same level.

We have started analyzing the mathematical properties of this algorithm, finding that it is well suited for analytical calculations. We are thus able to compute the way probabilities P must decay with level in order to satisfy modularity at each scale (i.e. to have more internal than external connections per module), the maximum decay allowed to produce a giant component network (85) and not fully dismember the structure, and the total degree distributions, to name a few.

The interest of considering these different network generating procedures is to find unify and extend the knowledge of MH networks and to classify their properties. This is a necessary step to then be able to systematically compare these idealized structures with those found in the natural and engineered biological systems that will be investigated in the proposed project. We will fist focus on characterizing their topological properties, such as their degree distributions, level of hierarchical modularity, clustering coefficients, centrality properties, etc. using the various algorithms that exist in the literature or defining new ones when necessary. We will also consider how these structures could develop and be observed in the real world by studying their resiliency to noise and to alternative network descriptions such as grouping sets of nodes into one unit, switching between directed and undirected nodes, or having access to only partial connectivity information.

5.2. Dynamics on MH Networks

The archetypical properties of processes that develop on MH structures will be investigated by studying a series of simple dynamical processes that we will implement on the node states. We will begin by carrying out analyses similar to those existing in disperse literature and extending them to the various MH topologies developed in **Section 5.1**. These include implementing on MH topologies Ising models (86), information flow models (87), a spin-glass model (88), synchronization models (89), and percolation models (90); with and without noise (91). We will seek to understand what properties are imposed on the dynamics by various MH topologies, trying to identify those that appear universal since they are found in different processes and variations of the MH topology. These investigations are already under way in collaboration with Prof. Brockmann, specifically on topologies generated by the RMHN algorithm above.
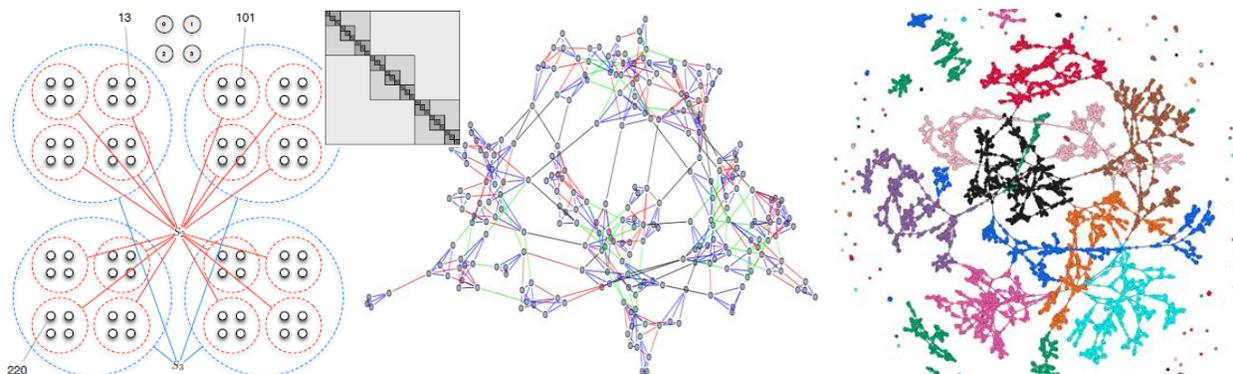
Figure 3: Examples of structure, topology, and dynamics on Random MH Networks (RMHN). Left: Structure of generating algorithm and of adjacency matrix (inset). Center: RMHN with 4 levels and 4 modules per level; links are colored by level. Right: Preliminary calculations of percolation on RMHM; modules initially percolate independently.

More general forms of nonequilibrium MH dynamics will then be studied by implementing ODE-based (that use ordinary differential equations) and Boolean-based dynamics on the networks generated by algorithms in **Section 5.1**. The ODE-based studies will provide a closer match to experimental systems, tracking in detail changes in the amounts of specific metabolites. Boolean-based dynamics allow the systematic analysis of all idealized processes supported by a given network structure, an approach first developed in the context of genetic networks (29,92,93). One of the simplest and best studied Boolean network models is the N-K model, or Kauffman model (94). In it, each node can only be in a 0 or 1 state and receives exactly K inputs from other, randomly selected nodes. The next state of each node is computed as a function of the state of the K nodes connected as its inputs, according to a truth-table generated at random for each node. A rich behavior emerges from this model (95). The resulting trajectories in phase space can be stationary, periodic, or chaotic. It was argued that living organisms must be at the critical interface between both phases to be robust enough to maintain performance under a broad range of random perturbations and also evolvable enough to adapt over time (96–101). Many extension of the N-K model have been considered (102,103). We will take advantage of the vast literature on this model and of its approach for exploring supported dynamics in phase space to analyze the processes supported by different MH topologies. We will benefit in this effort from interactions with external collaborator Prof. M. Aldana, an expert in the field (see letter attached).

By considering a set of archetypical processes on the network, we will also study how perturbations to stationary or periodic solutions propagate between modules and across the different levels of the network hierarchy. We will compare these results to those obtained for networks lacking MH structures and to experiments in the collaborations. We are particularly interested in exploring whether MH structures insulate modules from propagating perturbations and whether averaging effects within modules confers additional robustness to their resulting dynamics.

5.3.   Co-Evolving MH Networks

We will study the evolutionary dynamics of MH structures in network models by fully implementing versions of our GA (and variations of it based on our novel RMHN algorithm) as artificial life systems. We will first generate an ensemble of networks with desired architectures and then iterate the GA to evolve using selective pressures imposed on the dynamics of network states. We will test models of evolution under alternating fitness criteria, mimicking experiments that are being developed by external collaborator Prof. Joshua Leonard (Northwestern University), by introducing environmental perturbations as changes in the topology and dynamical rules. We will carry out various other numerical experiments, including:

- Test if and how the results from evolutionary algorithms depend on network representation: Evolutionary dynamics can be implemented into the links or dynamical rules, and can be executed on various types of networks (directed or not, Boolean, ODE-based, etc.). A limit case of purely topological evolution can even be implemented using only NAND gates (the "universal" NOT-AND function with which any logical function can be expressed). We will study how the evolved network structures depend on the specific description chosen, which representations are best suited for our experimental collaborations, and if there are unifying measures under which these disparate systems converge to the same results.

- *Design bio-inspired Adaptive-Network algorithms that generate MH topologies of interest:* The network-generating algorithms in **Section 5.1** were not designed as realistic ways in which MH topologies could have evolved in natural systems. In Adaptive-Network models (104), the topology and node states co-evolve, often converging to specific network structures. We will implement such models, searching to generate MH topologies with dynamical and linking rules that are physically and biologically plausible. We will benefit in this effort from the expertise of external collaborator Prof. Gross (University of Bristol), an expert in the field with whom we have been working on Adaptive-Network systems since 2010.

- *Perform artificial life experiments using Boolean and ODE-based MH dynamics:* We will implement our Boolean or ODE-based dynamics on an ensemble of MH networks and make them compete under various fitness constraints. We will examine how the resulting features depend on over which nodes (and at which levels of the MH structure) we impose fitness or competition. We will consider the emergence of structures, studying under which conditions independent, symbiotic, or collaborative structures may emerge, and if these can be interpreted as modules at a higher hierarchical level. The emergence of possible spatially segregated communities will be examined using the NSD model. If these are found, we will attempt to implement them experimentally with our collaborators. Finally, we will examine how evolution under changing fitness or environmental conditions may differ from evolution under fixed ones.

- *Characterize multi-level MH structures:* We will also consider theoretical MH structures with multiple levels. We are interested in studying the evolution of substructures by implementing a version of the GA where, instead of evolving structures with a given number of levels, we continue adding levels while pruning nodes that do not contribute to the dynamics of nodes at the highest level. This constraint may impose conditions on the scaling properties of the overall emerging structure, in a way somewhat analogous to the Kolmogorov scaling-laws of fully developed turbulence in fluid dynamics (105–107), which can be deduced from the conservation of energy as it flows between eddies at different scales. We will explore which algorithms produce statistically stationary networks and study their properties.

5.4.   The Network State Diffusion Model

In order to consider the effect of spatial dynamics in MH structures and to more closely match experimental systems, we define a dynamical *network state diffusion* (NSD) model that takes into account the physical constraint of having mainly diffusive interactions at the intercellular level of the interaction hierarchy. It was recently developed in collaboration with Prof. Brockmann and combines complex network dynamical systems with reaction-diffusion dynamics. Each reflects a different hierarchical level or scale in the MH structures, here given by a spatially distributed system of interacting cells (Fig. 4a). At the intercellular level, these are coupled by diffusive signaling. At the intracellular level, cell phenotype selection, differentiation and the dynamics of cell states result from the regulatory dynamics of interacting genes. These processes can be modeled using high-dimensional network-based dynamical systems, where the degrees of freedom describe gene expression levels that interact as a set of dynamic variables through an underlying network of regulatory interactions.

The dynamical system for each independent cell takes the form $\partial_t \mathbf{u} = f(\mathbf{u}; \mathbf{W})$, where elements W_{ij} of coupling matrix $\mathbf{W}$ denote interaction strengths between genes. Gene expression is denoted by the elements of the vector $\mathbf{u}$. A generic choice for the nonlinear dynamical system can be derived from underlying Michaelis-Menten kinetics, which yields Hill-type functions such as

$$\partial_t\, u_i = -\gamma\, u_i + \sum W_{ij}\, \frac{u_j^{\beta}}{\theta_j^{\beta} + u_j^{\beta}}$$

This type of approach for understanding cell fate and state has been applied in many single cell models, e.g. to address how the stability of sequential differentiation cascades depends on topological features of the underlying regulatory network. At the scale of cell populations, simplified reaction diffusion models are typically employed, where the concentration of signaling molecules are modeled by dynamical quantities that interact through simplified, low-dimensional phenomenological diffusion-based signaling descriptions. These models can account for tissue differentiation in simplified contexts but cannot account for each cell's internal response to external signaling profiles. NSD models combine both approaches and can be cast into the form $\partial_t \mathbf{u} = f(\mathbf{u}; \mathbf{W}) + \mathbf{D}\, \partial_x^2 \mathbf{u}$, where matrix $\mathbf{D}$ is assumed to be a sparse matrix of diffusion coefficients with non-zero elements corresponding to gene signals between cells. NSD systems

can generate a rich variety of dynamical phenomena. For example, if the individual cell components possess multiple stable solutions, these network states will "diffuse" through space and potentially trigger a propagating wave of state changes across a population of cells. NSD systems correspond directly to 2 level MH structures where local modules are spatial replicates of the intracellular coupling of interacting genes, and cross-module couplings are the diffusive links between cells mediated by signaling nodes.

The interest of NSD models is that they properly represent the distinct physical mechanisms that mediate interactions at different hierarchical levels, corresponding here to intracellular and intercellular levels. This will allow us to design simulations and experiments to test key questions on MH systems, such as the robustness of internal network $\mathbf{W}$ and possible dynamical outcomes of cell state interactions at the population level. NSD models can also address evolution on MH systems; genetic heterogeneity and thus phenotypic variability as a function of position can be introduced as an explicit dependence of the regulatory networks on spatial coordinates $\mathbf{W} = \mathbf{W}(\mathbf{x})$. Evolutionary dynamics can thus be modeled by considering slow topological modulations of $\mathbf{W}$, driven by selection rules on states of the entire system.
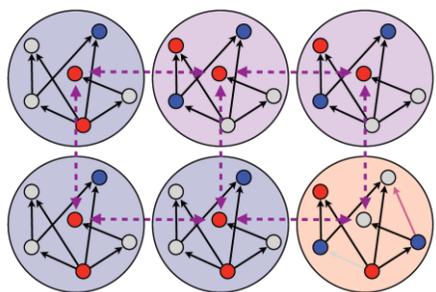

Figure 4: (a) Example of network state diffusion model; a spatial cell lattice of networks of interacting genes (black arrows). All cells are "clones" (they contain the same network), except for the cell in the lower right. Each network produces various stable activation states, illustrated by node colors. Cells interact via diffusive signaling that couples networks at the next hierarchical level (purple dashed lines), allowing the generation of spatial patterns.
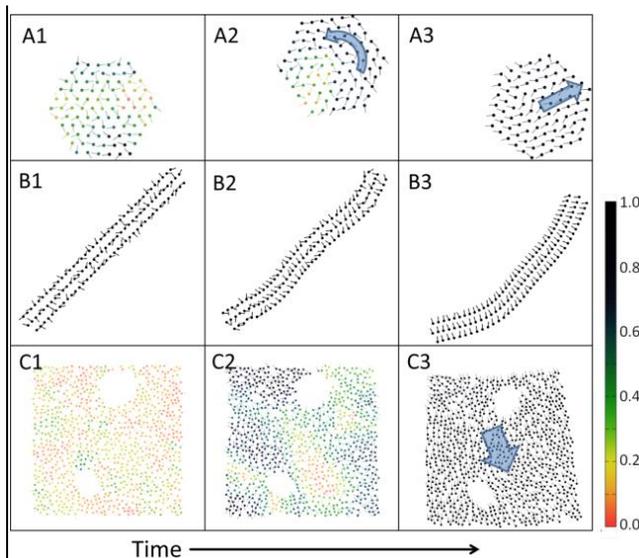
(b) Simulations of three active elastic modules, each with its own characteristic dynamics resulting from the AEM model (11,108). Nearby agents are linked by virtual springs. Coherent dynamics spontaneously emerge from initially randomly ordered active components. The color scale indicates degree of local order. MH structures and dynamics are produced by combining such modules into a hierarchy of increasingly larger AEM systems.

5.5. The Active Matter Approach

The Active Matter Approach will be based on a family of models of active agents (which have their own source of mechanical energy, such as self-propelled agents) connected through spring-like interactions. Mechanical analogies have a long tradition in physics as prototypical systems that are reasonably simple, have practical realizations, and help build intuition and explore classes of solutions. A relatively recent example is the case of the double-pendulum in the study of chaotic dynamics. We will focus below on a specific type of active matter model, but envision also implementing other variations.

The Active Elastic Membrane (AEM) model was recently introduced to describe groups of self-propelled agents permanently connected by virtual springs (11,108). Under certain conditions, individual velocities self-organize into growing regions of coherent motion through a mechanism based on standard elasticity, eventually reaching collective rotation or translation (Fig. 4b). This decentralized motion coordination is achieved without agents exchanging information on their velocities (which define here their states). Instead, interactions are based on positions (i.e. an integral of their states). Self-organization results from the cascading of energy to larger scales through the selection of collective elastic modes. The system thus naturally contains a physics- and selection-based self-organizing mechanism at the module level, which we can compare to consensus achieved through direct coupling of velocity states. It can represent an idealized case of morphological development and evolution.

Agents can be linked into structures with any shape, activity, or mechanical properties. We will simulate physical representations of MH structures by replacing on MH networks, like those on Fig. 2, nodes by self-propelled agents and links by springs. This produces MH dynamics on systems whose only phenotypic trait is motion in physical space. Modules become internally self-organized and then interact with other modules to achieve coherent motion at the next hierarchical level. This captures an essential feature of biological systems: the need to self-organize the energy produced at the individual component level before it can effectively do work, and power the next hierarchical level. Despite its simplicity, the MH-AEM model displays many characteristic features of MH systems. It will thus allow us to study in a concrete nonequilibrium physical framework the self-organization and evolution of MSMH structures.

We will implement interactive simulations of AEM structures composed of up to thousands of self-propelled agents that will illustrate visually MH dynamics and evolution and serve as a basis for our online game and crowdsourcing experiments (see **Section 8.1**). These simulations will place different AEM structures on a virtual arena to compete for randomly distributed resources or 'nutrients' in an artificial life setting. Since these AEMs self-organize into spontaneous coherent motion of their whole and parts that is determined by their structure, their design will determine their ability to reach these nutrients more effectively than others. We will mutate these structures and impose selective pressures based on this ability to find nutrients or other mobility-based fitness criteria, producing simulations where AEM-MH structures are visualized as they evolve. These will also help in outreach efforts and to develop intuition.

6. Experimental and Data-Driven Collaborations and Activities

6.1. Constructing and Perturbing Cellular MH Networks

Our collaboration with Prof. Leonard (Northwestern U) will connect concepts of origin and function of MH biological structures to ongoing synthetic biology experiments performed in his lab. The collaboration has identified specific questions on which MH theory may shed new light and experimental tests derived from theoretical work that can be implemented using the techniques available at the Leonard Lab.

In recent years, the Leonard Lab has unveiled mechanisms by which immune cells make "decisions" in complex environments at the level of individual and cell population levels. In this context, this collaboration will aim to (a) understand potential functions of natural immune networks and (b) predict perturbations to these that could produce specific (potentially therapeutic) functions. Understanding the role of MH structures on dynamics, robustness, adaptability, and evolution is thus tightly related to current questions investigated by Leonard, which include: how do our immune systems handle adapt collectively to a range of potential assaults? How do small differences in the function of "parts" (genes, receptors, etc.) within immune systems results in different collective behaviors of these multicellular networks?

The Leonard Lab has developed a suite of mammalian synthetic biology tools and techniques, which enable them to custom engineer input-output behaviors. They pioneered a Modular Extracellular Sensing Architecture (MESA) that allows engineering biosensor proteins that transduce detection of exclusively extracellular sensing and signaling cues (such as cytokines and other proteins) into changes in intracellular state. MESA does not rely on native proteins or pathways, and can be integrated into gene circuits to create devices that perform operations such as multiparametric logical evaluation of extracellular cues. This technology is ideally suited for studying MS structures in cellular systems, since it allows the engineering of two clearly identified hierarchical levels, the intracellular and the extracellular networks. This system can also be well described by the NSD model presented in **Section 5.4**.

We will also collaborate with Prof. Guet (IST, Vienna) in this effort for constructing and perturbing cellular MH networks. Prof. Guet also works on experimental synthetic and systems biology, but focuses on understanding systems of interacting genes and proteins that constitute themselves into genetic networks in bacteria, in particular in *E. Coli*, a model system of relative simplicity and for which powerful experimental genetic tools available. These bio-molecular networks are involved in a constant process of decision making and computation taking place at multiple spatial and temporal scales (from molecular-scale process that occur on the order of seconds to the division time of an organism and beyond), which can be identified as MH process. We will explore the measurable effects of underlying MH structures in the experimental systems studied by the Guet group. We will studying existing networks and construct new ones in living cells using synthetic biology approaches, aiming to uncover universal rules that govern the role of such MH dynamics and their underlying structures in biological genetic networks.

An emphasis of the experimental techniques of the Guet group is on understanding the molecular biology and physiology of the single cell, since often population level measurements mask the behavior of the individuals. To this end, they have developed minimally invasive in vivo techniques to quantitatively characterize the temporal dynamics of gene expression. These techniques will allow testing the connection between modular-hierarchical structure and function, using available data or designing new synthetic biology experiments that test predictions resulting from the theoretical aspects of this project.

6.2. Relating MH Structure and Function

Our collaboration with Prof. Amaral (Northwestern U) will aim to understand the relationship between MH structures and function in data from cell metabolic biomass production networks (39). Understanding and modeling the process through which microbes use nutrients in their environment to grow is of fundamental importance in cell biology, physiology, medicine, and bioengineering; and of practical importance for growing microbes in laboratories. The Amaral Lab has developed quantitative algorithms that have allowed them to identify the hierarchy of modules that characterizes the biomass production network of *E. coli* and other organisms, including a global metabolic network constructed from all corresponding available organismal networks. They have combined this analysis with an application, MetExplore, which allows interactive visualizations of systems of metabolites and reactions as a collection of embedded networks, where each node can be individually viewed as a lower-level network or module.

We have been collaborating with Prof. Amaral during the last few months, taking the first steps to understand if and how the hierarchy of modules that characterizes the structure of certain metabolic networks is related to the functional modularity of their dynamics. We are combining the Amaral Lab tools described above with Flux Balanced Analysis, in order to compare the visualized network structures to functional modules of biomass production that are activated under different conditions of metabolite availability. This will be a pioneering comparison of MH topology and MH dynamics, which we will develop into analyses of structure and function in a variety of experimental and theoretical MS systems.

6.3. Analyzing Coupled MH Networks of Invasive and Host Organisms

Our collaboration with Prof. Mandel (Northwestern U) will aim to improve our understanding of how modular processes at the bacterial scale and at the host organism scale are coupled in the beneficial colonization of Euprymna scolopes squid by Vibrio fischeri bacteria. This system is a valuable model of bacterial colonization as only a single bacterial species colonizes the animal's light organ, allowing for the type of reductionist analysis considered by this proposal in the natural context. The Mandel Lab has identified specific genes that have profound consequences on the colonization of the host and its subsequent development. They maintain a collection of strains from a variety of squid hosts which can be analyzed using bacterial genetics and comparative genomics approaches to discover bacterial genes that influence colonization. They also conduct phylogenetic and evolutionary approaches to determine the consequences of laboratory experiments in natural populations in an effort to understand evolutionary processes that underlie microbe-host associations.

Prof. Mandel will share data and provide expertise in the squid-Vibrio system. We will study in numerical models what the expected coupling of structural and functional variations of the host and guest organisms, given the MH perspective and compare results with these data. As coevolution of a symbiotic system must be successful at every stage of its development, the modular dynamical changes driven by mutations or adaptations of one of the organisms must be followed in parallel by the other. We will search for experimentally verifiable characteristic couplings between symbiotic species due to this process.

Our collaboration with Prof. Huang (ISB, Seattle) will consider a specific example model system of tumor biology that will help frame and search for potential applications of MH theory and models. The aim is to understand how tumor systems develop modules and higher levels of hierarchies that increase their resilience. His Lab works on gene regulator networks and on how alterations in their structure and dynamics lead to cancer. In the case of cancer, the progression of tumors is driven by a somatic evolution that is more complex than a simple competition of randomly mutating cells. This is due to the presence of a set of evolving modules of pathways that govern distinct modular functions (such as resilience to hypoxia, cell invasion, angiogenesis, colonization of distant tissues, etc.), which appear to accelerate evolution. Furthermore, it has been recently shown that cancer cells in a tumor can organize at a higher hierarchical level: they become functionally distinct, developing modules that take on distinct functions critical for the survival of the tumor, such defense against host immune attack or proliferation. In a recent

publication, Prof. Huang's group showed that it is the adaptation of these modules rather than single-cell evolution that can be responsible for the rapid development of cancer-drug resistance (109).

Prof. Huang will share with us an immense collection of existing and new data at the Inst. for Systems Biology of gene regulatory network dynamics of evolving tumor cells. This includes recently collected data on cell-cell communication networks and how they change as tumors evolve to become drug resistant. He will also develop new experiments that could test theoretical results on MH structures. The single-cell resolution measurements of gene regulatory network variables by Dr. Huang's lab is a perfect complement for the type of modular hierarchical models that are a central part of this proposal.
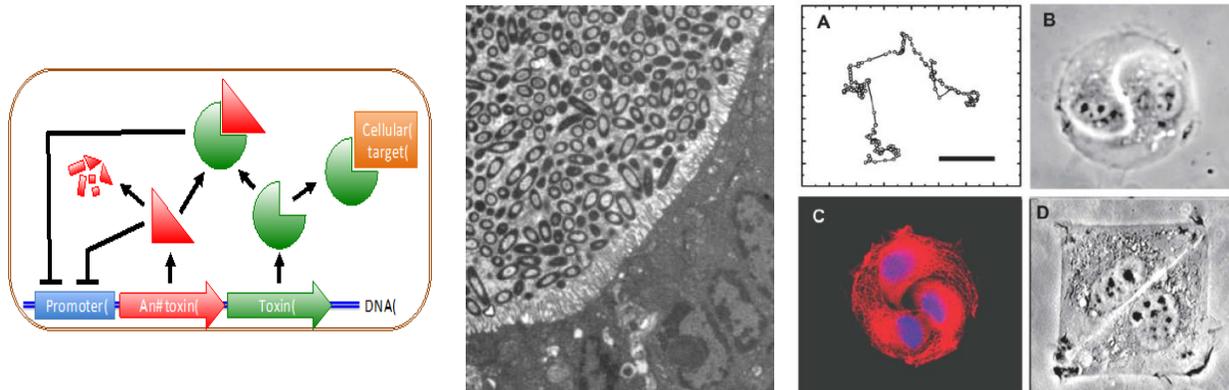

Figure 5: Illustrations of some of the systems that will be studied with external experiental collaborators. Left: Toxin-antitoxin modular process that is being implemented by the Leonard Lab. Center: Image of bacteria populating the squid's light organ in the symbiotic system studied by the Mandel lab. Right: Image showing cohort cell migration due to modular task specialization in a mamallian cell system at Huang Lab (Images courtesy of these collaborators).

7. <u>Integrated Research Objectives</u>

We will build on the theoretical approaches and experimental collaborations described above to study, in parallel, generic and specific features of the MH systems under investigation. We will examine their origins, dynamical behavior, and biologically relevant functions through an integrated research plan divided into three overarching aims, each addressing a set of specific questions through complementary investigations. Aim 1 is to generate and characterize MH structures, Aim 2 to explore their dynamics, and Aim 3 to examine how they evolve and adapt over time.

7.1. Aim 1: Develop framework for generating, classifying, and characterizing MH structures

We will first develop and characterize libraries of theoretically-generated MH interaction networks and compare them with interaction networks described by collaborators. We will develop metrics and classification schemes that facilitate the comparison of diverse systems. Key questions will include:

- How should we identify and classify MH topologies and functions? Can we find universality classes based on their topology, dynamics, or generating algorithms?
- Can we adequately describe some of the properties of theoretical MH topologies analytically?
- How sensitive are the MH network generating algorithms to noise or link failure? Do different MH topologies exhibit different levels of robustness (or other features) to structural perturbations?
- What topological properties characterize the natural MH networks identified by our collaborators?
- What duplication-based algorithms produce MH networks with topologies similar to those found in the natural and engineered biological networks studied by collaborators?

7.2. Aim 2: Explore the dynamics supported by MH structures

In this Aim we will systematically study the properties of dynamical processes supported by MH systems. We will identify their characteristic features by systematically studying processes on various fixed topologies with different degrees of MH organization. This will also help characterize MH networks (**Aim 1**) and generate dynamics for the artificial life simulations in **Aim 3**. Key questions are:

- What characterizes the dynamics of processes on MH systems and their response to external stimuli? Are the dynamics of network states different at different levels of an MH structure?
- How do dynamical perturbations at one level of an MH structure propagate within it and to other levels? Is there a time-scale separation that emerges from the MH topology (as claimed in (110))?
- How important are the physical constraints in spatial (NSD and AEM-MH) models for their resulting dynamics? Can we identify experimental systems with similar physical constraints and/or dynamics?
- Which theoretical structures best capture the dynamics of our collaborators' MH systems? How are their structures and dynamics coupled? (In particular in the system described in **Section 6.2**.)

7.3. Aim 3: Investigate the evolvability and adaptability of MH structures

Building on the products of **Aims 1** and **2**, we will investigate how different MH structures impact a system's ability to adapt and evolve under changing conditions. In our theoretical framework, evolvability and adaptability are not well distinguished; we can interpret the former as changes in network topology or dynamical rules (genotype) and the latter as changes of the attractor accessed by network states (phenotype). We will implement evolutionary algorithms and simple artificial-life simulations designed to study the emergence and stability of MH networks. We will focus on the selection and adaptability of the MH models described in this proposal, using different levels of modularity and hierarchical organization. This effort will be initially motivated by recent work showing that a variety of network systems evolve towards modular solutions when alternating fitness criteria are imposed (66–68). Specific question are:

- Do systems with higher modularity exhibit greater evolvability and/or adaptability under changing environmental conditions or fitness criteria? What mechanisms can enhance modularity?
- Can we extend the emergence of modularity to evolve MH structures that spontaneously develop two or more hierarchical levels? Can we design a synthetic biology system with our collaborators that demonstrate the adaptability and evolvability of MH structures?
- How stable are MH structures from an evolutionary viewpoint? How do theoretical and experimental MH systems react to mutations or stresses imposed at different levels in their hierarchies?
- Do MH systems imply specific couplings between host (higher, embedding hierarchical level) and microbiota (lower, embedded level) in a system like that described in **Section 7.3**?

8. Broader Impacts

The proposed research on how MH systems are structured, behave, and evolve could lead to breakthroughs in the beneficial utilization and manipulation of living systems, in the design of biologically inspired engineering solutions, and in the fundamental cross-disciplinary understanding of biological dynamics and self-organization. Our work could improve medical practice by helping identify critical nodes or structures in human or pathogen physiology that convey functional properties such as adaptability or robustness. Understanding how evolution has shaped biological MH networks to achieve desirable properties could help us design engineered systems that exhibit similar features. Such systems could include engineered biological functions (built using synthetic biology or other approaches) that enable the sustainable and cost-effective production of valuable chemicals, materials, or fuels. They could also include technological and robotic networks, where our increasingly interconnected world has a growing need for designs that are both robust and adaptable. Furthermore, our work will develop new approaches within nonlinear, statistical, and non-equilibrium physics that build upon the deep and extensive body of work in these areas to extend their applicability to MH systems and thus catalyze new contributions from these fields to the applications described above. Finally, we will build on our record of science outreach to conduct a series of specific activities that leverage the results of the proposed research to engage, educate, and inform constituencies ranging from students to the public at large.

8.1. Self-organized modular hierarchical motion – A crowdsourcing experiment

We will develop a crowdsourcing experiment based on project ideas, implementing an online game where participants are challenged to design MH-AEM structures (see **Section 5.5**) that compete with other MH-AEM's (that were either evolved or designed by other users) to achieve specific objectives. Note that many different virtual MH-AEM 'machines' can be engineered by adjusting the springs and self-propelled properties to create modules that self-organize to move, rotate, or deform in specific ways. The application will be designed to establish a library of user-created modules that can be combined to build

more complex machines. MH design will thus be encourages and naturally emerge (111,112), leading to MH structures. The AEM model has already been implemented as a Java application that will be set up in Amazon's Mechanical Turk or other crowdsourcing platform. This unique game-based crowdsourcing experiment will allow a fascinating comparison between evolved and engineered solutions to the same problems. It has the potential to generate widespread public interest (that we will encourage through press contacts), since it combines a mechanical analogy that illustrates the dynamical and evolutionary properties of MH systems for a broad audience with game-based crowdsourcing where participants can help our research by providing solution designs to evolutionary challenges.

8.2. Other outreach activities

We will maximize the impact of this work on diverse fields by actively disseminating its results to audiences that span the biological sciences, engineering, and physics. Dissemination will be enhanced through the interactions with multiple national and international external collaborators. We will target specialized physics and biology publications together with high-profile, broadly read journals such as Nature, Science, PNAS and PLoS. The postdoctoral researcher will present project results at large international physics meetings, as described in the mentoring plan. We will also target presentations at the highly interdisciplinary International Conference on Network Science (NetSci). A project-specific website will be established to provide information for both professional scientists and interested non-professionals. The PI has a track record of press and media interactions and cultivates relationships with journalist, which will facilitate the dissemination of project results to the broader public. In the past, this has led to news features in Wired Magazine, National Geographic and other high-profile national and international media. Moreover, Dr. Huepe is also a professional musician involved in electronic music and in various art/science projects that reach a broad nonscientific audience, and will incorporate concepts, images, and data from the project in his musical creations and performances. Finally, we will also increase broader impact on STEM education by participating in programs at Northwestern U that allow undergraduates to get involved in short-term research efforts, which will be mentored by the PI.

9. Project Management

This project has unique management requirements since it does not correspond to a traditional single-theory/experiment collaboration, but aims instead to compound existing knowledge and diverse experimental evidence from external collaborators and to use these to achieve integrative theoretical progress. The requested resources are to allow the PI and Postdoctoral Research Associate (PRA) the necessary time-dedication and provide the tools to develop theoretical and computational research and to interact with the experimental and data-driven external (not funded by this proposal) collaborators.

Due to the nationwide and international nature of the collaborations, the project will require effective management of long-distance scientific interactions. The PI is uniquely positioned to address this need, since he is based in Chicago and an established visiting researcher at the Northwestern University's Institute on Complex Systems (NICO – Northwestern University), while keeping the freedom to spend typically 2 to 4 months per year travelling for scientific visits, mainly in Europe. The PI will thus be able to interact in person with the external collaborators throughout the project. Our proposed budget includes resources to also enable the PRA to visit at least annually other national and international labs, in order to facilitate research with collaborators. It also allows for brief scientific visits of collaborators to Chicago.

The PI has an established track record of long-distance collaborations and co-supervision of trainees. As described in detail in the **Postdoctoral Researcher Mentoring Plan**, the PRA will be based at NICO, working under the PI's close supervision. Video conference meetings will be held at least every two weeks with one or more of the collaborators and include research updates from the PRA. We plan on having at least two work visits with each external collaborator during the duration of the project, where overall research progress, coordination, and potential applications will be discussed; and upcoming challenges and opportunities will be identified and addressed. The collaborators will provide experimental and diverse scientific interactions together with consultation to the PI and PRA, which will ensure the success of the project and to catalyze the dissemination of its products to diverse disciplines (see **Letters of Commitment to Collaborate** in **Supplementary Documents**). The tools and practices developed in previous experiences of successful long-distance and international collaborations by the PI (e.g. online collaborative resources) will be harnessed to ensure the effective integration of this project.