Collective Dynamics in Physical and Social Networks

Collective Dynamics in Physical and Social Networks

A dissertation presented

by

Alexander Isakov

to

The Department of Physics

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Physics

Harvard University

Cambridge, Massachusetts

April 2016

Dissertation Advisor
Professor L. Mahadevan

Author
Alexander Isakov

Collective Dynamics in Physical and Social Networks

Abstract

We study four systems where individual units come together to display a range of collective behavior. First, we consider a physical system of phase oscillators on a network that expands the Kuramoto model to include oscillator-network interactions and the presence of noise: using a Hebbian-like learning rule, oscillators that synchronize in turn strengthen their connections to each other. We find that the average degree of connectivity strongly affects rates of flipping between aligned and anti-aligned states, and that this result persists to the case of complex networks.

Turning to a fully multi-player, multi-strategy evolutionary dynamics model of cooperating bacteria that change who they give resources to and take resources from, we find several regimes that give rise to high levels of collective structure in the resulting networks. In this setting, we also explore the conditions in which an intervention that affects cooperation itself (e.g. "seeding the network with defectors") can lead to wiping out an infection. We find a non-monotonic connection between the percent of disabled cooperation and cure rate, suggesting that in some regimes a limited perturbation can lead to total population collapse.

At a larger scale, we study how the locomotor system recovers after amputation in fruit flies. Through experiment and a theoretical model of multi-legged motion controlled by neural oscillators, we find that proprioception plays a role in the ability of flies to control leg forces appropriately to recover from a large initial turning bias induced by the injury.

Finally, at the human scale, we consider a social network in a traditional society in Africa to understand how social ties lead to group formation for collective action (stealth raids). We identify critical and distinct roles for both leadership (important for catalyzing a group) and friendship (important for final composition). We conclude with prospects for future work.

Table of Contents

List of Figures

List of Tables

Citations to Prior Work

The work on the phase oscillator model in Chapter 2 was published as:

> **Isakov, A. and Mahadevan, L.** (2014). Synchronization in a stochastic Hebbian network of phase oscillators. *ArXiV*: 1404.2328

The work on interdependent population discussed in Chapter 3 was published as:

> **Vural D. C., Isakov, A. and Mahadevan, L.** (2015). The organization and control of an evolving interdependent population. *J. R. Soc. Interface* **12**(108), 20150044. DOI: 10.1098/rsif.2015.0044

The work on the recovery of locomotor bias after leg amputation in fruit flies discussed in Chapter 4 was published as:

> **Isakov, A., Buchanan, S. M., Sullivan, B., Ramachandran, A., Chapman, J. K. S., Lu, E. S., Mahadevan, L. and de Bivort, B.** (In press). Recovery of locomotion after injury in *Drosophila* depends on proprioception. *J. Exp. Biol*. DOI: 10.1242/jeb.133652. Available online at: http://jeb.biologists.org/content/early/2016/03/17/jeb.133652

The work on the emergence of violence in human social networks discussed in Chapter 5 is in preparation for peer review as:

> **Glowacki, L., Isakov, A., Wrangham, R. W., McDermott, R., Fowler, J. H. and Christakis, N. A.** (In preparation for *Proc. Natl. Acad. Sci. USA*). Social networks and the emergence of violence.

Additional work during my graduate career in which I was first or joint first author that does not appear as a chapter in this thesis has appeared as:

> **Isakov, A., Murdaugh, K. M., Burke, W. C., Zimmerman, S., Roche, E., Holland, D., Einarsson, J. I. et al.** (2014). A new laparoscopic morcellator using an actuated wire mesh and bag. *J. Med. Devices* **8**(1), 011009. DOI: 10.1115/1.4026294

> **Roithmayr, D., Isakov, A. and Rand, D. G.** (2015). Should law keep pace with society? Relative update rates determine the co-evolution of institutional punishment and citizen contributions to public goods. *Games* **6**(2), 124-149. DOI: 10.3390/g6020124

> **Isakov, A., Holcomb, A., Glowacki, L. and Christakis, N.** (2016). Modeling the role of networks and individual differences in inter-group violence. *PLoS One* **11**(2), e0148314. DOI: 10.1371/journal.pone.0148314

Several additional papers are currently in preparation.

Lovingly dedicated to my parents, Gregory and Leah Isakov.

Acknowledgements

I would like to start by warmly thanking my advisor, Professor L. Mahadevan. Over the years, we have discussed topics ranging from oscillators to the structure of human society, and it has been a true pleasure to work with a mentor who always encouraged me to expand my interests and to look for patterns throughout the natural world without being limited by the construct of "disciplines". Most of all, I would like to thank you for creating an atmosphere marked by warmth and kindness, and for allowing me exceptional intellectual freedom: my time with you has shaped me who I am as a scientist, and has set me on a wonderful course. I could not have asked for more!

I am likewise grateful to Professor Nicholas Christakis for advising me on the social science side, and for opening my eyes to the ways in which social networks mediate our lives. Thank you for sharing and continuing to share your warmth, patience, and deep insights into human nature with me – this has played a large role in my decision of the set of questions to pursue next. And, I am happy (though not surprised) that the environment you created in your new lab at Yale parallels the positive environment you fostered in Pforzheimer when I was an undergraduate at Harvard (including the nice way you treated me and my blockmates despite our perhaps sometimes being a tad too loud outside your residence a tad too late into the evening). I am delighted to have the opportunity to continue our work.

My graduate school experience at Harvard has been made exceptionally smooth by the wonderful guidance and administrative help I received from Dr. Jacob Barandes and Lisa Cacciabaudo in the physics department, and Claudia Stearns. My collaborations at Yale were expertly facilitated by Kimberly Kuzina. You are all excellent administrators and lovely people – thanks for answering all my questions promptly, accurately, and often (but I hope not too often) on short

notice. I was also fortunate enough to be awarded the National Defense Science and Engineering Graduate Fellowship, which funded me from 2012-2015 and gave me extended flexibility in my research.

Thank you to the many collaborators and mentors inside and outside of academia who inspired me, challenged me, provide me with opportunities for growth, worked with me side by side, and became my friends. I cannot possibly hope to list you all, but you know who you are. In academia, thank you in particular to Professor James Fowler, Dr. Luke Glowacki, Professor Benjamin de Bivort, Professor David Rand, Professor Dervis can Vural, Professor Hsueh-Ling Huynh, Professor Lene Hau, and Professor Jene Golovchenko. Outside of academia, I am beyond grateful to Professor Hans-Helmut Kotz, Dr. Garo Armen, George Castineiras, Stan Miroshnik, Lynn Rivenburgh, and Joe Eastin.

A special thanks to my mentees: Brian Sullivan, Akshitha Ramachandran, Ned Lu, and Amelia Holcomb – you were all terrific, and I'm grateful that you put up with letting me practice my managerial skills on you. We learned a lot working together and I hope that I was a generally good influence.

To my friends near and far I am grateful for more than I can relate, and I apologize in advance that I cannot list you all. Konrad Binienda, Zhao Chen, Alex Vai, Helmut and Inge Neukirch, Jake Connors, and Meredith McGregor: thanks for always having the time to talk and laugh, for forcing me to have fun, and for putting up with my weird sense of humor. Arnav Tripathy: thanks for too much, but at the very least for sweeping success in our joint pursuits (academic and otherwise) and for adventures of epic proportions. Kimberly Murdaugh (soon to be Dr. Kimberly Isakov): you are an amazing, exciting, awesome person, and a wonderful fiancée. I cannot do justice to how grateful I am to you. Thank you for always listening (even when I repeated

myself), for always being happy (hence making me happier, etc. – like a happiness-based perpetuum mobile), and for always being supportive. And for awesome adventures, of which there will be many, many more. Thank you likewise to Jackie Murdaugh, Kevin and Darlene Bernard, and Butch Corbitt – your kindness and support (and excellent Thanksgiving celebrations) were the highlights of each year.

Finally, I could not have done this without my family. You have given me infinite love, support, patience, advice, and inspiration. You challenged me to become better without judgment. You kept me grounded and focused on what is important. And you always did everything with a smile and a copious amount of humor – exactly the right amount. Mom (Dr. Leah Isakov) and Dad (Gregory Isakov): you're the best parents ever – thanks for always being insightful, for being proud of me, for actually reading my papers, and for helping me make the right choices. Dedushka Yasha (Professor Yakov Duboshinsky): I'm sorry you are no longer with us, and am grateful that you set me on the path to physics. I think of you as a shining example of what a scientist should be. I know that you would have approved of my thesis, and that our many conversations will continue to guide my scientific interests. Babushka Ira (Irina Duboshinsky), Dedushka Alec (Albert Isakov), Uncle Dania (Dr. Danil Duboshinsky), and Uncle Boris (Boris Duboshinsky): thanks for your love and support. Anna and Michael Isakov: you are awesome siblings, and I will spend more time with you when I am finished with this dissertation. Thank you for your love and for listening to me talk about my work. Hopefully I will be able to offer some useful advice when you go to graduate school.

"Nature does nothing uselessly."
- Aristotle, *Politics*

1

Introduction

Motivation

Much of the rich behavior observed in the real world arises as a consequence of a number of similar units interacting in a structured way. One need not look far for examples. In this work, I will reference a number of related scientific studies, thus playing a role in a citation network where "units" are papers (or scientists) and interactions are citations of prior papers. This network, which exhibits surprisingly regular statistical properties (Newman, 2001), represents the "scientific tradition" and can help both track and predict the evolution of research. On social networks, where agents are people and interactions are structured by e.g. friendship (Rapoport and Horvath, 1961), marriage (Padgett and Ansell, 1993), or being on the same corporate board (Davis and Greve, 1997), we find that there are predictable patterns of behavioral spread that in some instances lead to new norms or otherwise affect society at a large scale. For example, Davis and Greve found that practices such as "golden parachutes" in corporations spread from board to board until they became a new norm. And, it is now well-established that friendship ties up to three degrees affect our health, from directly influencing friends, to indirectly influencing friends-of-friends and friends-of-friends-of-

friends to quit smoking (Christakis and Fowler, 2007) to nudging them towards harmful eating behaviors that show themselves on a large scale in the "obesity epidemic" of recent years (Christakis and Fowler, 2008).

Likewise, there has been great theoretical and practical interest in physical and biological networks. One particularly exciting area of research is the organization of the brain. Here, the units are neurons and the connections are synapses. The first animal to have its neural wiring diagram ("connectome") completely explored was *C. elegans* (White et al., 1986). Despite having only 302 neurons (almost completely invariant across animals) and roughly 7000 synapses, these worms exhibit a range of complex behavior in addition to crawling (Wen et al., 2012; Paoletti and Mahadevan, 2014), such as sensing and reacting to e.g. heat (Garrity et al., 2010) and light (Ward et al., 2008), etc. At the same time, the incredible complexity of human behavior is governed by massive neural networks that have on the order of 10^{10} neurons (Herculano-Houzel, 2009).

While there has been a surge of research in collective dynamics on physical, biological, and social networks (for a classic broad overview, see (Newman, 2010)), an area that has been relatively underexplored is that of feedback between agents and network structure. Recently, biological and social network motivations have provided the need to consider activity-dependent interactions. The idea is easy to state: the underlying network dictates who interacts with whom, but specific interactions today change the strength of the interaction tomorrow; in other words, the network can itself changes depending on the interactions of the agents. Despite its seeming simplicity, this line of thinking holds great promise for moving towards more accurate modeling of many network-based phenomena because, in reality, networks grow and evolve in response to the activities that are occurring on them. For example, it is now known that changes in connections between neurons are related to the relative times between firing or the synchronization between them (Caporale and Dan,

2008; Stam et al., 2010). Even in the classical examples of synchronization in well-mixed clapping audiences (which can be viewed as an all-to-all network) (Néda et al., 2000), electromechanical oscillator systems (Landa and Duboshinsky, 1989), and circadian rhythms (Saper et al., 2005), one can discern that the coupling changes as a function of the current state of the system.

This work is organized as follows. We begin by introducing a twist to a classic example of a simple physical system that exhibits interesting collective dynamics (synchronization), the Kuramoto model (Kuramoto, 1984). Here, the units are phase oscillators, and the interaction consists of slowing down faster neighbors and speeding up slower neighbors. We investigate a generalized version of the model that introduces Hebb-like couplings (Hebb, 1949) that evolve according to a stochastic differential equation on various topologies. Using numerical simulations, we show that even with identical oscillators, there is a regime in the nearest-neighbor coupling topologies and complex network topologies where oscillators flip between in-phase and anti-phase states relative to their neighbors. Phase diagrams show the transition probabilities as a function of the noise strength and rate of evolution of network coupling, and we build a minimal theoretical model to understand these transitions.

In neural networks, this corresponds to having neurons which fire together to wire together (slowly). At the scale of brain regions, similar models have been put forward for studying regimes of cortical activity (background, resting-state, and epileptiform) (Sadilek and Thurner, 2015). And, in social phenomena, this may correspond to the ties between people changing based on the behavior of the people themselves (Yi et al., 2013); this general mechanism of network assortment has been proposed as a general mechanism for sustaining cooperation in structured populations (van Veelen et al., 2012).

In the next chapter, we apply dynamic networks and an evolutionary dynamics perspective to ask how interdependence evolves. Starting with Darwin, biologists have asked how populations evolve from a low fitness state that is evolutionarily stable to a high fitness state that is not. Specifically, there has been enduring interest in the emergence of cooperation and multicellularity where the fitness of individuals often appears in conflict with that of the population. Theories of social evolution and evolutionary game theory have produced a number of fruitful results employing single-strategy two-body framework (where the two-body condition can be interpreted more broadly, such as "agent-society" in the case of public goods games). These studies provide insight into how altruistic and cooperative behavior can emerge to maximize the fitness of the group despite the apparent fitness advantage of cheating individuals, and a general overview can be found in e.g. (Nowak, 2006; Perc et al., 2013), with a more experimental view on human cooperation in particular in e.g. (Rand and Nowak, 2013). In the context of the evolution of cooperation, these models typically investigate the outcome of repeated runs of the prisoner's dilemma between pairs of agents that have two strategies, cheating and cooperating, or of public goods games where cheating is interpreted as contributing less to society than your fair share, and cooperation is contributing at least your fair share. Variants of such models include using structured interactions, coupled populations, coevolution, stored reputation, punishment, and preferential or random partner choosing (West et al., 2006; Gross and Blasius, 2008; Roca et al., 2009; Perc and Szolnoki, 2010; Wu et al., 2010; Rand et al., 2011; Isakov and Rand, 2012; Perc et al., 2013; Wang et al., 2013; Roithmayr et al., 2015).

However, real life is more complicated in a number of ways. First, many actual games are massively multi-player (Connor, 2010; Archetti and Scheuring, 2012; Gokhale and Traulsen, 2014). The fitness of an organism may depend on its simultaneous relationship with multiple players. Second, biology allows for a much larger variety of internal states beyond cooperating or defecting.

For example, the genetic make-up of an organism may be suitable for cooperation with only an exclusive few, while some organisms may be incapable of defecting or cooperating all together. Third, real social evolution leads to highly organized dependence structures beyond the homogeneous mixtures or aggregates of cooperator-defector states that are often seen in evolutionary game models. From biochemical to societal scales, life organizes itself in highly complex arrangements of cliques, communities, cycles, and hierarchies.

So, we extend the traditional view by considering a multi-player, multi-state evolutionary game, in which the fitness of an agent is determined by its relationship to an arbitrary number of other agents. These changing dependences (agent strategies) form an evolving network structure. We show that populations organize themselves in one of four distinct phases of interdependence depending on one parameter, selection strength. We find that some of these phases involve the formation of specialized large-scale structures. But, cooperation in this manner is not always a universally desirable outcome. For instance, bacteria cooperating to make biofilms create interesting, complex, and (unfortunately) more antibiotic-resistant structures that cause many serious and hard to treat infections (Costerton et al., 1999). Thus, we also explore how manipulating some agents' strategies through an external perturbation (disrupting cooperation in a limited way) affects the whole network structure and identify parameters where the whole network is likely to collapse.

In our next study, we explore a system with a smaller number of units, but where each is of greater complexity: the locomotor system of a fruit fly. In a world where locomotion is critical for survival and reproduction for most animal species, a wide range of motor strategies are present: walking, swimming, crawling, gliding, and flying (Dickinson et al., 2000). Of course, the specific behavioral details of movement vary wildly even in the case of a specific modality such as legged locomotion, as seen in bipeds (Vaughan, 2003), quadrupeds (Alexander, 1984), various hexapods

(Cruse, 1976; Full and Tu, 1991; Grabowska et al., 2012; Mendes et al., 2013; Couzin-Fuchs et al., 2015), and octopods (Blichkan and Full, 1987). Even when the number of limbs is held constant the pattern of limb placement during locomotion (gait) can vary greatly within and between species, as reviewed in (Holmes et al., 2006) and (Borgmann and Buschges, 2015). One common thread in the study of locomotion is the importance of proprioception, the ability to sense (e.g. via stretch receptors) the relative position of moving body parts. The importance of proprioception in locomotion has been shown in a suite of experimental and theoretical studies across many animals, including cats (Lam and Pearson, 2001), mice (Akay et al., 2014; Takeoka et al., 2014), insects (Bässler, 1977; Bässler et al., 2007; Borgmann et al., 2009; Mendes et al., 2013), sea slugs (Jahan-Parwar and Fredman, 1978), and nematodes (Wen et al., 2012; Paoletti and Mahadevan, 2014).

In a natural setting, it is common for an organism to suffer injury to the appendages mediating locomotion over its lifetime. We combine these ideas and investigate the role of proprioception in recovery from injury in fruit flies, using *Drosophila melanogaster* as a model organism due to the rich collection of genetic and transgenic tools available in this species for mechanistic inquiry. In particular, we examine the recovery of walking following leg amputation. By analyzing video recordings, we find that whereas flies explore open arenas in a symmetric fashion pre-amputation, foreleg amputation induces a strong turning bias away from the side of the injury immediately after the surgery. However, unbiased walking behavior returns over time in wild type flies, while recovery is significantly impaired in proprioceptive mutants even after 72 hours of observation. To identify the biomechanical basis of this locomotor impairment and recovery, we then examine individual leg motion (gait) at a fine scale. We also build a minimal mathematical model that links neurodynamics to body mechanics during walking. Combining our experimental results and the model shows that redistributing leg forces between the right and left side enables the observed recovery.

Altogether, our study suggests that proprioceptive input from the intact limbs plays a critical role in the behavioral plasticity associated with locomotor recovery after injury. In other words, the units (legs) have to change their individual behavior to achieve the desired result (unbiased walking) at the level of the whole fly, and the ability to correctly interpret how to tune the leg behavior likely comes at least in part from the sensory feedback loop provided by intact proprioception mechanisms.

In our last study, we return to a larger network of very complex agents: humans. Intergroup violence is common, worldwide, and harmful; it has also played an important role in the evolution of human behavior. Global annual deaths from warfare, for example, range from 0.5 to 1 million, and this does not include non-fatal physical and mental injuries (Lozano et al., 2012). Despite this harm, warfare may have had an important role in the evolution of modern human behavior in part because it often requires solving a collective action problem (Kelly, 2005; Choi and Bowles, 2007; Bowles, 2009).

Our closest relatives the chimpanzees engage in behavior strikingly akin to human raiding parties where small groups seek to ambush and kill members of other chimpanzee groups (Wilson et al., 2014). And, many hunter-gatherer societies have small-scale warfare with neighboring groups. These conflicts consist primarily of surprise raids and ambushes in which members of the attacking groups are only infrequently killed or injured (Wrangham and Glowacki, 2012; Gat, 2015). Unlike modern state warfare, small-scale warfare occurs outside of formal institutions, such as militaries, and there are no explicit requirements for participation. Successful warriors commonly receive individual benefits, such as status or captives, but their contributions can also create group benefits including access to territory and rival reduction (Glowacki and Wrangham, 2013). The production of group-level benefits from warfare creates a collective action problem because participants engage in

personally costly activity that produces benefits shared by the entire group – including non-participants (Mathew and Boyd, 2011; Glowacki and Wrangham, 2013).

The ability to solve such collective action problems is an important factor in the success of our species. Understanding how groups solve risky collective action problems outside of formal institutions can yield fundamental insights into the origins of human prosociality and the emergence of group-functional behavior in humans. Yet, there is little naturalistic behavioral data on how small-scale populations solve collective action problems (von Rueden et al., 2015). Individual differences have been hypothesized to promote costly cooperation across both humans and other primates (Modlmeier et al., 2014; McAuliffe et al., 2015; von Rueden et al., 2015; Barta, 2016). Recent theoretical work shows that heterogeneity within a group may increase within-group cooperation (Gavrilets, 2015) and contribute to group success, including in between-group competition (Gavrilets and Fortunato, 2014; Molleman et al., 2014). Studies with wild chimpanzees support the importance of individual differences in solving collective action problems. For example, a few "impact" chimpanzees show exceptional motivation to participate in boundary patrols and hunting (Gilby et al., 2008; Gilby et al., 2015). When chimpanzees hunt as a group, the presence of such highly motivated "impact individuals" who initiate group hunts appears to catalyze group hunting events by subverting prey defenses and lowering the costs of hunting to other chimpanzees (Gilby et al., 2015). Similar group dynamics with key individuals may be at work in other instances of collective action in animals suggesting the importance of individual differences in promoting risky collective action (Modlmeier et al., 2014; McAuliffe et al., 2015). However, data from human populations to inform these models is scarce.

We study the formation of stealth-raiding parties in a group of East African nomadic pastoralists engaged in small-scale warfare. We collect personal characteristics and social

relationships, including mapping their social networks. We show that the initiation of raids depends on the presence of key individuals who participate in many raids, have more friends, and occupy more central positions in the network. However, raid participants are recruited from the whole population, not just from the friends of leaders. Moreover, non-leaders have a larger total impact on raid participation than leaders, in spite of the greater connectivity of leaders. In other words, group formation (and the decision-making for joining or not joining) is driven by more complicated processes than simple direct recruitment. Overall, we find that leaders matter more for raid initiation than for participant mobilization. Since different social positions in the network play different roles in the emergence and amplification of group formation for between-group violence, there are opportunities in future work for asking questions in the same spirit as our other studies: both about controlling violence in networks with targeted behavioral interventions and about how particular roles, such as leadership, arise in the context of networks evolving in response to the prior behavior of particular agents. Studying such questions will lend deeper insight into the dynamic world around us.

Computation

Simulations for Chapters 2 and 3 were performed in Matlab 2012a (The MathWorks Inc., 2012). For Chapter 4, image analysis was implemented using Matlab 2012a with the Image Processing and Statistics Toolboxes. Scripts for determining locomotion turning bias, the calibration curve path simulation and the physical model simulation were also implemented in Matlab. All other analyses were performed using the statistical software R (R Foundation for Statistical Computing, 2014). For Chapter 5, all analyses were performed in R. Network figures were generated using Cytoscape (Shannon et al., 2003). For Chapters 4 and 5, we used the following R packages for analysis and generating figures/tables: igraph, HMM, mixtools, vegan, lfe, plyr, ggplot2, R.matlab,

and stringr (package descriptions, including full author information, are available at: *http://cran.r-project.org/web/packages/*).

"μηδὲν ἄγαν [nothing in excess]"
- Inscription on the Temple of Apollo at Delphi

2

Synchronization of Phase Oscillators on a Dynamic Network

Introduction

A fundamental question in studying self-organizing phenomena is determining what broad classes of conditions lead to which types of possible states of synchronization. From a mathematical point of view (Golubitsky et al., 2004; Golubitsky et al., 2005), symmetry considerations can provide some guidelines for possible equilibrium states, with certain limitations on the allowable types of systems. Perhaps one of the simplest general models of synchronization is the case of coupled phase oscillators. The framework of the eponymous model was analyzed and popularized by Y. Kuramoto and has since seen extensive study in physics, biology, game theory, and other disciplines as a stepping stone towards understanding more complex systems (Wiesenfeld et al., 1998; Mertens and Weaver, 2011; Yin et al., 2012; Timms and English, 2013). In the simplest case, the N oscillators are described by

$$\dot{\phi}_i = \omega_i - \frac{K}{N}\sum_{j \in S_i} \sin(\phi_i - \phi_j), \qquad i = 1, \ldots, N \tag{2.1}$$

where S_i is the set of neighbors of oscillator i, not including itself. Here $\phi_i(t)$ is the phase of oscillator i at time t and ω_i is its natural frequency. The coupling K is usually assumed to be constant and positive, leading to attractive interactions: slower oscillators are sped up by their neighbors and faster oscillators are slowed down by their neighbors.

This model acts as a springboard to a set of natural generalizations such as changing the connectivity topology S, drawing ω_i from various distributions, adding noise terms to the equation, and considering more general coupling functions (Acebrón et al., 2005). In order to better understand synchronization in such a physical system when taking into account agent-network interactions, we consider the case when the coupling strength obeys its own dynamics on a time-scale which is slow compared to the phase dynamics of the oscillators (Aoki and Aoyagi, 2009; Yuan and Zhou, 2011; Skardal et al., 2014), in the spirit of a Hebbian network (Hebb, 1949).

Model

We started with a modified Kuramoto model where the slowly evolving coupling obeys the simplest expected symmetry laws and is affected by a noisy environment, more closely mimicking a real biological or social system. For simplicity, we assumed identical oscillators ($\forall i \in \{1, \ldots, N\}\ \omega_i = \omega$) and rescaled natural frequencies by ω. To ensure that each neighbor contributes a fraction proportional to the number of neighbors to the frequency of oscillator i, we let $\alpha = \frac{1}{|S_i|+1}$; here the form of the denominator allow the oscillator to "affect itself". The set of equations that describes this system is

$$\dot{\phi}_i = 1 - \alpha_i \sum_{j \in S_i} K_{ij} \sin\left(\phi_i - \phi_j\right) + \sigma_1 \zeta_{1;i}, \qquad i = 1, \dots, N \tag{2.2}$$

$$\dot{K}_{ij} = \begin{cases} \epsilon\left(-K_{ij} + \cos\left(\phi_i - \phi_j\right)\right) + \sigma \zeta_{2;i,j} & j \in S_i \\ 0 & \text{Otherwise} \end{cases} \tag{2.3}$$

The long time-scale $\tau \sim \epsilon^{-1} \gg 1$ describes the slow evolution of the network coupling. We note that the symmetry of coupling with $K_{ij} = K_{ji}$ is preserved by the given functional form. The noise ζ is Gaussian with $\langle \zeta \rangle = 0, \langle \zeta_i(t) \zeta_j(t') \rangle = \delta(t - t') \delta_{ij}$ and standard deviation σ_1, σ, where $\sigma_1 \ll 1$ to prevent the system from locking when $(\phi_i - \phi_j) \in \{0, \pi\}$. Thus, the model is characterized by the two parameters ϵ, σ. In order to complete the model description, we specified a coupling topology. Initially, we considered a ring topology (Figure 2.1A) and a 2-dimensional nearest neighbor model with periodic boundary conditions (Figure 2.1D).

Below, we relax the symmetry assumptions and discuss the case of a modified preferential attachment model (Barabási and Albert, 1999) whose mean mimics the properties of the 2-dimensional lattice. For the ring, labeling oscillators by $i = 0, 1, \dots, N - 1$, the non-vanishing coupling variables are $K_{i,\mathrm{mod}(i \pm 1, N)}$. This labeling carries over analogously to the 2-dimensional case by adding a second index to the oscillators to arrange them "on a grid". From now on all subscripts are given modulo N, so oscillator i connects to oscillators $i \pm 1$ for the ring and oscillator (i, j) connects to $(i \pm 1, j)$ and $(i, j \pm 1)$ in the 2-dimensional case. Thus, $|S_i| = 2, 5$ and $\alpha = \frac{1}{3}, \frac{1}{5}$ for the respective topologies.

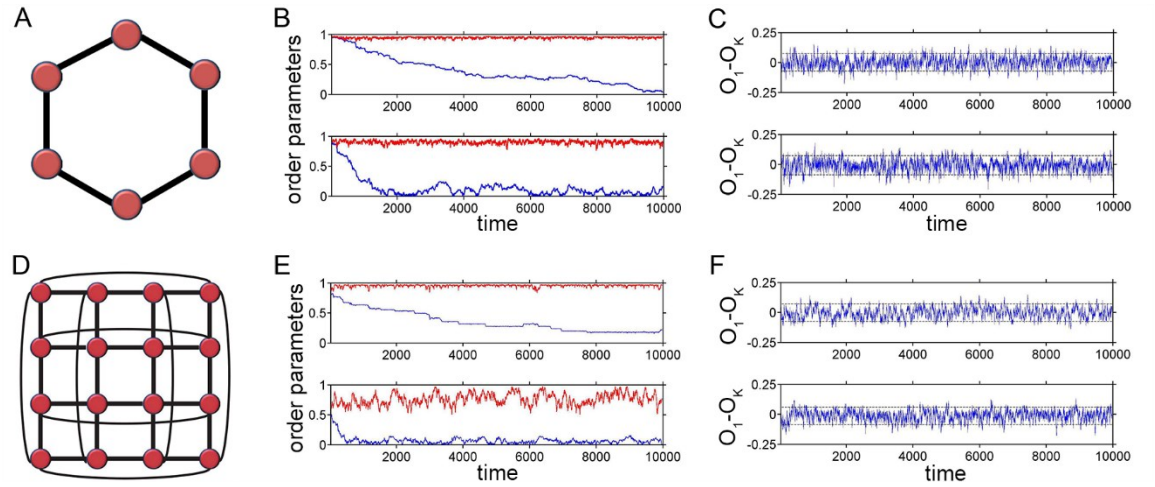

Figure 2.1. Topology and dynamics of oscillators with nearest neighbor coupling in one and two dimensions. A. Illustration of ring topology – nodes (circles) with connecting lines are coupled. **B.** Order parameters O_1 (blue) and O_2 (red). $\epsilon = 0.1$. **C.** $O_1 - O_K$ (blue). $\epsilon = 0.1$. The dashed lines encompass 90% of the data points on either side of 0. **D.** Illustration of 2-dimensional nearest neighbor topology with periodic boundary conditions. **E.** Order parameters O_1 (blue) and O_2 (red). $\epsilon = 0.05$. **F.** $O_1 - O_K$ (blue). $\epsilon = 0.05$. delineate 90% of the data points on either side of 0. All parameters are $N = 100$, $\sigma = 0.20, 0.25$ (top to bottom within panels); top row is for ring topology, bottom row is for 2-dimensional nearest neighbor topology; data starts at $t = 30$ to remove transients. The system can maintain overall synchrony (high O_2) even at small ϵ and large σ. O_1 tends to move down in a slow step-wise fashion for lower σ at a fixed ϵ, indicating the existence of relatively rare "discrete" events. At higher σ, O_1 drops very rapidly, so the oscillators quickly become arranged randomly in relative phase/anti-phase. O_K tracks well with O_1 – the difference looks randomly distributed around 0 and is within tight bounds. They are good mutual predictors even for relatively small ϵ and large σ.

Discussion

We explored the system behavior by solving the stochastic Equations (2.2), (2.3) iteratively using the Euler-Maruyama method with scaled time-step $\Delta t = 0.1$ for 10^5 steps, starting with initial conditions $K_{ij} = 0$ and $\phi_i = 0$ for all i, j. Random initial conditions did not change any qualitative results. For all simulations, $N = 100$ oscillators, approaching the thermodynamic limit.

We found that a set of appropriate oscillator order parameters for the ring topology are

$$O_m = \frac{1}{N}\left|\sum_{j=1}^{N} e^{im(\phi_j - \phi_{j-1})}\right|, \qquad m = 1,2, \tag{2.4}$$

where $i = \sqrt{(-1)}$. Longer-range order parameters were also tried but were not as successful at capturing the results. Values of O_2 close to unity imply that all oscillators are coherent, with $\phi_j - \phi_{j-1} = \{0, \pi\}$, while values of O_1 close to unity imply that most oscillators are in phase, with $\phi_j = \phi_{j-1} = 0$. For the coupling coefficients, an order parameter that captures the link to the dynamics of the system of oscillators (in particular, to O_1) is the simple expression

$$O_k = \frac{1}{N}\left|\sum_{i=0}^{N-1} K_{i,i-1}\right|. \tag{2.5}$$

Figure 2.1B shows three representative plots of the order parameters $O_{m=1,2}$ using $\epsilon = 0.1$ and $\sigma = 0.20, 0.25$ (top, bottom within the panel, respectively), starting at $t = 30$ to remove the transient from the deterministic initial conditions. As σ increases, O_2 remains near 1 but has a larger spread. Thus, the effect of noise on coupling is not to throw the system into disarray – there is still a sense of synchronization. However, as the noise σ increases, the coupling is more disordered and the oscillators can move from in-phase to anti-phase relations with their neighbors. Then, the oscillators switch between two possible discrete phases (a form of synchrony) but move between them incoherently. Of course, for very large σ (outside of the range shown), there is an uninteresting incoherent regime characterized by O_2 decreasing significantly from unity.

Figure 2.1C shows the difference $|O_1 - O_K|$ for the same two sets of parameters. These order parameters remain within a relatively tight band – the dashed lines delineate the range for 90% of the data points. The fact that the lines are highly symmetric about 0 and have magnitude near $\pm$0.1 even in the high σ case suggests that O_K tracks O_1 very closely.

The order parameters generalize directly to the 2-dimensional case by adding a second index and changing the constant in front to the total number of couplings (number of unique connected pairs). Figure 2.1E-F show $O_{m=1,2}$ and $O_1 - O_K$ in the 2-dimensional topology for $\sigma = 0.20, 0.25$ (top and bottom within the panels, respectively). A similar pattern as in the ring topology emerges, where oscillators have a high relatively high overall synchrony and tend towards arranging in random phase/anti-phase relations in time. Again, O_K follows O_1 closely, as seen by tight bounds of the dashed black lines representing the cutoff for 90% of the positive and negative points. In fact, even though ϵ is lower (here, $\epsilon = 0.05$, as compared to $\epsilon = 0.10$ for the ring) and the σ is the same, the lines form a tighter band in the 2-dimensional nearest neighbor topology. This opens up the intriguing possibility of reasoning about the order parameter of the coupling strength, which is in general hard to observe, based on the more easily observable synchronization of the oscillators, especially in more realistic scenarios where a 2-dimensional structure provides a better model.

To understand the temporal dynamics of synchronization, we examined the logarithm of the time required for O_1 to "crash" (defined as first achieving a threshold value 0.1 of the initial value $O_1(0) = 1$) as a function of σ for representative values of ϵ in the ring and 2-dimensional topologies (Figure 2.2A-B). All results were averaged over 10 runs, and dashed lines correspond to the best fit. A value was not plotted if O_1 did not reach 0.1 during the full simulation time. This suggests that for small σ oscillators switch between alignment and anti-alignment rarely, while for large σ switching is common. Guided by this intuition, we explored the possibility of distinct synchronization regimes in the system as a function of the noise strength σ and the rate of change of coupling ϵ.

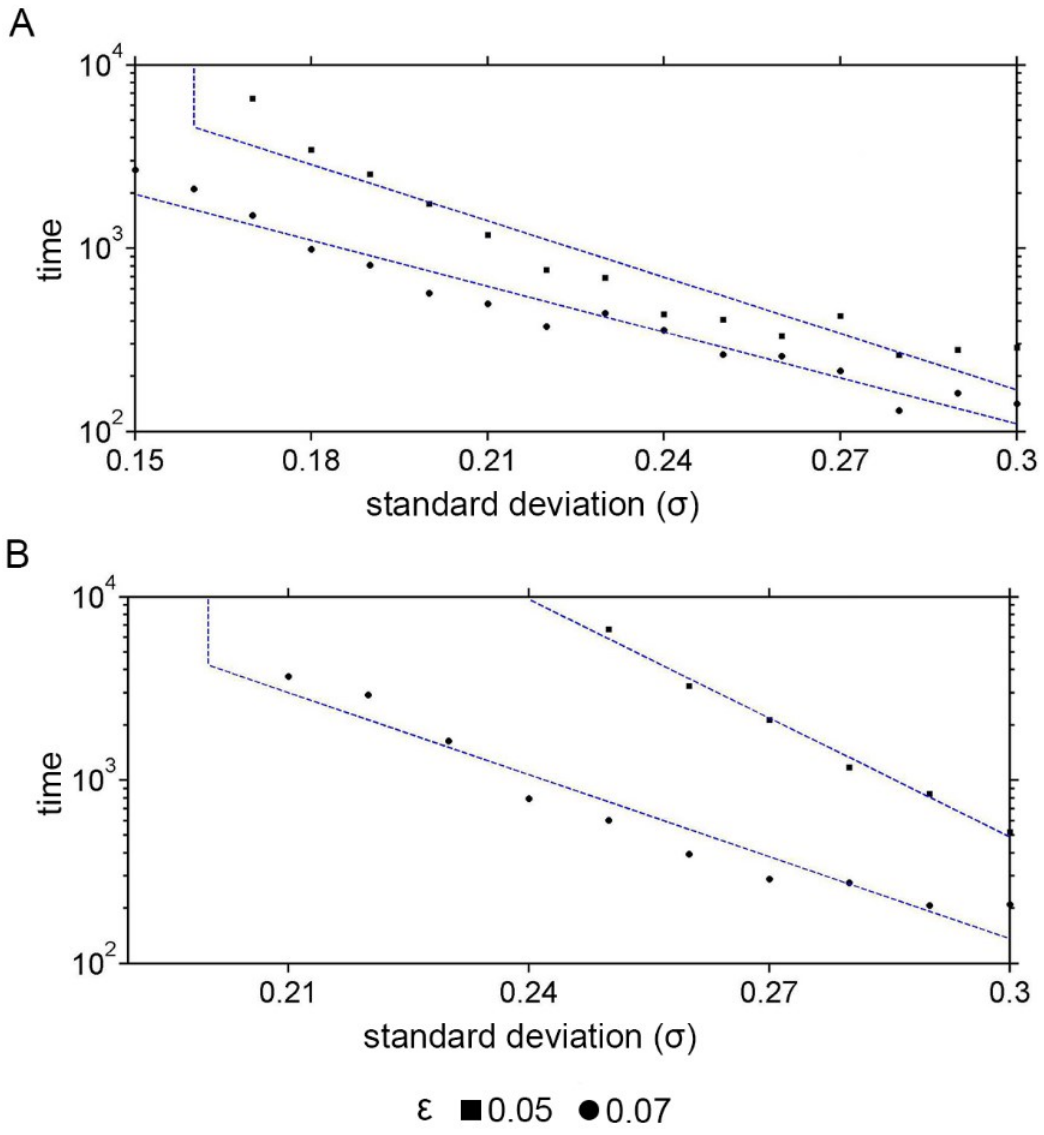

Figure 2.2. Time for desynchronization in 1-dimensional and 2-dimensional topologies as a function of noise strength for different values of the rate of change of coupling. A. Ring topology. **B.** 2-dimensional nearest neighbor topology. Time required for O_1 to reach 10% of its starting value of 1 as a function of σ for various values of ϵ (logarithmic scale). The dashed vertical lines represent the point before which O_1 did not reach 0.1 during the full simulation run-time. Non-vertical dashed lines are best fits. $N = 100, \epsilon = 0.05, 0.07$ (circle, square). The time to crash is close to exponential in σ, indicating the existence of distinct regimes. Topologies with more connections take longer to crash for the same parameter values.

Phase diagrams in the $\epsilon - \sigma$ plane indicate that the system can operate in two distinct regimes split approximately along a straight line in the ring (Figure 2.3A) and 2-dimensional topologies (Figure 2.3B). We define a flip to be an event where the difference of phases between

two neighbors changes from $\Delta\phi \pmod{2\pi} \in \left[0, \frac{\pi}{3}\right] \cup \left[\frac{5\pi}{3}, 2\pi\right]$ to $\Delta\phi \pmod{2\pi} \in \left[\frac{2\pi}{3}, \frac{4\pi}{3}\right]$ or vice versa. Contracting the bounds did not affect the results. Results were averaged over the number of pairs over 10 runs. In the "synchronous regime" (above the line in Figure 2.3), the oscillators achieve a steady state where after an initial fluctuation they choose to be in a phase relation of 0 or π with their neighbors, depending on the couplings, and freeze. However, for larger values of σ for a fixed ϵ, they dynamically switch between in-phase and anti-phase relations with their neighbors. We call this the "flipping regime" (below the line). Aside from a fast transient, even in the flipping regime there is no "incoherent" motion – a particular oscillator i simply changes from being closely synchronized with its neighbors to being anti-synchronized. Thus, the system maintains overall coherence while each oscillator can move between two states. For the ring topology, the black line is $\epsilon = 0.9\sigma - 0.085$, while in the 2-dimensional case the line is $\epsilon = 0.5\sigma - 0.05$. The region corresponding to completely asynchronous, noise-dominated motion is not shown, although Figure 2.1E shows that the system begins to approach that regime towards the bottom right of the phase diagram. It is interesting to note that the line in the 2-dimensional topology has a significantly smaller slope than in the 1-dimensional case. Indeed, in the limiting case of an all-to-all connection topology, no flips were observed on the time-scale of the simulation in this parameter regime.

The oscillator behavior can be intuitively seen as follows: since they are identical and have fast dynamics compared to the coupling, they quickly arrange themselves into a steady state. When the coupling between neighbors changes such that the anti-phase state becomes more stable than the in-phase state (or vice versa), the relevant oscillators quickly rearrange themselves to the new state, resulting in a high O_2. However, as the relative coupling strengths between an oscillator's neighbors shift more quickly, there are more flips between the two states the oscillator wants to adopt relative to its neighbors, leading to a lower O_1. If σ is too large, the effects of neighbors are

dominated by noise and the system is thrown into incoherence. Further, if there are too many neighbors, in order to have flipping occur one needs to have the majority of oscillators flip at once (since otherwise the other majority will tend to prevent flipping). Thus, topologies with many neighbors lack the flipping regime. In other words, the flipping regime occurs in a special intermediate region and can be encapsulated as follows: there need to be few enough neighbors that their individual effects matter and the coupling between oscillators has to have sufficient noise to promote flipping but not so much that oscillators consistently have only random interactions and fall completely out of coherence.

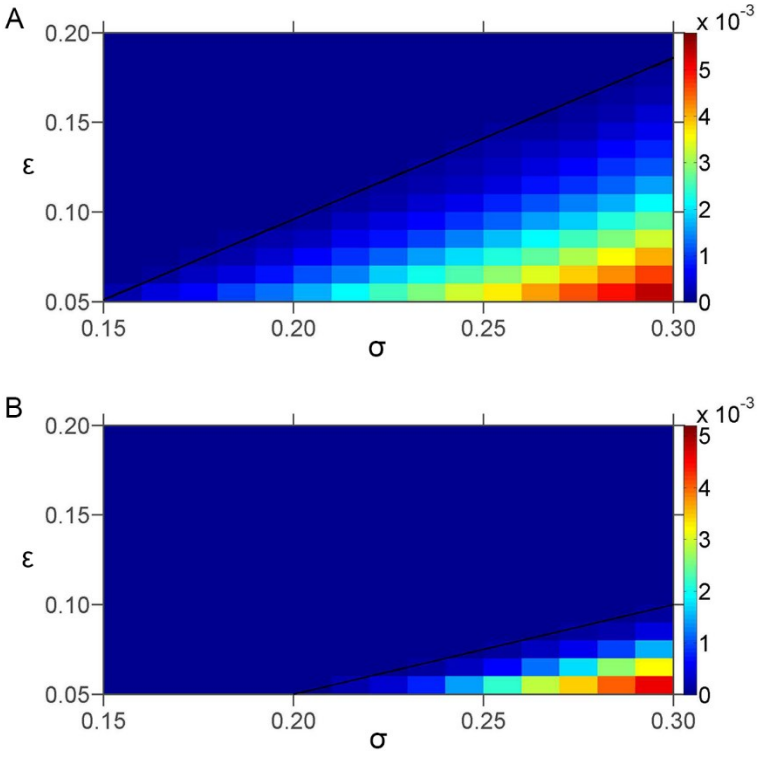

Figure 2.3. Phase diagrams for synchronization/desynchronization in 1-dimensional and 2-dimensional nearest neighbor networks as a function of the noise strength and rate of change of coupling. A. Ring topology. **B.** Nearest neighbor topology. In both panels, color corresponds to the frequency of switching – blue is low, red is high. Values are per oscillator per unit time. Black lines correspond to the approximate division between the synchronized regime (no switching) and the dynamic (switching) regime. $N = 100$. Two distinct regimes are visible: a synchronous regime above the line and a "flipping" regime below the line.

These observations naturally led to considering the case of topologies with structural randomness to better understand how spatial variation interacts with the temporally noisy dynamics, which is of paramount importance in neural systems (Mi et al., 2013). We used a modified preferential attachment model with a tuning parameter p: with probability p, a new node attaches randomly to one of the previous nodes, and otherwise it attaches to a previous node i with probability $\frac{k_i}{\sum_{j=1}^{i-1} k_j}$, where k_i denotes the degree of node i.

Then, we added a "backbone" of a ring lattice to allow comparison with the 2-dimensional lattice. We considered three cases: each new node makes $l = 1, 2, 3$ attachments to prior nodes. The case $l = 1$ corresponds to a mean degree of approximately 4, as in the lattice case. The case $l = 2$ corresponds to a minimum of 4 neighbors (up to removal of duplicate links when adding the ring lattice backbone, which was negligible) with a mean degree of approximately 6. The degree distribution when $l = 3$ has a mean of approximately 8.

The phase diagrams for flipping times in $p - \sigma$ space are shown in Figure 2.4, considering flips only between neighbors along the ring, with $\epsilon = 0.05$. While the results are seen to be independent of p, they are strongly dependent on l. For $l = 1$, the flipping rates are generally of the same order of magnitude as in the 2-dimensional lattice, while for $l = 2$ flipping begins at $\sigma = 0.27$ and is of significantly smaller magnitude compared to the same σ when $l = 1$. The flipping regime is almost extinguished when $l = 3$ – flipping rates are on the order of 10^{-4}. Thus, although spatial variability itself may not be usable as a valid method for the fine control of temporal dynamics, the synchronization of this system is highly robust to the underlying network topology and depends on the mean of the underlying degree distribution. This result is in line with the previous analysis where we found that the synchronization effects are only a function of local connectivity. Thus, even in the

case of high spatial variance where some nodes are highly connected (and hence are not expected to flip often), there are so few of these nodes compared to those of low degree (below the average, hence more likely to flip) that their effect on the macroscopic system behavior is largely negligible. However, increasing the minimum degree of all nodes makes all nodes less likely to flip, impacting the macroscopic behavior. This robustness result is interesting because it helps us understand the wide-spread appearance of synchronization in real-life networks, which can exhibit any number of interesting topologies that also have slowly varying coupling strengths.

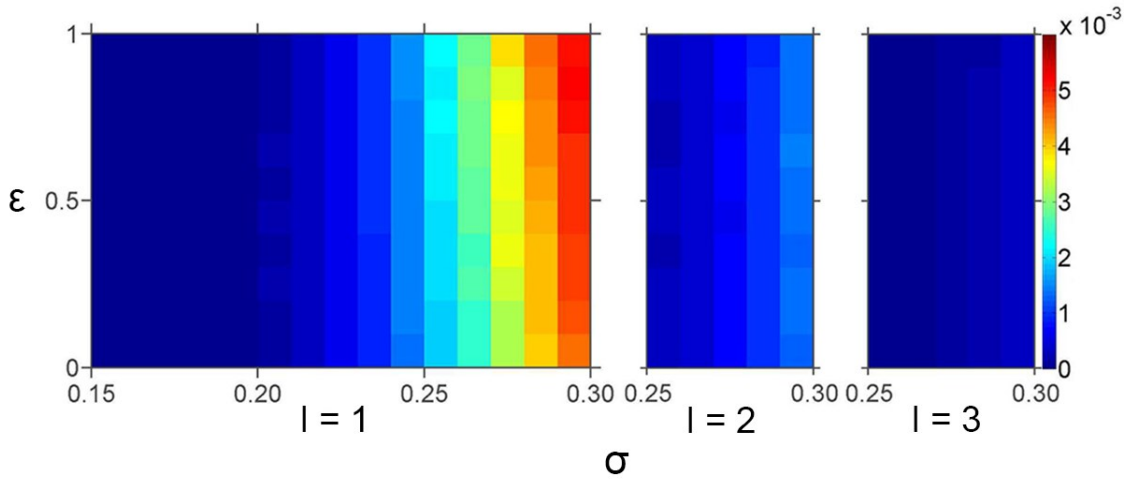

Figure 2.4. Phase diagram for synchronization/desynchronization in a random network with preferential attachment and different values of mean degree. Phase diagrams showing the number of times oscillators switch from an in-phase to an anti-phase state relative to their neighbors on the ring as a function of σ, p for three different values of l. Values are per oscillator per unit time. $N = 100, \epsilon = 0.05$. The amount of flipping is found to be independent of p but highly dependent on l. When $l = 1$ (same mean degree as the 2-dimensional lattice), the amount of flipping is approximately a factor of 2 less than in the 2-dimensional lattice. In the case where $l = 2$ (same minimum degree as the lattice, mean degree = 6), flipping is largely extinguished, and at $l = 3$ there is no consistent flipping even at the highest value of $\sigma = 0.30$.

For an analytical perspective on the system dynamics as a function of ϵ, we derived a useful approximation for $O_1(t)$ based on the results of the numerical experiment. Since the order parameter for the coupling O_K tracks O_1, we estimate $O_1(t) \sim O_K(t) = \frac{1}{N}\sum_{i,j} K_{ij}$, dropping

absolute value signs from the definitions since they were only introduced into the order parameters for convenience. Summing Equation (2.3) over all neighbors i, j yields

$$\sum \dot{K}_{ij} = \epsilon\left(-\sum K_{ij} + \sum \cos\left(\phi_i - \phi_j\right)\right) + \sum \sigma \zeta_{2;i,j} \tag{2.6}$$

Since $\langle \zeta_{2;i,j} \rangle = 0$, the last term vanishes in the large N limit. Approximating $\mathrm{sign}\left(\cos(\phi_i - \phi_j)\right)$ by $\mathrm{sign}\left(K_{ij}\right)$ (since synchronization between neighbors occurs on a faster time-scale than the change in coupling) yields

$$\frac{d}{dt}\sum K_{ij} = -\epsilon \sum K_{ij} + \epsilon \sum \mathrm{sign}(\mathrm{K_{ij}}) \tag{2.7}$$

In the no flipping regime, the two terms on the right hand side cancel as each K_{ij} settles into an average steady state of ± 1. In the flipping regime, flips occur relatively rarely so the last term on the right hand side can be approximated as $\epsilon\beta$, where β is approximately constant and bounded by the number of couplings (number of unique oscillator pairs). The solution to this equation is $\sum K_{ij} = Ce^{-\epsilon t} + \beta$, where C is determined by initial conditions. Indeed, the bottom graphs in Figure 2.1B, E show an approximately exponential decline in O_1 towards a state where there is overall synchrony in the system but each set of neighbors may be aligned or anti-aligned.

Conclusion

Our study of two-way oscillator-network interactions points to a number of directions for future work. One area to explore is applying the model to experimentally observed topologies. It would be promising to look at the correspondence between the order parameters for the oscillators and the coupling strength with a view to reasoning about coupling strength from direct observations of e.g. neural networks. A local tie between coupling and the phase/anti-phase relation in the ring and 2-dimensional nearest neighbor topology is given by $\mathrm{sign}(K_{i,i-1})$. In a typical run, this measure

correctly predicts the synchronization relation between two neighbors ($\mathrm{sign}(K_{i,i-1}) = \pm 1$ when $i, i-1$ are in in-phase and anti-phase states, respectively) approximately 90% of the time even for the extreme case $\epsilon = 0.05, \sigma = 0.30$, and even more accurately farther away from the bottom right corner of Figure 2.3A-B. Another promising line of research is to consider the proposed model in the context of chimera states (Abrams and Strogatz, 2004; Abrams et al., 2008; Martens et al., 2013). Experimentally, it may be possible to test the phase diagram predictions with an electromechanical or laser model using feedback as the analogy to a strengthening or weakening coupling.

Inspired by the idea of agent interactions changing a network structure, in the next chapter we turn to a model of an evolving, interdependent population. There, the network is directly defined by the (evolving) agent strategies themselves, and we will follow a similar methodology of characterizing different operating regimes. And, we will go further towards answering the question of control: how does a perturbation to the network in a given state affect the final outcome?

"He who thus considers things in their first growth and origin… will obtain the clearest view of them."
- Aristotle, *Politics*

3

Organization of an Evolving Interdependent Population

Introduction

Cooperative behavior, as exemplified by multicellular life, seems to have evolved at least 25 times independently – once for plants, once or twice for animals, once for brown algae, and possibly several times for fungi, slime molds and red algae (Grosberg and Strathmann, 2007). On shorter time scales, the social composition of eukaryotes such as *Saccharomyces cerevisiae*, and biofilm-forming bacteria such as *Pseudomonas aeruginosa* can dramatically change in a brief period (Diggle et al., 2007; Gore et al., 2009; Jiricny et al., 2010; Ratcliff et al., 2012). In a related context, tumor formation is a rare example of the transition, taking place in the reverse direction, from a multicellular to an essentially unicellular lifestyle. Interestingly, cancer cells end up cooperating by collectively secreting angiogenic factors, and it seems possible, at least in principle, that there may even be cheaters (i.e. those who do not secrete the growth factors) among this collection of cooperating cheaters (Axelrod et al., 2006; Nagy et al., 2007).

Without compromising the simplicity and tractability offered by traditional evolutionary game theory, we built an evolutionary model in which the fitness of an agent is determined, not by the outcome of a two-player two-state game, but instead a multi-player multi-state one. Thus, the focus is not the typical cooperator–defector ratio in the population, but rather the large-scale structure of all exchanges. In other words, we considered the *interdependence* between agents or groups of agents within a genetically heterogeneous population. We asked how independent agents become interdependent through the simple laws of evolution, whether positive selection is a necessary or a sufficient condition for the formation of interdependence, what kinds of interdependent structures are stable/unstable, and how these structures and processes depend on evolutionary parameters.

Accordingly, the present model offers a clear framework for classifying and categorizing different regimes of interdependence, as well as allowing for careful control of evolutionary parameters that may be influencing recent non-intuitive empirical outcomes (Kohler et al., 2010). We determined which kinds of external perturbations promote anti-sociality (e.g. in order to eradicate biofilms) and which other kinds can inhibit anti-sociality (e.g. as to suppress or reverse tumor growth) by simulating the introduction of selfish/altruistic strains into a population or the administration of anti-sociality/sociality promoting drugs. Finally, we evaluated the success rate of these evolutionary interventions as a function of the original population structure, drug dose, fraction of drug-resistant agents and reproduction speed of the target species.

Model

The multi-state multi-player game can be best visualized as a network of $N \gg 1$ agents connected by directional edges. An edge from A to B indicates that A contributes to the fitness of B at the cost of its own. Unlike the typical evolutionary game theoretic models where the state of a player i is binary (cooperator/cheater), here the player states ψ_i are characterized by high-

dimensional vectors, i.e. $\psi_i = \{x_1, x_2, \ldots, x_N\}$ with $x_j \in \{0, 1\}$ indicating whether i provides a fitness benefit to j. The evolutionary dynamics is governed by the following assumptions:

1. The fitness $\omega(x_i)$ of a node i is assumed to be a monotonically increasing function of received net benefit $x_i = bn_{i,in} - cn_{i,out}$, where $n_{i,in}$ and $n_{i,out}$ are the number of incoming and outgoing edges for node i, respectively. The parameter $\beta = \frac{b}{c}$ quantifies the benefit of an edge (to the receiver) relative to its cost (to the provider).
2. Every generation, the r most fit nodes produce offspring that replace the r least-fit nodes. Reproduction preserves all edge relationships of the parent, i.e. parents and offspring connect to the same agents.
3. There is a small mutation probability p per generation with which edges are added/removed randomly.

A schematic of the model is shown in Figure 3.1. The model has four parameters, kept constant throughout the course of evolution: population size N, mutation probability p, number selected for replacement r, and the relative benefit β. For every run, we tracked the total number of edges $E(t)$ as a function of generation number t. $E(t)$ is a measure of the interdependence of the population as well as the average fitness (the latter follows from $\langle\omega\rangle = \sum_i \frac{\omega_i}{N} = \frac{(\beta-1)E(t)}{N}$, which can be positive or negative depending on the value of β). In addition, we studied the community structures and genetic composition within the population, which are defined in terms of the connectivity matrix C of the network. Here, $C_{ij} = 1$ if j depends on i and $C_{ij} = 0$ otherwise, represented by black and white pixels, respectively, in array plots. Simulations were run for $N = 200$, with the relevant number of degrees of freedom being the number of edge slots $N^2 = 4 \times 10^4$. This ensured that the evolutionary transitions are not accidental fluctuations.

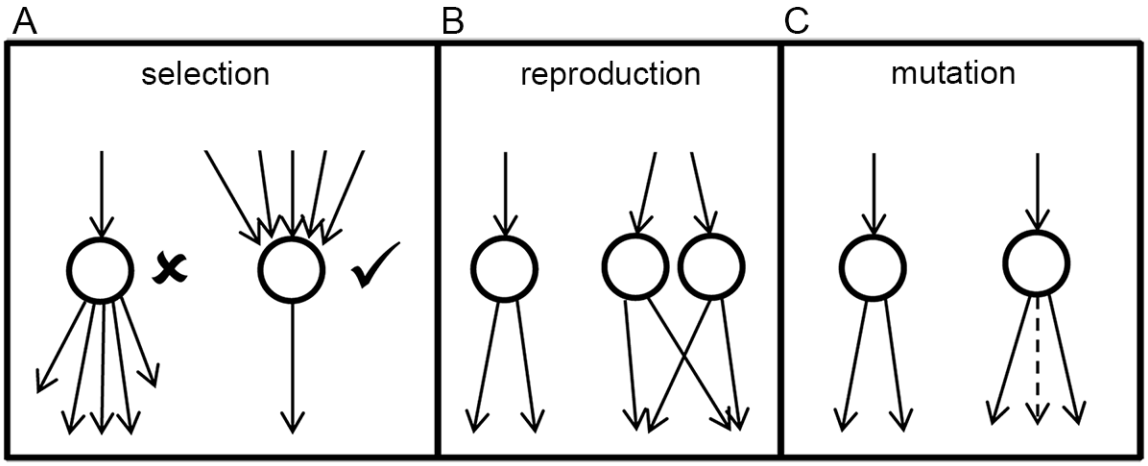

Figure 3.1. Schematic of the evolutionary dynamics model. A. Selection step. The fitness of a node $\omega(x_i)$ is a monotonically increasing function of edge influx $x_i = bn_{i,in} - cn_{i,out}$, where $n_{i,in}$ and $n_{i,out}$ are the number of in and out edges for node i. **B.** Reproduction step. The r fittest nodes replace the r least fit nodes. Reproduction preserves all in/out relationships of the parent. **C.** Mutation step. A small number of edges are randomly added/removed every generation, with probability $p \ll 1$.

Results

The evolutionary dynamics and final interdependence states depend on the values of β and relative selection pressure $\frac{r}{m}$, where $m = N^2p$ is the expected number of mutations per generation (which is equal to the number of mutants if $p \ll 1$). The parameter space can be divided into four regimes: neutral constructive ($r \ll m, \beta > 1$), selective constructive ($O[r] \sim O[m], \beta > 1$), neutral destructive ($r \ll m, \beta < 1$), and selective destructive ($O[r] \sim O[m], \beta < 1$). In all cases, the formation of an edge is beneficial for one of the nodes and deleterious for the other. As the value of β determines the change in average fitness per edge, $\frac{d\langle\omega\rangle}{dt} = \frac{c(\beta-1)}{N}$, one may intuitively expect $E(t)$ to decrease for $\beta < 1$ and increase for $\beta > 1$ as long as selection strength is finite. However, we will see that this expectation will not always be satisfied, particularly when $\beta > 1$.

First, consider the straightforward case of $\beta < 1$. Here the formation of an edge is more deleterious to its originator than beneficial to its target, and the fitness of the population changes

by $b - c < 0$ per edge. Thus, in the long run, only if the selection is weak ($r \ll m$) can such deleterious edges accumulate, and a random interdependence network forms with $E = \frac{N^2}{2}$. As expected, increasing $\frac{r}{m}$ causes the network to become sparse and fragmented, and all structure vanishes as $O[r] \sim O[m]$.

Now, let us consider the constructive regime $\beta > 1$, which produces distinct phases of complexity (Figure 3.2A-F). The long-term behavior of $E(t)$, which can be viewed as a proxy for average fitness as well as for interdependence and complexity, depends non-monotonically on selective pressure. For small values of $\frac{r}{m}$, the asymptotic value $E(t \to \infty)$ decreases with $\frac{r}{m}$ (Figure 3.2A-C). However, if selective pressure exceeds a critical point, there are sudden transitions between well-defined discrete levels (Figure 3.2D-F). As selective pressure is increased further, there are increasingly larger fluctuations around these levels.

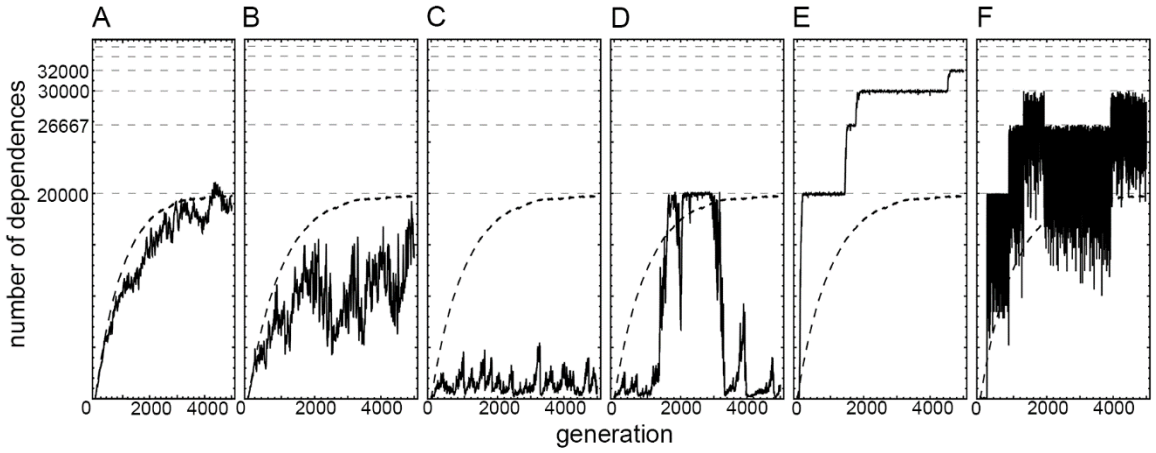

Figure 3.2. Regimes of constructive evolution for different values of selection strength. A. Number of edges $E(t)$ as a function of time t. Mutation probability is kept constant such that $N^2 p = 20$. The dashed straight lines indicate the stable number of edges corresponding to an integer number k of equal-sized bunches, $E = N^2(1 - 1/k)$. $N = 200$, $\beta = 1.01$. The dashed curved line is the outcome of the fully neutral simulation. In this panel, $r = 2$ (slowest reproduction). **B.** $r = 4$. **C.** $r = 9$. **D.** $r = 10$. **E.** $r = 15$. **F.** $r = 100$.

These asymptotic states can be described in more detail by connectivity matrices (Figure 3.3A) and phylogenetic trees (Figure 3.3B) for varying levels of $\frac{r}{m}$. The phylogenetic trees were obtained by quantifying the similarity distance $D_{ij} = \sum_k |C_{ik} - C_{jk}| + \sum_k |C_{ki} - C_{kj}|$ between all pairs of nodes i and j. In other words, if i and j receive from and provide to the same nodes, they are considered to be genetically related, consistent with the reproduction rule (compare Figure 3.1).

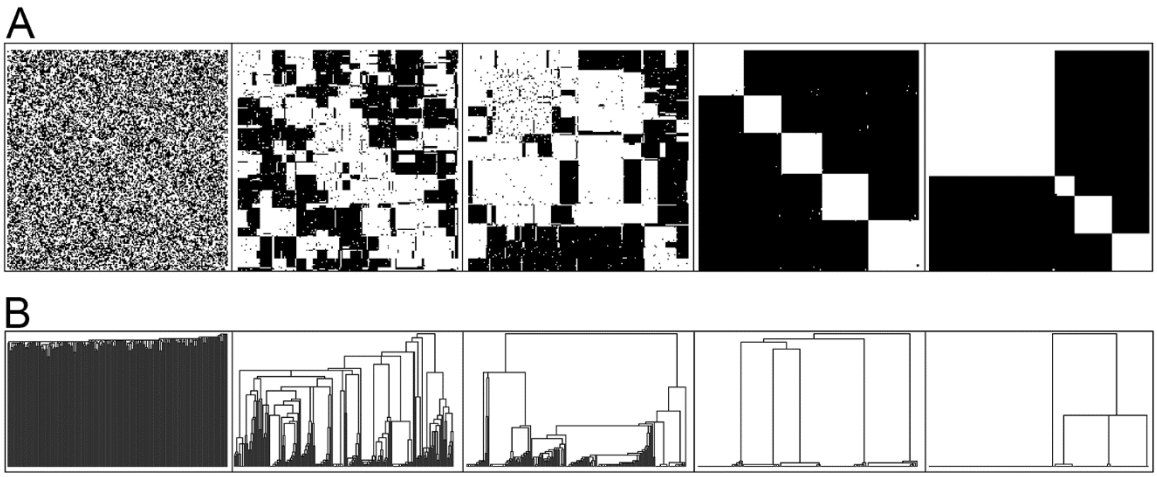

Figure 3.3. Interdependence and genetic composition in the constructive regime for different values of selection strength. A. Connectivity matrices for different selection strengths. Selection strength $r = 0, 2, 4, 15, 100$ (left to right). The connectivity matrix element C_{ij} is marked by black if individual i provides fitness to j, and left white if there is no exchange. For all panels $N = 200, N^2p = 20, \beta = 1.01$. There is a dramatic difference in the final organization of the population depending on r. There is an onset of "bunch" (anti-community) formation even in the weak selection limit (compare first two panels). The number and definition of bunches increases with higher selection strength (fourth panel). In the strong selection limit bunches compete with each other, leading to size heterogeneity (fifth panel). **B.** Phylogenetic trees corresponding to the connectivity matrices. Tree linkages are formed according to smallest inter-cluster dissimilarity defined by the L^1 norm.

While the destructive $\beta < 1$ regime produces either random or sparse networks, the constructive regime $\beta > 1$ can be summarized in terms of a sequence of complex phases governed by the value of $\frac{r}{m}$ (Figure 3.4): a transition from cooperation to competition between individuals (Figure 3.2A-C) is followed by unstable interactions between individuals and "bunches" (Figure 3.2D), followed by a transition from cooperation to competition between "bunches" (Figure 3.2E-

F). We define a *bunch* to be the opposite of a graph theoretical community; a group of nodes that form denser connections towards other groups, than they do within (compare the last two panels in Figure 3.3). Dense outwards connections and sparse intra-connections are the key qualitative characteristic of a highly specialized system. For example, nearly all energy spent by a heart muscle cell is directed at serving other tissues. The same holds true in a specialized society, e.g. a lawyer spends more of his time defending non-lawyers. The interdependence structures seen in the last two panels of Figure 3.3 conform to these biological and social examples of specialization.

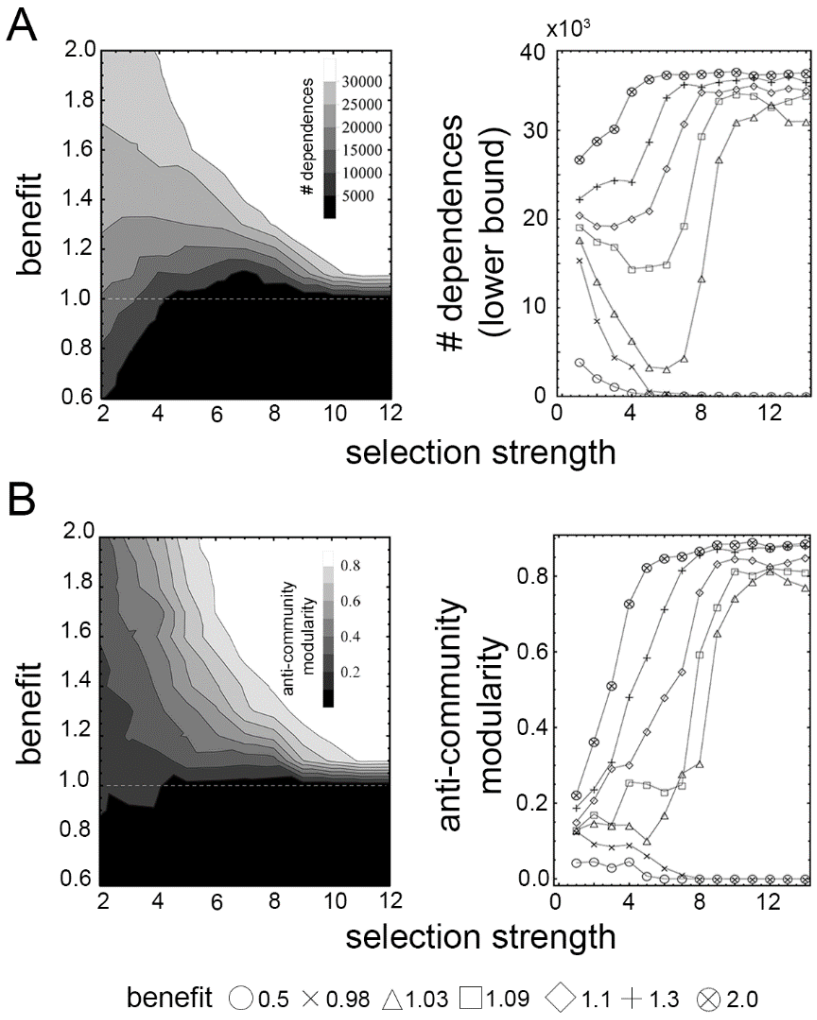

Figure 3.4. Phases of interdependence as a function of selection strength and benefit. A. Phase diagram showing the asymptotic dependence number E (left panel), and the same phase diagram traversed in the horizontal direction (right panel). Since some of the phases are highly

Figure 3.4 (Continued) dynamic, E is taken to be the minimum number of dependences in a large time-window in the long-time limit. **B.** Phase diagram showing bunch modularity (left panel), defined by exchanging 0 and 1 in the connectivity matrix and determining community modularity, and the same phase diagram traversed in the horizontal direction (right panel).

Next, we moved towards an understanding of the *control and manipulation* of the evolution of interdependence, which is now experimentally possible (albeit with mixed success) in biomedical and ecological settings. For example, the sociality of *P. aeruginosa* can be manipulated by drugs that suppress a microbe's production of a common good (iron scavenging siderophores). As the microbes that are resistant to the drug will altruistically continue to produce the expensive siderophores, they are taken over by their selfish counterparts affected by the drug (Diggle et al., 2007; Sandoz et al., 2007; Mellbye and Schuster, 2011). As a result, the iron-deficient population can be easily annihilated by the host's immune system (Boyle et al., 2013). Note that the evolutionary fate of the drug-resistant group would have been the opposite, had the drug been an antibiotic instead of a quorum blocker. On the other hand, there have also been experiments yielding the opposite outcome, where the drug aggravates the infection instead of impairing it, presumably by levelling the relative advantage of cheaters (Kohler et al., 2010). We used our model to quantify these mixed outcomes. Social evolution is complex, and its manipulation and control requires a detailed quantitative understanding of the evolutionary outcomes of varying initial states and system parameters.

To manipulate the sociality of a highly interdependent population, we started with an initial network that has a given community structure, selected a fraction η of the population and blocked a fraction γ of their outgoing connections of those that are selected. Following this perturbation, we tracked the evolution of the network to see if the perturbation would cause the entire population to lose all connections (which, for $\beta > 1$, amounts to minimal fitness). If $E(t)$ dropped to and remained at zero we counted it as a success, and we determined the fraction of successes for every

parameter value. If $\eta \ll 1$, the perturbation can be interpreted as an external introduction of a new strain/species, or a novel mutation which introduces a very small number of selfish individuals in the population. If $\eta \simeq 1$, the perturbation can be interpreted as a drug or event that affects nearly everyone, such as the quorum blocker discussed earlier. Accordingly, the quantity γ can be interpreted as the *dose* of the drug, or the degree of "selfishness" of the newly introduced species/strain. Figure 3.5A-B show the dependence of success as a function of η (empty versus closed plot markers correspond to $\eta = 2\%, 98\%$), initial population structure k (quantifying the number of bunches), and γ in fast- and slow-reproducing species, respectively.

It is important to distinguish between two very different mechanisms that can bring a population back to its pre-perturbed state. The first is determined by the time required for interdependence to evolve anew from $E = 0$. The original factors causing the establishment of cooperation in the first place is present regardless of the perturbation, and the effect of even the strongest drug ($\eta = 1, \gamma = 1$) is to simply reset the evolutionary clock. The second mechanism is evolution through the repopulation of the drug-resistant fraction, which happens much faster, on reproductive time scales. To clearly distinguish between these two mechanisms, we set $p = 0$ in Figure 3.5; using a non-zero p scales down all the success rates but does not otherwise change the qualitative dependence on η, γ, or k.

We observed a number of interesting features in the response of the population to external perturbations. For $\eta \simeq 1$, there is a non-monotonic dependence of success rate to β for populations with few bunches: for slow reproducing populations a moderate dose works as well as, or better than, a strong one. For larger numbers of bunches, and faster reproduction rates the non-monotonicity vanishes: the stronger the dose, the better the outcome. A second remarkable outcome was the degree to which a few individuals can make a difference: targeting $\eta = 2\%$ of the

population is as effective as targeting $\eta = 98\%$ of the population provided the drug has a high enough dose. This is because few selfish individuals, as is the case in tumors or invasive species, can devastate an entire population. Finally, we observed a very strong dependence of the success rate on the initial community structure. With increasing k and r, this difference vanishes.

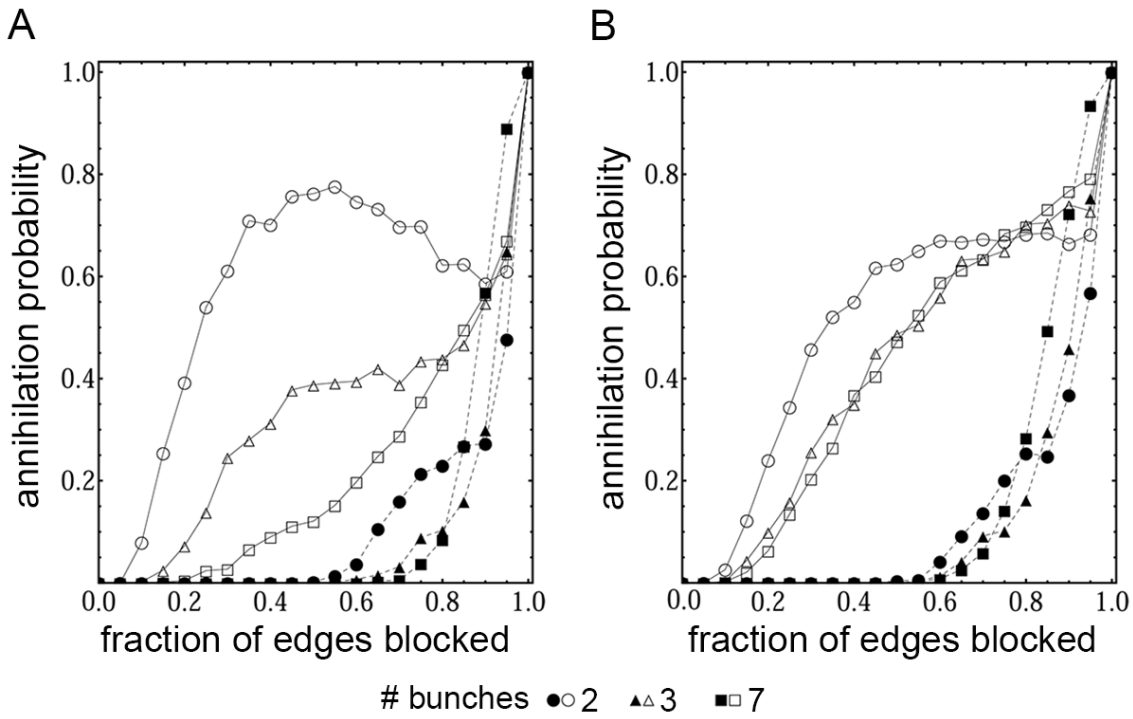

Figure 3.5. Manipulation and control of social evolution. A. Slow reproduction case ($r = 15$). For both panels, $\eta = 2\%$ (empty markers) and $\eta = 98\%$ (filled markers) of the nodes in the population ($N = 200$) are marked as susceptible to perturbation. The horizontal axis shows the proportion γ of the out edges that are blocked in the susceptible population. The vertical axis shows the probability that the initially cooperating population collapses into a weak, non-cooperating one (i.e. to $E = 0$). Here, $p = 0$ to distinguish between the spread of the resistant subpopulation η and novel mutations that occur after the perturbation (increasing p does not change the qualitative behavior). The dashed curve shows that even a small number of anti-social individuals are damaging to the whole population and suggests treatments where individual cells are targeted. The strong dependence of annihilation probability on k for small k might explain why quorum blockers are effective against some biofilms forming bacteria but not others. Interestingly, for $k = 2$, $P(\gamma = 0.5)$ is higher than $P(\gamma = 0.95)$, suggesting that for small number of cooperative bunches smaller perturbations may be more effective than larger ones. The effect of the drug strongly depends on the community structure for slow reproducing species. **B.** Fast reproduction case ($r = 60$). Here, the effect of the drug is less dependent on the number of bunches in the starting population.

Discussion

The phase diagram for the evolution of interdependence is shown in Figure 3.4 and compactly summarizes the results. Figure 3.4A shows the number of dependences, while Figure 3.4B quantifies their structure through "bunch modularity". We define the latter by exchanging $1 \leftrightarrow 0$ in the connectivity matrix and determining the community modularity. In the following, the regimes of interdependence are labeled with the letters A-D: regimes A and B refer to cooperation and competition between individuals, while regimes C and D refer to cooperation and competition between bunches, respectively.

Cooperation between individuals

In the neutral regime ($r \ll m$), additions and deletions of edges are equally likely. Thus, if the population starts fully independent, $E(t)$ increases until the network is fully randomized with $\frac{N^2}{2}$. This increase is statistically irreversible and is analogous to the scenario described in (Stoltzfus, 1999; Gray et al., 2010; Vural et al., 2014). In the neutral regime, individuals have high fitness due to the benefits of indirect reciprocity; interdependence emerges not due to higher fitness but simply due to higher likelihood. Figure 3.1A shows the dynamics and final outcome of nearly neutral evolution. Despite following a similar trajectory to fully neutral evolution ($E(t) \sim N^2(1 - e^{-2pNt})$, indicated by the dashed curve in Figure 3.2A-F), the connectivity matrix shows the onset of community formation; the interdependence structure and genetic composition of the population is far from random (compare Figure 3.3A-B).

As the selection strength is increased ($r < m$ but not $r \ll m$), the fluctuations in $E(t)$ are amplified. This is caused by the random formation of nodes for which the number of in-edges are different than out-edges. However, the system is self-stabilizing (Figure 3.2A-B); e.g. when the fit defectors reproduce, they typically replace their unfit providers, which in turn reduces their own

fitness. Consequently, they are taken over by the fair and fit nodes that dominate the $r \ll m$ population. In Figure 3.3A-B (third subpanels), two large reciprocating groups can be distinguished. They are taken advantage by smaller-scale opportunistic sub-populations. It is also possible to see a smaller sub-cooperative group sustaining itself within a larger cooperative group.

Competition between individuals

As r approaches $r_c \sim \frac{m}{2}$ from below, the fluctuations in $E(t)$ become comparable to $E(t)$ itself. Here the selective competition is just high enough to allow for small cooperative communities to form and grow at a rate much higher than random chance, but also high enough for cheaters to spread over their providers in one step, beyond recovery (Figure 3.1C). Although regimes A and B have similar destabilizing factors, their re-stabilization is very different. The drops in $E(t)$ in regime A can recover through *re-population*, over time-scales approximately $\frac{1}{r}$. By contrast, regime B exhibits system-size losses from which the only way to recover is *re-mutation*, over longer time scales determined by approximately $\frac{1}{m^2}$, as the smallest cooperative group requires two mutations.

Comparing A to B reveals that higher selection strength in this case leads to lower fitness; had one mixed the stronger selected population B with the weaker selected one A, the former would be driven to extinction. The behavior of B is similar to that expected in a classical prisoner's dilemma, which emerges from the model as a special case – survival of the fittest produces the globally least-fit outcome.

Formation of specialized bunches

As r is increased above the critical point $r_c \sim \frac{m}{2}$ higher-level structures start to form. While the connectivity matrix C is sparse and random for r just below r_c, metastable bunches begin to

form at $r > r_c$, the number and stability of which increases with r. The sudden jump in the edge number in Figure 3.2D-E is analogous to that found in (Jain and Krishna, 2001).

For a very large window of selection strength $O[m] \sim r < O[N]$, the system can only maintain certain discrete values of E. These are the stable configurations corresponding to an integer number of equal-sized bunches (k) given by the relation $Ek = N^2\left(1 - \frac{1}{k}\right), k = 0, 1, \ldots, k_{max}$ (Figure 3.2D). The maximum number of bunches k_{max} is determined by the mutation rate, $k_{max} \sim \frac{N}{m}$ (i.e. so that in steady state there is one mutation per bunch per time step). However, we have observed k transiently increasing to 50% higher than this value. Note that the degree of interdependence (and hence the average fitness) in the strong selection limit well exceeds that in the weak selection limit.

It is interesting that in the limit $r, m \sim 1$, structures more complex than bunches can form. These include hierarchies (smaller bunches within a bunch), cycles (3 or more groups providing to one other), and hierarchies of cycles (cycles within a cycle). In this limit, the dynamics of $E(t)$ still exhibit discrete steps similar to Figure 3.2E, but with more possible metastable plateaus corresponding to unequal-sized matrix blocks.

Competition between bunches

With increasing r the fluctuation in the number of edges around the stable k starts increasing, and there is destructive competition similar to that near the phase boundary of B; however, now the competition is between the bunches rather than the individuals, which creates significant size differences between them. These fluctuations can lead to one bunch replacing another, causing E_k to make large transitions between different values of k. Despite the apparent noise (Figure 3.2F) the dependence structure remains in a highly ordered state with high reciprocity

(Figure 3.3A-B, fifth panel). As r is increased further, the competition between bunches causes fluctuations comparable to the size of bunches, i.e. a small bunch can increase in size by spreading over others until another metastable structure re-evolves.

Conclusion

We constructed a simple model that allowed us to study the coevolution of self-replicating interdependent structures, and reported multiple evolutionary transitions as β and r were varied. This model is quite general and has few assumptions – the fitness function is only assumed to be an arbitrary increasing function of x and there are only two relevant parameters governing the dynamics (selection strength $\frac{r}{m}$ and relative benefit $\frac{b}{c}$). The population size N does not make a qualitative difference as long as both $Np, r \ll N$. Furthermore, the value β does not make a qualitative difference apart from whether it is larger or smaller than unity, and no quantitative difference if $|1-\beta| < \frac{1}{N}$. Unlike the typical simplified models of evolutionary game theory, we did not assume that an individual's behavior is the same towards all others (although some individuals can end up in a state where they give to all and receive from all). In this respect, the states allowed in this work are a generalization of the two-state models common in the literature. Thus, we hope that this model can serve as a guiding framework for understanding the emergence of sociality.

Even in this simple case, we observed a number of surprising phenomena. First, even the weakest selection strengths ($m \gg r$) can produce interdependence structures that are far from random. Thus, assumptions regarding "random interdependence" invoked by neutral evolutionary arguments may be too strong (Stoltzfus, 1999; Gray et al., 2010; Vural et al., 2014). Second, we observed the natural emergence of specialized bunches and multi-scale structures from the simple

laws of evolutionary dynamics. As we probed the response of the system to various selection strengths, we saw regimes of random interdependence, competition between nodes, cooperation between nodes (bunches), and competition between bunches.

Thirdly, we found that the regime $\beta > 1, r > 0$ does not itself guarantee complex interdependence. There exists a "dead zone" within the constructive regime (Figure 3.2C) due to competition between agents. This non-monotonic dependence of E on selection strength can have important implications in medicine. For example, biofilm populations may be induced into a less virulent non-cooperative state by decreasing the selective pressure, so that a cooperative film behaving as Figure 3.2E evolves into an intermediate non-cooperative state behaving as Figure 3.2C. This may be experimentally verified in *P. aeruginosa* by increasing the available iron while keeping their population constant by limiting their carbon source.

Another interesting result was the observation of a non-monotonic connection between the dose of anti-social drugs and the successful annihilation of cooperativeness. Indeed, the model exhibits a "contagion" effect which allows the manipulation of a few individuals to have population-wide effects. It has been noted that introducing several selfish mutants (or using an anti-social drug effective on a few individuals) may be far more effective than manipulating an entire population (Kohler et al., 2010) and is consistent with experimental observations (Diggle et al., 2007; Sandoz et al., 2007; Mellbye and Schuster, 2011).

Finally, when mutation rate was set to zero, we observed that the behavior of $E(t)$ resembles that of classical population dynamics. The dynamics between providers and receivers becomes qualitatively similar to that between the predators and prey of a system like Lotka-Volterra. If the initial condition is a randomized connectivity matrix and $p = 0$, $E(t)$ commonly tends to a fixed value and oscillates around it.

Owing to its generality and applicability, this model has room for many natural extensions. For example, the distribution of parameters and fitness functions in a more realistic model could include spatial, temporal, and individual heterogeneity. For instance, if we conceptualize the fitness as resources such as food, a more appropriate fitness model could be a concave function rather than a linear function to account for saturation. Further, the quantities p, b, and c can be dynamic as they are themselves, to an extent, subject to evolutionary forces. This can lead to an interesting set of potential future studies exploring connections between interdependence and evolvability/efficiency. Another factor not taken into account here is the possibility of the change in population size due to statistical fluctuations (e.g. due to a time-dependent energy input, or infection/predation). Such natural extensions could be appropriate to address systems in ecology, structured biological population, and provide insight into complicated social trends.

Moving from large evolutionary networks with multiple simple agents to a smaller system with more complicated units, in the next chapter we focus on the adaptation of the fruit fly locomotor system to injury. In particular, we combine experimental results and modeling to determine that proprioception (that is, feedback to the legs from the environment with which they interact) plays an important part in the ability to tune agent (leg) dynamics in order to recover from a turning bias observed immediately after leg amputation.

"Natural forces within us are the true healers of disease."
- Hippocrates

4

Recovery of Locomotion After Injury in *Drosophila*

Introduction

Given the varied environment in which organisms move, injury to locomotor systems is very common in nature. Therefore, it is not surprising that animals will often prioritize leg safety in locomotor strategies (Birn-Jeffery et al., 2014). However, damage can be unavoidable; if locomotor systems were not robust to damage, or were incapable of plasticity, limb injury would pose an insurmountable challenge to survival. In humans, a number of studies have shown that damage to the control mechanism (e.g. spinal cord injury) can be overcome to an extent by training using manually assisted signals to the limbs which reorganize the spinal network and allow it to adapt (Harkema, 2001; Dietz et al., 2009). Plasticity leading to locomotor recovery after spinal cord injury is seen in animal models such as rats as well (Ballerman and Fouad, 2006). After direct injury or amputation of a limb itself, animals (including mammals) can recover mobility over time (Kirpensteijn et al., 1999) – indeed, three-legged dogs and cats walking and even running are familiar images. In humans, a number of medical interventions such as prosthetic limbs after amputation or reconstructive surgery (Bosse et al., 2002) can help patients recover mobility. Thus, the locomotor

system is remarkably robust. Moreover, the idea of plasticity after limb injury is not limited to the animal world: even in engineered systems such as legged robots, instilling the ability to recover locomotion after injury is an active research topic (Christensen et al., 2013; Cully et al., 2015).

We ask the questions of (a) whether and (b) how the recovery of locomotion takes place after a significant biomechanical injury (leg amputation) in fruit flies. While many walking parameters have been characterized for freely walking *Drosophila* (Strauss and Heisenberg, 1990; Mendes et al., 2013; Berman et al., 2014), fewer studies have considered recovery of walking after injury. A notable exception is (Wosnitza et al., 2013) – even immediately after amputation of a fly's hind leg, these authors observed several important changes in an amputated fly's behavior that allowed it to continue walking but at a slower speed and with a shift in leg stepping patterns (gait coordination). Whereas here we examine both the immediate effects of injury as well as recovery in locomotion over time, Wosnitza et al. focused on behavior immediately after hind leg amputation. Intriguingly, a study where fly walking was impeded by adding weights to the body (Mendes et al., 2014) found evidence for adaptation of step parameters over time to maintain coordinated walking, as well as increased sensitivity to load and other locomotor defects in proprioceptive mutants. These observations beg for the exploration of recovery after amputation over longer times.

Therefore, we examined the immediate and days-long recovery of walking behavior after leg amputation in *Drosophila melanogaster*. By using video recording before and after injury, we show that amputation impairs exploratory locomotion, i.e. the paths followed in an open arena. Specifically, amputation of the right foreleg induces a counter-clockwise bias to exploratory locomotion. Interestingly, unbiased locomotion recovers well over time in wild type flies, but this recovery is significantly hampered in proprioceptive mutants. To understand how this might happen, we start by considering individual leg motion (gaits), quantifying them from high speed video of walking flies

before, immediately after, and for several days after amputation. Typically, gaits in hexapods are categorized into three distinct patterns, "tripod", "tetrapod", and "wave" or "non-canonical" (Hughes, 1952). The tripod gait is considered to be the alternating movement of two groups of three legs, with the legs in each group simultaneous taking off. These groups are traditionally defined as legs (135)(246) – with groups separated by parentheses and leg numbering as in Figure 4.1A.

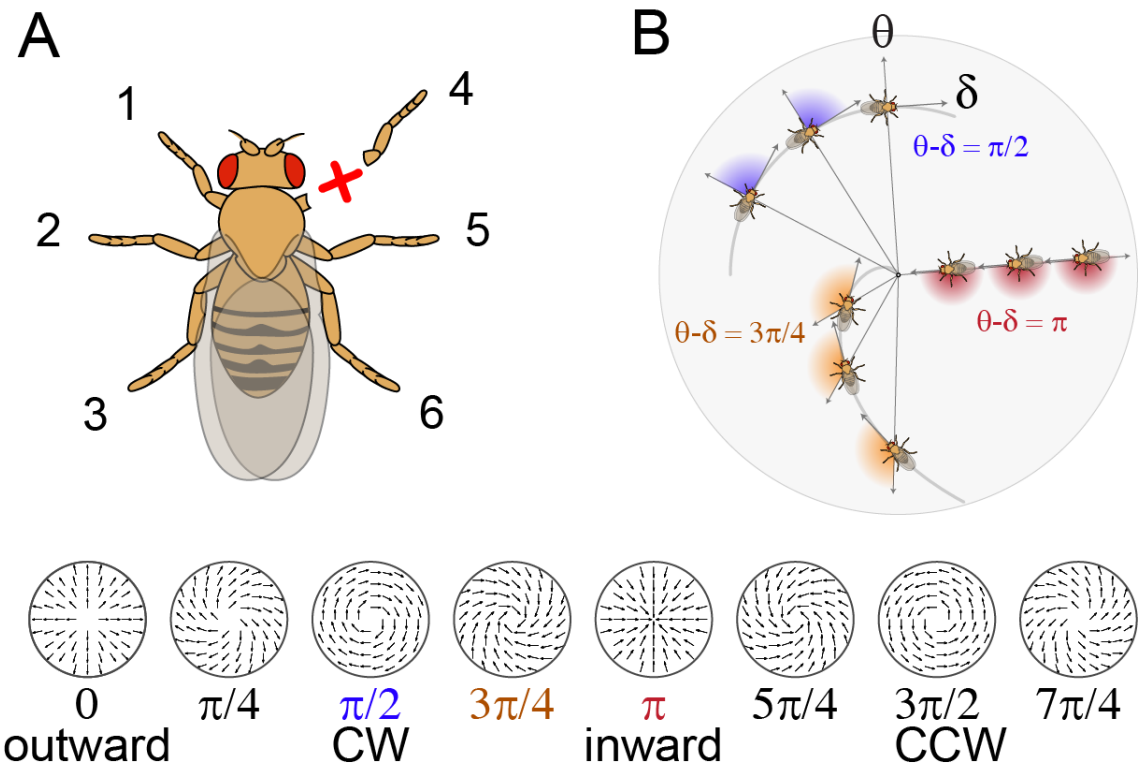

Figure 4.1. Amputation protocol and schematic of walking bias parameters. A. Amputation protocol: the right foreleg is removed between mid-femur and the femur-tibia joint. The given leg numbering is used throughout the paper. **B.** Schematic of analysis parameters used in constructing histograms of walking bias and calculating μ. The value of $\theta - \delta$ characterizes the degree of clockwise/counter-clockwise behavior.

The tetrapod gait consists of three groups of two moving legs each, with the legs again taking off simultaneously. Here, the groups can be arranged as either (15)(26)(34) or (24)(35)(16), due to a left-right symmetry. In a traditional wave gait, the legs proceed forward along a side before switching to the next side, as in (3)(2)(1)(6)(5)(4), though a number of sources such as (Kain et al.,

2013; Mendes et al., 2013) label gaits only as "tripod," "tetrapod" or "non-canonical". In reality, strict boundaries between gaits are not always well-defined.

We observed all of these gaits in our video analysis and found that injury resulted in permanent changes to their relative frequency during walking. This presents a puzzle: if injury permanently alters gaits, what mechanism explains the recovery in exploratory locomotion turning bias? Using a neuromechanical model strongly informed by experimental data, we show that amputated flies may redistribute the forces applied by the legs to enable the observed recovery in the absence of gait recovery, suggesting that a consequence of proprioceptive defects is the inability to precisely control leg forces. Altogether, combining behavioral observations and gait analysis of normal and proprioception deficient flies with a physical model provides us with a mechanistic description of recovery of locomotion after injury in *Drosophila*.

Materials and methods

Fly strains and care

Flies were housed on modified Caltech medium in temperature controlled incubators on a 12-h/12-h light/dark cycle. Flies mutant for *nanchung* and *inactive* were procured from the Bloomington Drosophila Stock Center (nan^{36a} BDSC #24902 and iav^{3621} BDSC #24768; Bloomington, IN, USA). Canton-S was our wild type strain. Flies were 4-8 days post-eclosion when experiments began. All experimental flies were female.

Centroid tracking in open arenas

Three-by-three arrays of 5.08 cm (2 inch) diameter arenas were fabricated from clear acrylic cut with a laser engraver. 10 cm high walls between neighboring arenas were frosted with a random orbital sander to prevent flies from viewing each other. Four-day-old wing-clipped flies were placed

into the arenas and allowed to walk freely for 2 hours. Arenas were uniformly illuminated from below by an array of LEDs covered by a diffuser fabricated from two sheets of 0.64 cm (1/4 inch) thick clear acrylic frosted on both sides by sanding. Arenas were imaged from above by 2MP digital cameras and the X-Y position of individual flies' centroids were identified and tracked by custom software written in LabView. The next day, the right foreleg was amputated within the femur and flies were tested again 1 hour, 24 hours, 48 hours and 72 hours post-injury. More proximal amputation resulted in higher mortality. More distal amputation risked leaving the animal with enough leg to support itself. The direction of motion was inferred as the angle between centroids of successive frames. In order to characterize turning bias, we considered the tangential component of the velocity ($\theta - \delta$) relative to the center of the arena (Figure 4.1B). To exclude edge artifacts, data collected within 80% of the radius of the arena was analyzed.

Gait experiments

Single, lidded, circular arenas were fabricated from acrylic. Individual four-day-old females of wild type (N = 56), *inactive* (N = 17), and *nanchung* (N = 15) were placed inside and the camera was refocused on a region roughly 2 cm x 3 cm. Wings were not clipped for gait experiments. Arenas were illuminated as in centroid tracking experiments and video was collected at 60 Hz using FlyCap software. The data collected from each trial consisted of two videos per fly, taken when the fly was performing a quick, straight run segment (subjectively assessed during data collection). In post-processing, the faster and straighter of the two videos was chosen for analysis. Both videos were rejected if the straight run segment had fewer than 3 full strides uninterrupted by pauses or large angular reorientations. Several animals perished over the course of the experiment. The five time-points had N = 56, 52, 52, 50, 51 (wild type), N = 17, 17, 16, 16, 16 (*inactive*), and N = 14, 13, 15, 15, 14 (*nanchung*). Before each assay, flies were anesthetized with CO2 and then allowed at least 45

minutes of recovery. On the first experimental day, each fly performed two assays. First, a pre-amputation assay was performed. Flies were subsequently anesthetized and amputated. On subsequent days (day 1, day 2, and day 3 post-amputation) each fly performed one assay.

Statistics

For between-group (wild type to mutant) parameter comparisons, P-values at corresponding time-points were found by unpaired t-tests with Welch's correction to the degrees of freedom. Change within-groups was assessed by regression on time post-amputation using all available data for each time-point, with P-values corresponding to an F-test against the null hypothesis that the slope is not significantly different from zero.

Gait video analysis

Video analysis was performed in several steps (schematized in Figure 4.2). Movies were first temporally cropped to encompass the full straight run and exclude all other frames. The cropped movie was then run through a semi-automated tracking algorithm to determine the fly centroid and the endpoints of the legs (Figure 4.3). We then reviewed every frame and either accepted the automatic recommendation or hand-corrected the leg endpoints. Then, an algorithm automatically sorted legs by calculating the angle between the centroid to head vector and the centroid to leg vector.

Finally, we binarized leg motions into "swing" (off the ground) and "stance" (on the ground) for determining gaits. To choose the motion threshold, we noted that apparent motion in the end position of a leg has two components: true leg motion and experimenter/measurement error when clicking on leg endpoints. We used a Gaussian mixture model to decompose the observed distribution of leg motion into these components and chose a threshold of 8 pixels/frame (Figure

4.4). We also excluded frames that indicate four legs moving (<1%) from analysis.

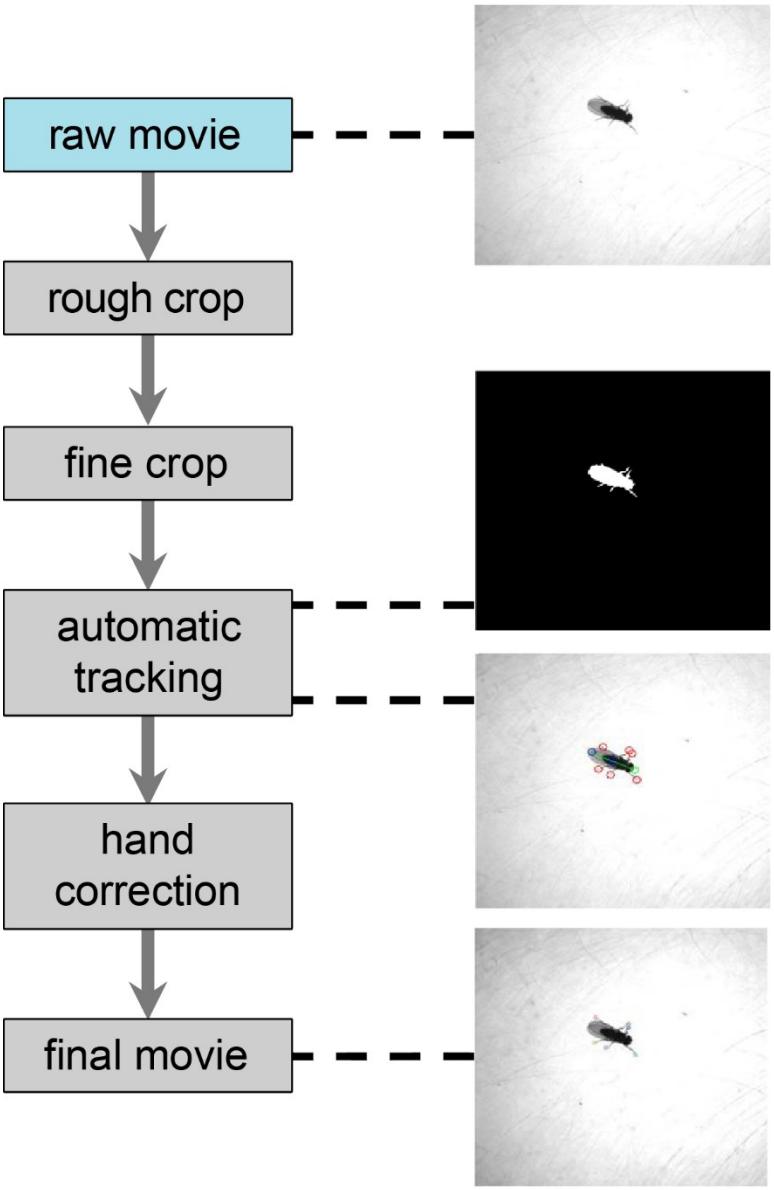

Figure 4.2. Video analysis process. Starting with an 8-bit grayscale movie, we perform an automatic rough temporal crop to remove the frames before the fly appears. Then, a fine-grained crop is performed by manually looking at still frames and determining the first and last frames that constitute the run. Then, we run the movie through an automatic tracking algorithm (Figure 4.3) and perform hand verification/correction on each frame. The final result is a movie with all leg positions tracked.

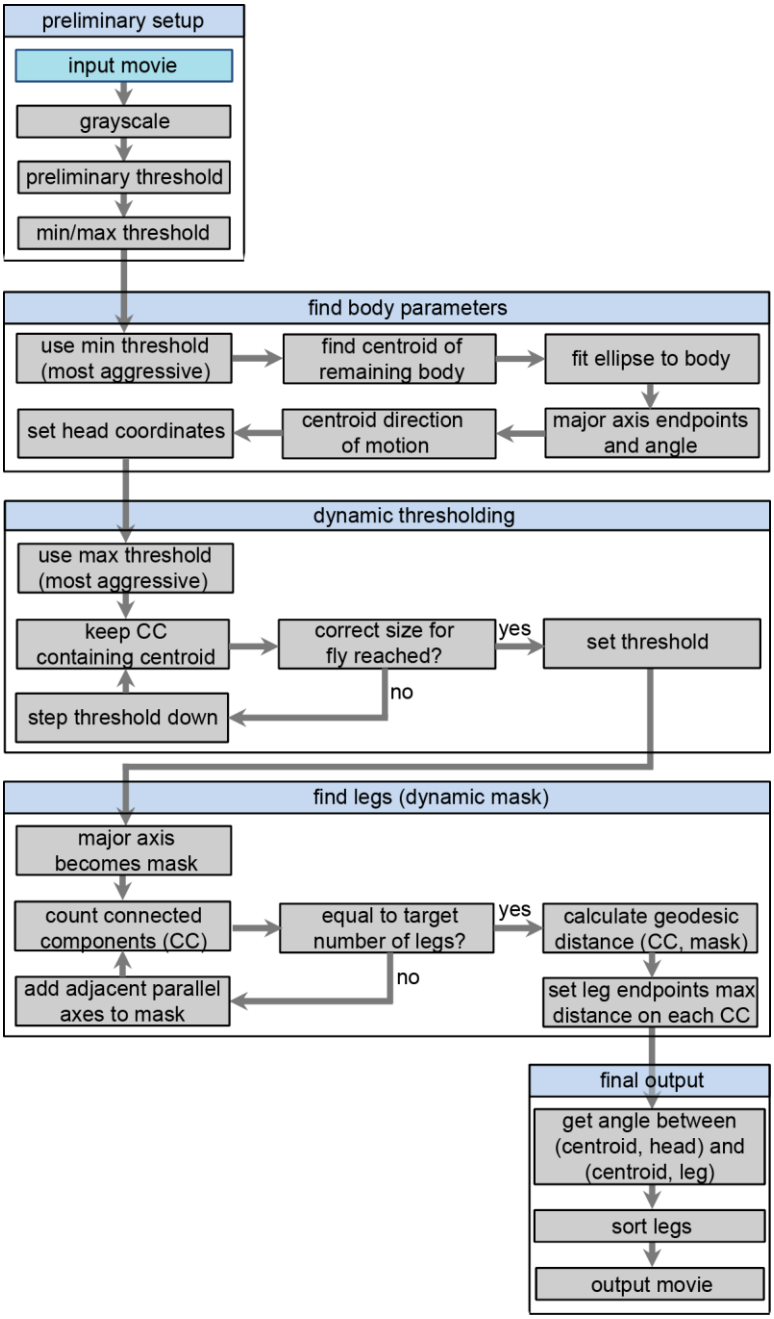

Figure 4.3. Automatic tracking algorithm logic flow chart. The algorithm takes a grayscale movie and applies the Otsu method (Otsu, 1979) to quickly find a preliminary threshold for converting the frames to binary images. It then sets a maximum and minimum threshold based on the preliminary threshold. Using the most aggressive threshold to remove all "non-body" points (including legs), we find the centroid of the body and the major axis by fitting an ellipse. Front/back symmetry is broken by computing centroid motion and choosing the head as the extremal point on the axis in the direction of centroid motion. Then, we use a dynamic thresholding approach by sweeping through the acceptable threshold values from maximum (least aggressive) to minimum (most aggressive) and automatically identifying the connected component (CC) of interest as the one containing the centroid. The acceptance criterion is based on the number of pixels retained in the

Figure 4.3 (Continued) CC with the fly being reasonably close to what is expected of a fly (acceptable parameter values were found through testing). Once the threshold is set, we keep only the CC containing the fly centroid. We then use what we term "dynamic masking" to find leg endpoints. The major axis becomes the first mask (the pixels are removed) and we count the number of CC. Until the number of connected components is equal to the required number of legs (6 legs pre-amputation, 5 legs post-amputation), expand the mask by adding the closest parallel axis to the central axis at either end of the current mask and recalculate. That is, remove more and more of the fly body in "slices" parallel to the major axis until the appropriate number of connected components remains. Then, calculate the minimum geodesic distance from the mask to all points within each CC. The leg endpoints are chosen to be the pixels with the maximal distance measure on each connected component (if several pixels share the same distance within one connected component, choose randomly). This algorithm can yield accuracies greater than 95%, but only with high resolution images. In the case of our camera, the algorithm tracked <10% of frames (i.e. all 5 or 6 legs) sufficiently accurately, so most frames were hand-corrected.

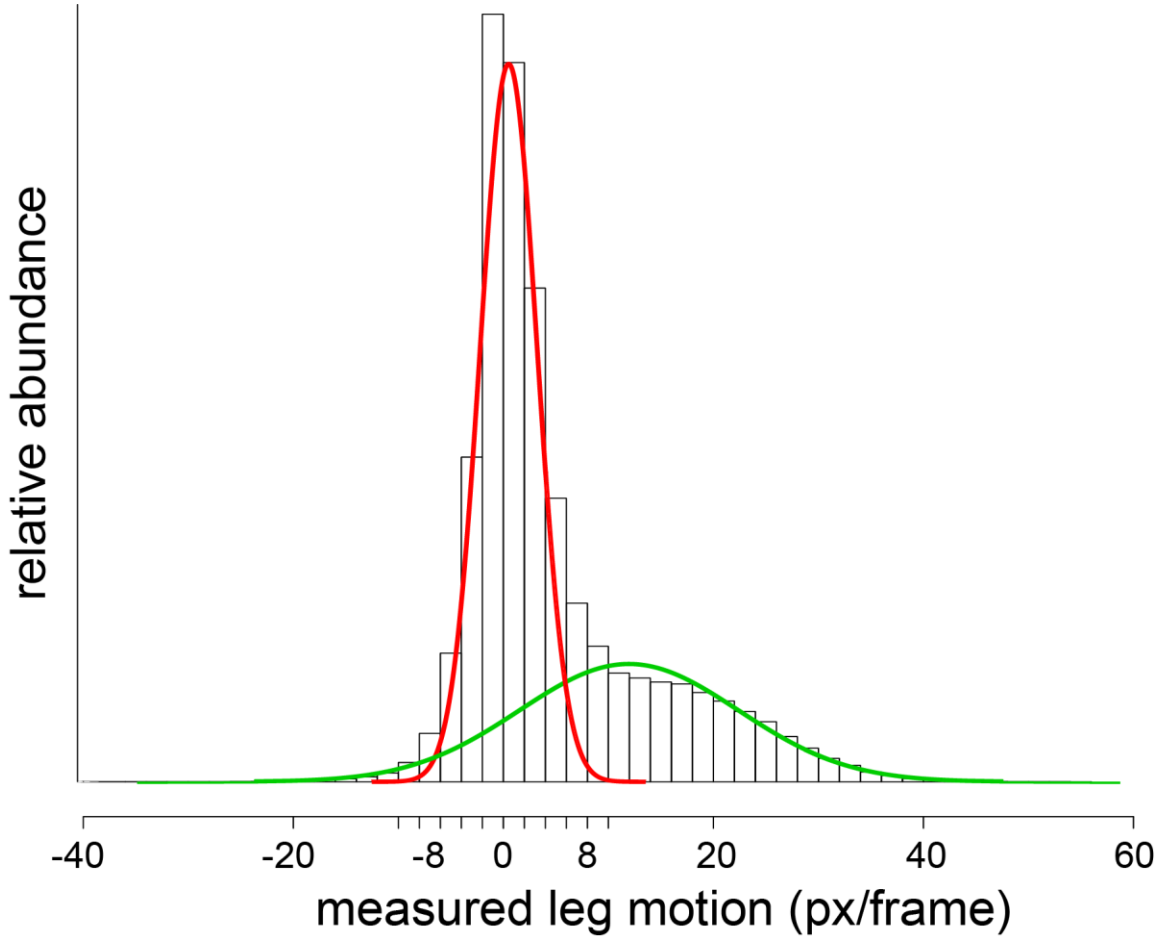

Figure 4.4. Histogram of frame-to-frame leg velocities with a Gaussian mixture model overlay. Frame-to-frame leg motion based on semi-automated video tracking appears to be bimodal and composed of two components: experimental tracking error and true motion. We used a

Figure 4.4 (Continued) Gaussian mixture model with two underlying distributions to fit the data. The left distribution (red) corresponds to tracking error, the right (green) to true motion. We chose a threshold of 8 pixels/frame to distinguish real motion from measurement errors below the 3 SD mark of the left distribution (mean = 0.51, SD = 2.70).

Hidden Markov Model

A Hidden Markov Model (Baum and Petrie, 1966) assumes unobserved ("hidden") internal states for a system (e.g. gait), each of which results in emitting a measured signal from a set (e.g. number of observed legs moving) with some probability. While a common way of characterizing gaits is by consider the number of legs moving simultaneously in each frame (Kain et al., 2013; Mendes et al., 2013), this methodology has several weaknesses, including potentially introducing artifacts due to imaging (misclassifying gaits due to imperfect simultaneity in take-off) that require smoothing (Figure 4.5).

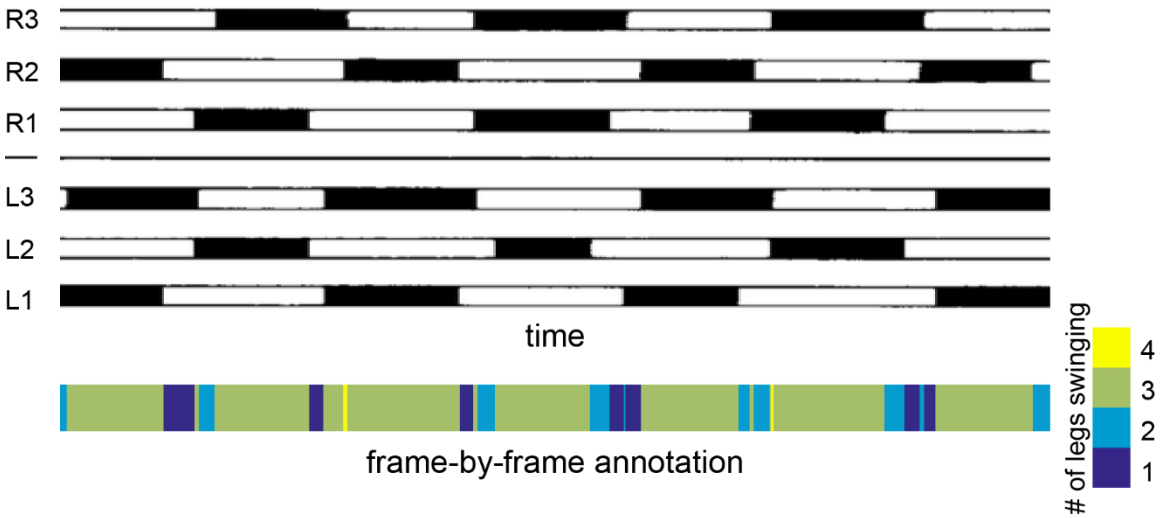

Figure 4.5. Canonical tripod gait with frame-by-frame annotation. Canonical tripod gait plot from (Strauss and Heisenberg, 1990). Black indicates legs in swing phase, white legs in stance phase. Using frame-by-frame annotation as in (Kain et al., 2013) leads to frames with greater than and fewer than 3 legs moving, which requires smoothing (Mendes et al., 2013). An HMM provides an algorithmic way to smooth gait annotations.

Figure 4.6A shows an example of annotated movie frames. These can also be visualized in stride-stance plots (Figure 4.6B) that show the legs in stance (white) and swing (blue) as a function of time (compare the top panel of Figure 4.6B and Figure 4.5). Instead, we used an HMM-based approach. We used 3 hidden states (1-leg, 2-leg, and 3-leg gaits) and four observed states (0-3 legs moving). A schematic of the HMM with emission and transition probabilities for wild type flies pre-amputation is shown in Figure 4.6C. To fit the parameters of this model, we used the Baum-Welch algorithm.

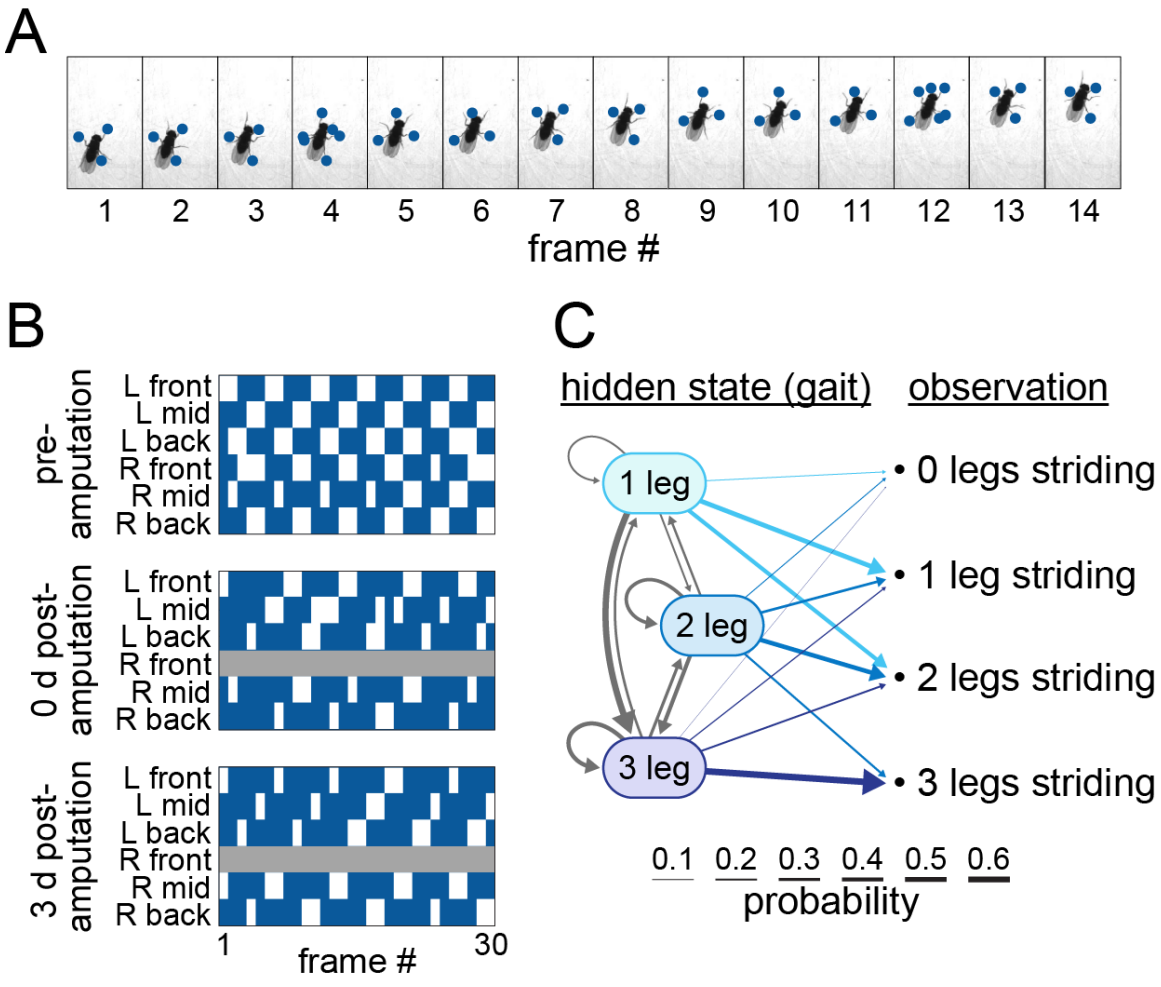

Figure 4.6. Example of gait and schematic of Hidden Markov Model (HMM). A. Example of a fly with legs moving in alternating tripod gait. Blue circles indicate that a leg will remain stationary in the next frame. **B.** Examples of stride-stance plots for a wild type fly at pre-amputation (top), immediately post-amputation (middle), and three days post-amputation (bottom). Frames are on the

Figure 4.6 (Continued) horizontal axis and legs are on the vertical axis. White indicates that a leg is in swing in that frame, blue indicates that a leg is in stance, and gray indicates amputated legs. The top panel shows a typical tripod gait. On day 0 post-amputation, there is a mix of non-canonical gait as well as potentially remnants of tetrapod and tripod. On day 3 post-amputation, a clearer tetrapod-like pattern emerges. **C.** A schematic of the HMM used to determine gaits. Three hidden states (1-leg, 2-leg, 3-leg gaits) each have a probability of emitting frames with 0, 1, 2, or 3 legs in swing phase. Weights correspond to probabilities for wild type flies pre-amputation. See Figure 4.7 for transition probabilities.

First, we aggregated all fly information by strain, stratifying into pre- and post-amputation and found emission probabilities. Then, we obtained transition probabilities for each strain and day (Figure 4.7). Finally, to obtain gaits, we fit internal states using the Viterbi algorithm as follows. At each step through the chain, the algorithm has stored the probability of being in every hidden state at the previous step, having come along the likeliest path so far. It then calculates the probability of being in each hidden state at the current step given the probabilities of being in each hidden state at the previous step (using the transition probability) and the observed value at the current step (using the emission probability of the observation), and remembers the hidden state that is most likely from this set of options. This state is appended to the likeliest path. The process repeats until the end of the chain, creating the full likeliest path through the hidden states. For a high-level introduction to HMMs with an application to biology, see (Eddy, 1996).

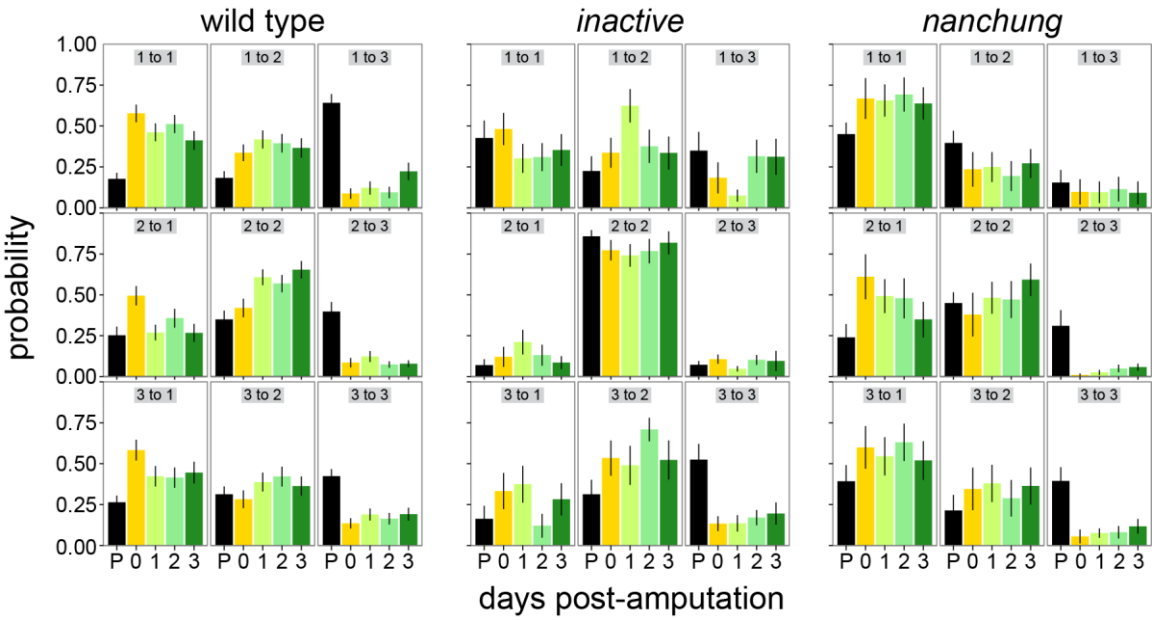

Figure 4.7. Transition probabilities between gaits by strain by day from the HMM. Numbers with gray background correspond to gait type (that is, "1 to 2" means "1-leg gait to 2-leg gait"). Error bars are ±1 SEM.

We verified the results of the HMM on a frame-by-frame analysis (Figure 4.8) of leg motion in the pre-amputation data. While the boundaries between "canonical" gaits even in intact animals are not always sharp, a frame-by-frame analysis of the most frequent pattern of leg striding during 3-leg gaits pre-amputation was the standard alternating tripod (135)(246) and the most frequent patterns of 2-leg and 1-leg gaits correspond respectively to traditional tetrapod and wave patterns (Figure 4.9). We found qualitative consistency between the methods: more frames (5%) were labeled tripod by the HMM overall than with the traditional approach, as expected, and 83% of frames agreed on tripod labeling with a frame-by-frame approach; more importantly, 15% of frames with 2 legs moving and 10% of frames with 1 leg moving were labeled tripod. Also intuitively, the state transition probabilities from 1-leg, 2-leg, and 3-leg gaits to the 3-leg gait from the HMM were high, suggesting that this gait was both more common and more persistent when it occurred (Figure 4.7). This is not surprising, since we selected for fast locomotor bouts and higher speed is associated with greater persistence of an alternating tripod (Strauss and Heisenberg, 1990; Mendes et al., 2013). A

comparison of statistical significance of post-amputation trends also yielded similar results to frame-by-frame analysis (Figure 4.8).

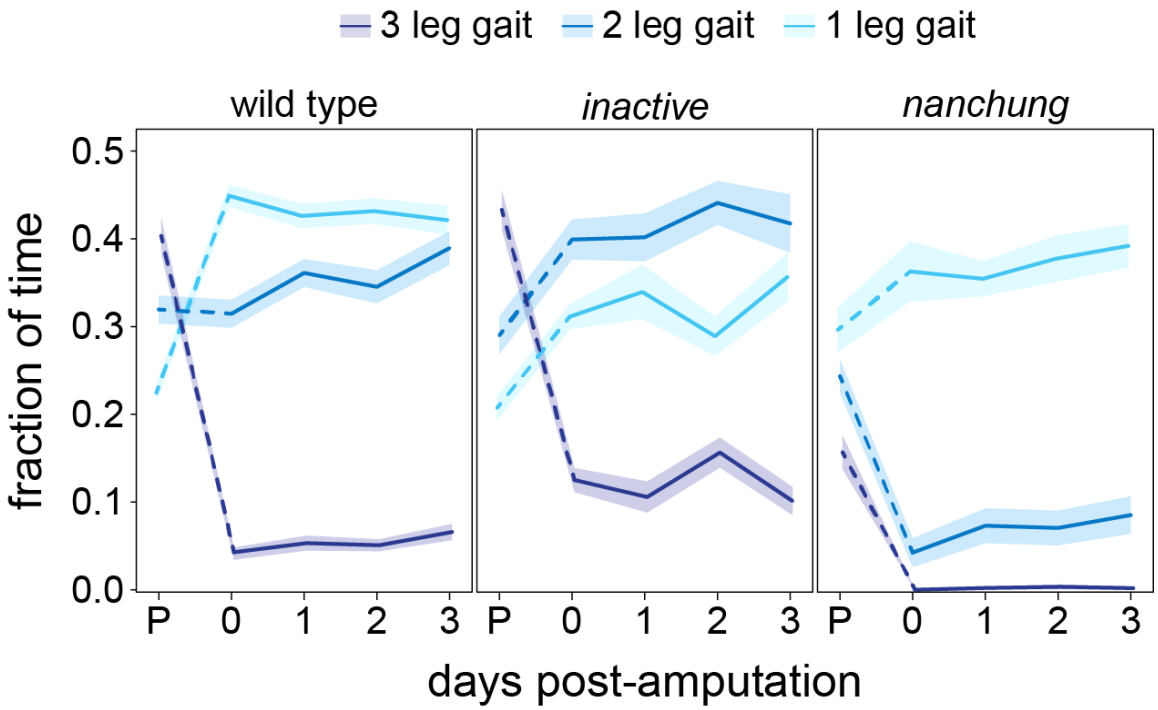

Figure 4.8. Gait frequency with frame-by-frame analysis. Frequency of 1, 2, 3-leg gaits from frame-by-frame analysis of moving legs (rather than HMM analysis). We see a qualitatively similar pattern to Figure 4.15A. A statistical analysis reveals that, as in the case of the HMM, there is no significant difference in any gaits over time post-amputation for either *inactive* or *nanchung* ($P > 0.168$ for the slope in all conditions, F-test). In wild type flies, while both the HMM and the frame-by-frame analysis do not show significant change of the 3-leg gait over time ($P > 0.072$ for both) and both show a statistically significant increase in 2-leg gait over time ($P < 0.008$ for both), the HMM suggests a significant decrease in the 1-leg gait ($P = 0.004$), while statistical significance is absent in the frame-by-frame analysis ($P = 0.222$).

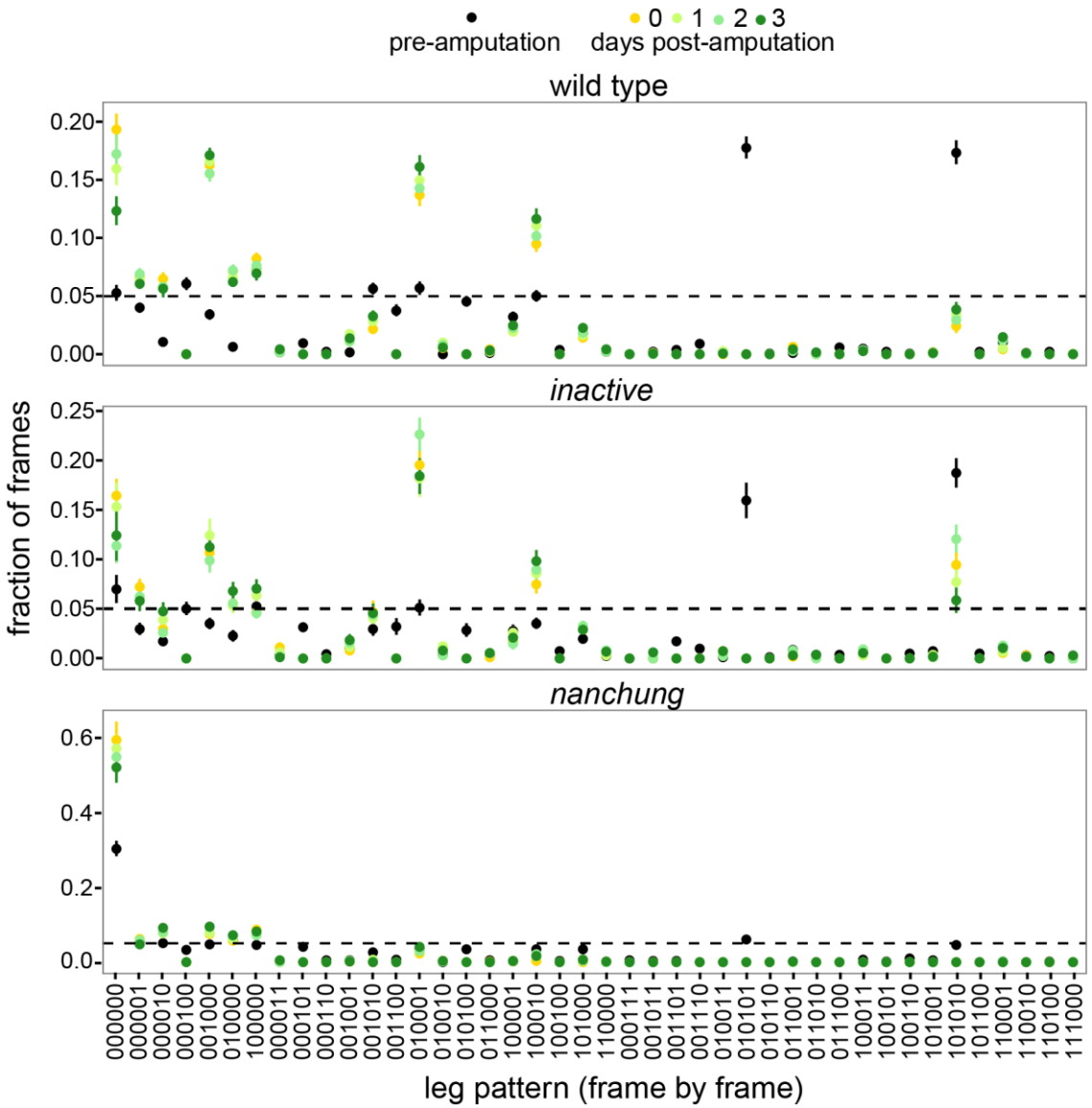

Figure 4.9. Frequency of specific leg patterns in a frame-by-frame analysis. Six-digit label indicates which legs are moving. Digits are ordered the same way as the legs in Figure 4.1A from bottom to top. A 1 represents leg in swing phase while a 0 represents leg in stance phase. E.g. 101010 indicates that legs 1, 3 and 5 are in swing phase and legs 2, 4 and 6 are in stance phase – an instance of alternating tripod gait. The dashed line is a visual guide at 5%. We see that largely canonical gaits are exhibited pre-amputation.

Neuromechanical model

Complementing the experiment, we also built a neuromechanical model that takes into account body motion, leg motion, and a neural controller (see Figure 4.10A for a schematic, with capital letter variables corresponding to the body, n represents neural modules, and other lowercase

letters represent the legs, and indices run from $1:6$. Without loss of generality, we let module 1, corresponding to the front left leg, be the "clock" relative to which all other leg phases were measured. When a module is excited beyond a threshold, it drives its corresponding leg to lift off the ground (enter swing phase). Legs relax quickly (relative to the excitation duration) to their respective forward-most position relative to the body. When the module activity drops below the threshold, the leg is placed down on the substrate and exerts force on the body until the next neuronal excitation lifts the leg. Body dynamics obey Newton's laws.

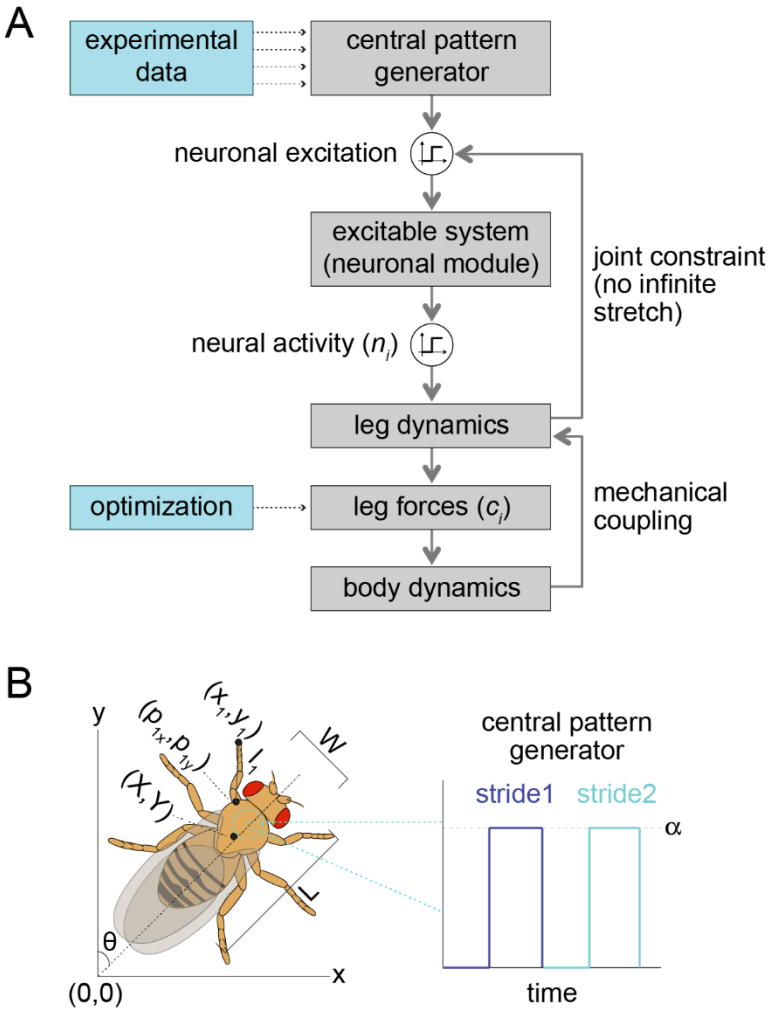

Figure 4.10. Schematic of neuromechanical model of walking. A. Schematic of the theoretical model. Data was used to fit all relevant parameters except leg forces (Table 4.1), which were fit through an optimization procedure. A central pattern generator sends a signal to the excitable

Figure 4.10 (Continued) system (neuronal module). Once the neuron reaches a threshold, it fires and the leg that it controls lifts up and follows its own dynamics relative to the body to prepare for the next step. When the neuron drops below the threshold, the leg exerts forces on the body, which undergoes Newtonian dynamics in (x, y, θ). A joint parameter, coupling back to the neural system, prevents infinite stretch. A full stride occurs when all 6 legs (5 in the amputated case) and the central pattern generator repeats its pattern. **B.** Schematic showing model notation corresponding Equations (4.1)-(4.9), and Table 4.1. A signal from the central pattern generator (located in the fly's ventral nerve cord in the thorax) to one limb is shown in the inset and mediates the neuronal dynamics that control that limb.

Following the notation used in Table 4.1 and Figure 4.10B, we can then write the governing equations coupling the neurodynamics to the body and limb mechanics, as

$$M\ddot{X} = \left(\sum_{i=1}^{6} f_{ix}\right) - B_1\dot{X} \tag{4.1}$$

$$M\ddot{Y} = \left(\sum_{i=1}^{6} f_{iy}\right) - B_1\dot{Y} \tag{4.2}$$

$$I\ddot{\Theta} = \left(\sum_{i=1}^{6} m_i\right) - B_2\dot{\Theta} \tag{4.3}$$

$$\begin{bmatrix} x_i^* \\ y_i^* \end{bmatrix} = R_\Theta\left(\begin{bmatrix} \lambda_i l_i^* \cos(\theta_i^*) \\ l_i^* \sin(\theta_i^*) \end{bmatrix} + \begin{bmatrix} p_{ix} \\ p_{iy} \end{bmatrix}\right) + \begin{bmatrix} X \\ Y \end{bmatrix} \tag{4.4}$$

$$\tau_L\dot{x}_i = (x_i^* - x_i)H(\hat{n} - n_i) \tag{4.5}$$

$$\tau_L\dot{y}_i = (y_i^* - y_i)H(\hat{n} - n_i) \tag{4.6}$$

$$\begin{bmatrix} f_{ix} \\ f_{iy} \end{bmatrix} = R_\Theta\begin{bmatrix} 0 \\ c_i \end{bmatrix}H(n_i - \hat{n}) \tag{4.7}$$

$$m_i = (x_i - X)f_{iy} - f_{ix}(y_i - Y) \tag{4.8}$$

$$\dot{n}_i = \frac{1}{\tau_n}\left(-n_i - v_S(H(l_i - (1+s)l_i^*) + H(-l_i + (1-s)l_i^*)\right) + \alpha\mathbb{I}_{\mathrm{mod}(\omega t - \phi_i, 1) > (1-\delta_i)} \tag{4.9}$$

where Equations (4.1)-(4.3) determine the location of the body center of mass and its orientation as a function of the leg forces and torques, Equation (4.4) determines the positions of the tips of the legs in terms of the location and orientation of the body, Equations (4.5), (4.6) characterize the over-

damped dynamics of the legs as a function of the neuronal dynamics, Equations (4.7), (4.8) characterize the forces and torques exerted by the legs, and Equation (4.9) characterizes the neuronal dynamics controlling the legs. To ensure that the effective leg lengths did not exceed their total lengths and prevent unrealistic stances, we also imposed some physical length limits via feedback into neural excitation.

Due to the number of legs and degrees of freedom, the model necessarily has a number of parameters. Of the 48 independent equation parameters, 13 were fit for each fly and day (mean ± SD for wild type pre-amputation): the excitation pulse frequency (ω) (11.4 ± 1.8 pulses/second), the proportion of time per stride that each leg spent in stance ($\delta_{i=1:6}$) ([0.62, 0.70, 0.66, 0.65, 0.70, 0.68] ± [0.07, 0.06, 0.06, 0.07, 0.06, 0.06]) and the phase of each leg relative to leg 1 ($\phi_{i=1:6}$) ([0, 0.58, 0.17, 0.48, 0.13, 0.60] ± [0, 0.07, 0.07, 0.06, 0.07, 0.09]). The other parameters were fixed from average values reported in the literature and observed in experiment.

Table 4.1. Model symbols and descriptions. Gray background denotes variables, light yellow background represents mathematical functions, dark yellow background denotes parameters that set an overall scale/threshold, blue background denotes derived variables/parameters, light green background denotes parameters found from literature/experiment, red background denotes tuning parameters. Model units are given on the scale of the fly (b.l.u. = body length unit = 2.5 mm, b.m.u. = body mass unit = 0.25 mg, a.u. = arbitrary unit for threshold). Values reported for parameters that vary in the simulation $(\omega, \phi_i, \delta_i)$ are means for wild type flies pre-amputation.

Symbol	Description	Initial Condition, Definition, Value
$\boldsymbol{X, Y}$	Center of mass position (x, y coordinates)	$X(0) = Y(0) = 0$ $\dot{X}(0) = \dot{Y}(0) = 0$
$\boldsymbol{\Theta}$	Body angle (from vertical)	$\Theta(0) = \dot{\Theta}(0) = 0$
$\boldsymbol{x_i, y_i}$	Leg endpoint (x, y coordinates)	$x_i(0) = x_i^*$ $y_i(0) = y_i^*$

Table 4.1 (Continued)

$\boldsymbol{n_i}$	Neuron activity	$n_i(0) = \begin{cases} 1, i = 1,3,5 \\ 0, \ i = 2,4,6 \end{cases}$
$\boldsymbol{t}$	Time	0 s
$\boldsymbol{H(x)}$	Heaviside function	$H(x) = \begin{cases} 1, x \geq 0 \\ 0, x < 0 \end{cases}$
$\boldsymbol{\lambda_i}$	Left/right index function	$\lambda_i = \begin{cases} -1, i \in \{1,2,3\} \\ 1, i \in \{4,5,6\} \end{cases}$
$\boldsymbol{R_\theta}$	Clockwise rotation matrix (θ is positive in the counterclockwise direction)	$\mathrm{R}_\theta = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix}$
$\mathbb{I}_C$	Indicator variable	$\mathbb{I}_C = \begin{cases} 1, C = \text{True} \\ 0, C = \text{False} \end{cases}$
$\boldsymbol{M}$	Mass	1 b.m.u.
$\boldsymbol{L}$	Body length	1 b.l.u.
$\boldsymbol{\tau_L}$	Leg relaxation time-scale	10 ms
$\boldsymbol{\tau_N}$	Neuron relaxation time-scale	10 ms
$\boldsymbol{v_S}$	Neuron relaxation amplitude if stretch exceeds bounds	1 a.u.
$\boldsymbol{\hat{n}}$	Neuron firing threshold	0.9 a.u.
$\boldsymbol{\alpha}$	Neuron excitation amplitude when excited by the central pattern generator	1 a.u.
$\boldsymbol{I}$	Inertia	0.01 b.m.u. x b.l.u.^{2}
$\boldsymbol{m_i}$	Torque from leg	$m_i = (x_i - X)f_{iy} - f_{ix}(y_i - Y)$

Table 4.1 (Continued)

$\boldsymbol{f_{ix}, f_{iy}}$	Force from leg i in the x, y direction	$\begin{bmatrix} f_{ix} \\ f_{iy} \end{bmatrix} = R_\Theta \begin{bmatrix} 0 \\ c_i \end{bmatrix} \mathrm{H}(n_i - \hat{n})$
$\boldsymbol{l_i^*}$	Relaxed leg length (relative to body)	$l_i^* = \begin{cases} 0.59, i = 1{,}4 \\ 0.66,\ i = 2{,}5 \\ 0.42,\ i = 3{,}6 \end{cases}$
$\boldsymbol{\theta_i^*}$	Relaxed leg angle (relative to body)	$\theta_i^* = \begin{cases} 1.19,\ i = 1{,}4 \\ 0.33, i = 2{,}5 \\ -0.69,\ i = 3{,}6 \end{cases}$
$\boldsymbol{l_i}$	Leg length	$l_i = \left\lvert \begin{bmatrix} x_i \\ y_i \end{bmatrix} - \left(R_\theta \begin{bmatrix} p_{ix} \\ p_{iy} \end{bmatrix} + \begin{bmatrix} X \\ Y \end{bmatrix} \right) \right\rvert$
$\boldsymbol{W}$	Body width	0.34 b.l.u.
$\boldsymbol{B_1}$	Translational damping. Body is in over-damped regime	1.5 b.m.u./s
$\boldsymbol{B_2}$	Rotational damping. Body is in over-damped regime	1.5 b.m.u. x b.l.u^2/s
$\boldsymbol{p_{ix}, p_{iy}}$	Position of leg-body attachment point (x, y coordinates)	$p_{ix}(0) = \lambda_i * 0.05$ $p_{iy}(0) = \begin{cases} 0.20,\ i = 1{,}4 \\ 0,\ i = 2{,}5 \\ -0.11, i = 3{,}6 \end{cases}$
$\boldsymbol{x_i^*, y_i^*}$	Relaxed leg endpoint position (x, y coordinates)	$x_i^*(0) = \lambda_i \begin{cases} 0.27, i = 1{,}4 \\ 0.67,\ i = 2{,}5 \\ 0.37,\ i = 3{,}6 \end{cases}$ $y_i^*(0) = \begin{cases} 0.75, i = 1{,}4 \\ 0.21,\ i = 2{,}5 \\ -0.37,\ i = 3{,}6 \end{cases}$
$\boldsymbol{s}$	Maximum stretch ratio ("physical joint")	2
$\boldsymbol{\omega}$	Excitation pulse frequency	11.4 strides/s
$\boldsymbol{\delta_i}$	Proportion of time leg is down per stride	[0.62, 0.70, 0.66, 0.65, 0.70, 0.68]

Table 4.1 (Continued)

$\boldsymbol{\phi_i}$	Excitation pulse phase relative to ϕ_1 (when leg is lifted)	[0, 0.58, 0.17, 0.48, 0.13, 0.60]
$\boldsymbol{c_i}$	Force magnitude	optimization

We used this model to simulate multiple strides with a first-order method for numerical simulation with a step size $h = 0.001$. After a short transient (2-3 strides), the walking behavior converged to a steady state and generated a speed of approximately 0.65 body lengths/stride, consistent with published results (Wosnitza et al., 2013) and our own measurements. As in experimental observations, flies with right foreleg amputation have a counter-clockwise locomotor angular bias.

Angular bias-turning bias calibration

In order to convert model angular bias output to experimentally observed turning bias in the arena, we built a calibration curve. We simulated simple rules for walking flies in an arena: start at a random place in the arena facing in a random direction. Move at a constant speed in the chosen direction. At each step k, choose a new direction using a local rule: $\phi_k = \phi_{k-1} + \phi'$, where $\phi' \sim N(\beta, \Delta)$. Here, β is the angular bias (the horizontal axis of the calibration curve) and Δ is random heading drift (the standard deviation of the normal distribution). If a wall collision occurs, choose a new heading by masking out the original heading distribution (disallowing angles that would result in a wall collision) and renormalizing. If this would cause a machine precision error, choose the new heading to be the angle between the current position and the center (θ) modified by $\pm\left(\frac{\pi}{2} + \varepsilon\right)$, where ε is a small parameter and the negative sign is chosen with probability $\frac{|\theta + \frac{\pi}{2} - \phi_{k-1}|}{\pi}$, with the numerator lying in the range $(0, \pi)$. The paths formed by this procedure qualitatively mimic

those of real flies. As in the calculation of μ – the weighted average ($\theta - \delta$) Figure 4.1B) – from our experimental arena data, we calculated μ from simulated paths by only examining behavior when the fly is within a distance $0.8R_a$ from the arena center, where R_a is the arena radius.

To build the calibration curve, we first matched heading drift by fitting simulated $\theta - \delta$ histograms at a fixed bias $\beta = 0$ to the corresponding wild type pre-amputation histogram from experiment and minimized least square error. Since the fitted heading drift ($\Delta = 0.035$) did not give a large enough dynamic range for μ to recapture all experimental results, we shifted it to be as close as possible to the fit while capturing the necessary dynamic range ($\Delta = 0.029$). The paths generated by this compromise heading drift value are still qualitatively reasonable (Figure 4.11). Finally, we built the curve (Figure 4.12) by sweeping β and plotting μ (averaged over 50 runs, each run consisting of 10^4 simulation steps). Fitting a quadratic function gives an excellent fit over the range of interest ($R^2 = 0.999$). We used the analytical expression for the fit as the final calibration curve.

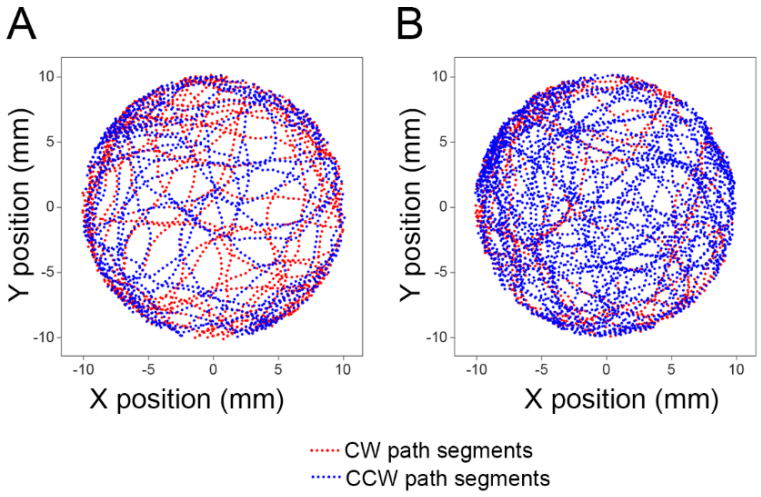

Figure 4.11. Random walk simulation path examples. A. Typical simulated paths (left panel): angular bias = 0, (right panel): angular bias = 0.03. Blue indicates a clockwise path segment and red indicates a counterclockwise segment. Thus, tuning the angular bias parameter allows us to mimic the turning bias imparted by amputation.

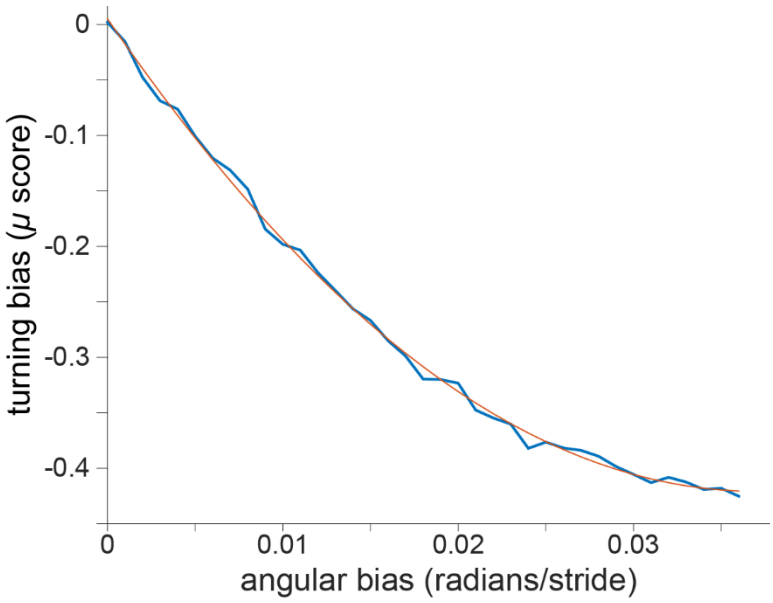

Figure 4.12. Calibration curve for converting between model output (angular bias) and arena locomotion turning bias (μ). Heading drift was chosen to be closest to a best fit to pre-amputation wild type turning bias histogram data (with bias = 0 in the simulation) while permitting the necessary dynamic range for μ. The blue curve is average value over 50 runs for each value of angular bias. The red curve is a quadratic line of best fit ($R^2 = 0.999$).

Optimization

The optimization goal was to find a set of leg forces acting on a fly with averaged parameters to match the experimental μ from the calibration curve (for each of the three strains and each day determined separately). We ran the model to 5 strides (to steady state) and took the angular bias in the last stride to calculate the energy of the proposed solution. We ran a simulated annealing optimization procedure (Kirkpatrick Jr et al., 1983) using a normalized annealing schedule with initial temperature $T_1 = 1$, an annealing interval of $I_a = 75$ steps, and a multiplicative cooling factor $C_a = 0.9$. A full simulation was run without re-heating to $N_a = 1.5 \times 10^3$ steps, which was found to be sufficient to consistently converge. At each step k, the number of forces that were changed for the proposal was $\lceil 3T_k \rceil$. Starting with $c_i = 0.25$ on each leg, we allowed the forces on the left side to change (allowing all forces to change leads to similar results but takes longer to converge):

$c_i^{proposed} = c_i^{current} + u$, where $u \sim U(-T_{current}\tau_m, T_{current}\tau_m)$ and τ_m is a parameter that makes the proposal distribution tighter. Proposals for each leg were constrained to lie in the interval $f_a = [0.1, 0.25]$ (again for faster convergence – in practice, the average optimal forces on the left side were not even close to the lower bound even for the lowest target μ).The energy of a proposed solution was the distance between the absolute values of the target bias given the real μ (from the calibration curve) and the angular bias as calculated from model output (angle difference during the last stride/distance moved during the last stride). The acceptance probability was a Boltzmann function with normalized energy and a multiplicative "convergence factor" γ: $p = e^{-\frac{\gamma(E_{proposed} - E_{current})}{T_{current} E_{current}}}$. Symbols used in the calibration curve procedure and simulated annealing are summarized in Table 4.2.

Table 4.2. Calibration curve and simulated annealing symbols and descriptions. Green background is for symbols used in calibration curve simulation and blue background is for symbols used in simulated annealing.

Symbol	Description	Value
N_r	Number of runs	50
N_c	Number of simulation steps	10^4
ϕ	Heading direction	varies
β	Heading bias	varies
Δ	Heading drift	0.029
R_a	Arena radius	10.2 cm
N_a	Number of steps in annealing	1.5×10^3
I_a	Annealing interval	75
C_a	Cooling factor	0.9

Table 4.2 (Continued)

T	Temperature	varies
f_a	Proposal interval	[0.1, 0.25]
τ_m	Proposal tuning parameter	0.15
E	Energy of solution	varies
γ	Acceptance tuning parameter	5
p	Acceptance probability	$e^{-\frac{\gamma(E_{proposed}-E_{current})}{T_{current}E_{current}}}$

Results

Proprioception mediates locomotor recovery after injury

First, we investigated path-level behavior of adult *Drosophila* before and after amputation of the right foreleg. Individual flies were allowed to explore circular arenas for 2 hours, and the X-Y positions of their centroids were tracked and recorded. Approximately 24 hours later, the right foreleg was amputated and flies were recorded in the arena 1 hour, 24 hours, 48 hours and 72 hours post-injury (all sample sizes were between 9 and 30). Visual inspection of characteristic paths of wild type flies (Figure 4.13A) suggests that injury caused behavior to change from 1) paths composed of roughly equal portions of clockwise and counterclockwise segments to 2) highly biased, slow walking in the direction opposite to the leg that was removed immediately post-amputation and then 3) back to a largely unbiased walk three days post-amputation.

To provide a quantitative characterization of locomotor bias, we measured the "mu score" (Buchanan et al., 2015): the weighted average direction (μ) of the tangential component of the velocity relative to the center of the arena (Figure 4.1B). The vast majority of paths were around the

edges of the arena rather than directly inwards or outwards. Consequently, μ is largely invariant across spatial scales used in its calculation (i.e. the frame rate or mean interval used to determine direction of motion) within a broad range (Figure 4.14), suggesting that it is a robust measure of bias.

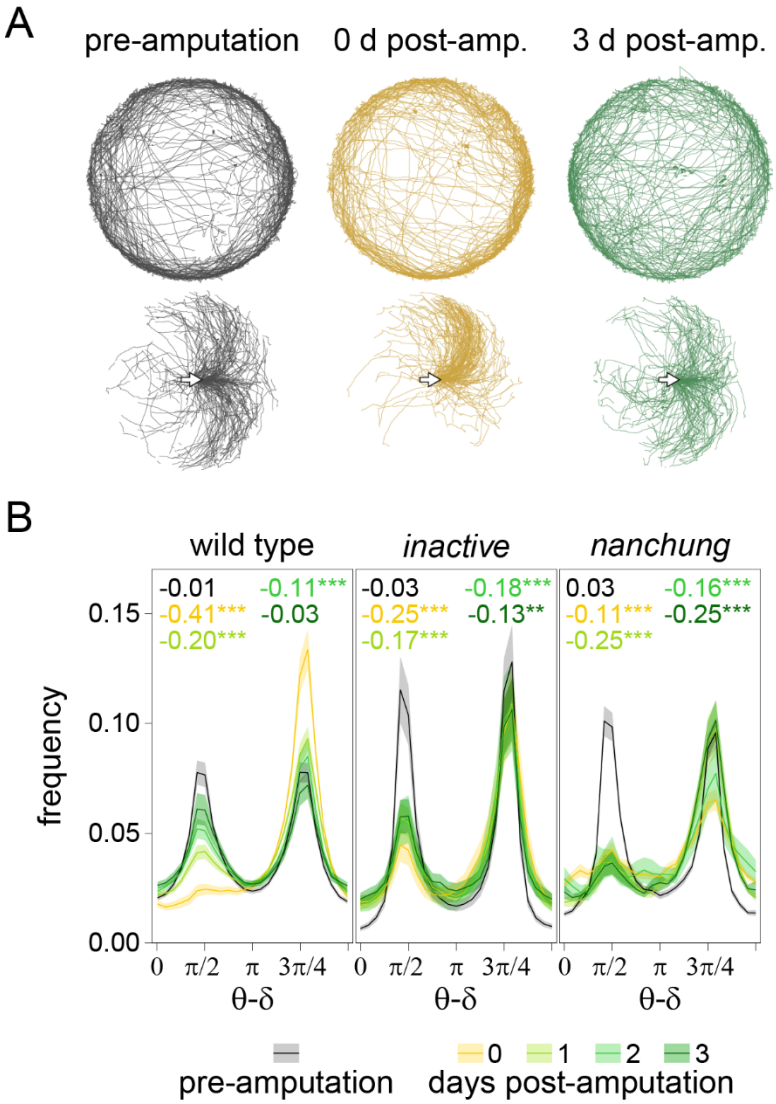

Figure 4.13. Turn bias recovery analysis. A. Representative sample fly paths over time (top panel). Gray is pre-amputation, yellow is day 0 post-amputation, green is day 3 post-amputation. Paths divided into segments of equal length (bottom panel), and aligned to start all in the same direction (arrow). The strong turning bias immediately post-amputation is evident. **B.** Histograms of turning behavior. Inset numbers indicate average μ value. Histogram symmetry about the center indicates unbiased behavior and μ values close to 0. Shaded regions indicate ±1 SEM ($9 < N < 30$ for all experimental groups). From a bootstrapping analysis, we find that the distribution for wild

Figure 4.13 (Continued) type flies is not significantly different between pre-amputation and day 3 post-amputation (P = 0.372), while the distributions are significantly different between pre-amputation and all post-amputation days in *inactive* and *nanchung* (P < 0.001). Stars indicate significance levels, here and elsewhere: *, P < 0.05; **, P < 0.01; ***, P < 0.001.

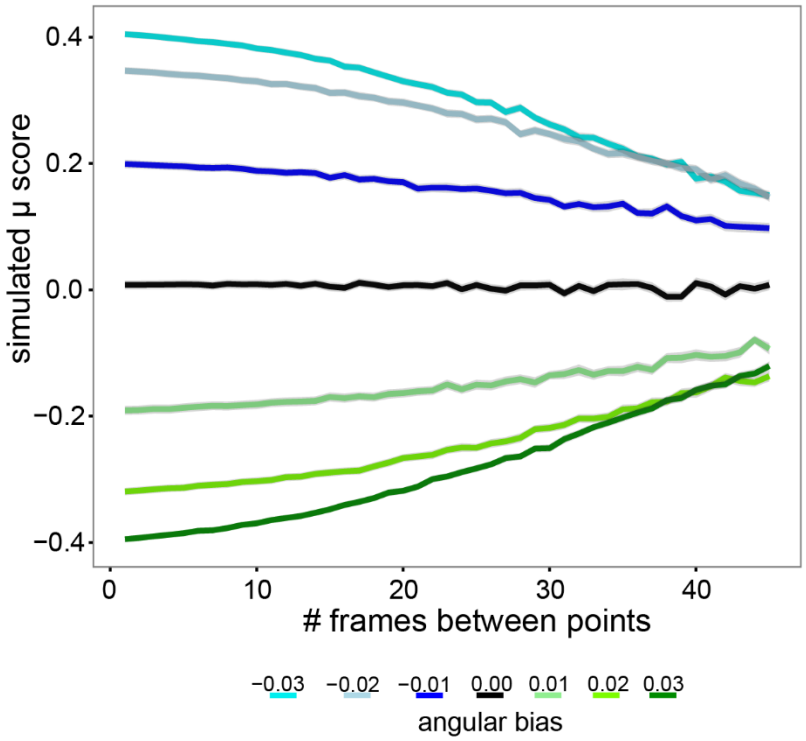

Figure 4.14. The μ score is broadly scale invariant. Calculation of μ at different effective frame rate (number of points used) from simulated fly paths on an arena. $\Delta = 0.029$, $\beta \in [-0.03, 0.03]$, $N = 100$ runs per point. Errors are ±1 SEM We find that μ is largely invariant over an order of magnitude difference in frame rate sampling.

A score of $\mu = 0$ corresponds to perfectly unbiased locomotion (flies moving clockwise and counterclockwise to the same extent) while $-1 < \mu < 0$ corresponds to an overall counterclockwise bias and $0 < \mu < 1$ corresponds to an overall clockwise bias. For wild type animals, we find that on average they start unbiased before amputation (μ = -0.006), develop a very strong bias immediately post-amputation (μ = -0.410), and steadily recover towards an unbiased state over the next three days (μ = -0.031) (Figure 4.13B). To compare μ values pre- and post-amputation, we performed a bootstrapping analysis, generating 10^5 $(\theta - \delta)$ histograms resampled from each strain's respective

pre-amputation histogram and calculating μ. We computed 5×10^5 experimental resamples, drawing random subsets with sizes matching the sample sizes of the post-amputation experiments of those μ values ($9 < N < 30$ for all experimental groups). The number of instances (k) out of 5×10^5 in which the pre-amputation distribution produced average μ values as or more extreme than observed in the post-amputation experiments was recorded. To be conservative in our estimation of the P-value, the upper bound on P at which k instances would be expected with probability greater than 0.025 was used prior to a P-value correction for 12 multiple comparisons using the formula $P^* = 1 - (1 - P)^{12}$. The difference in pre- and post-amputation distribution is not statistically significant after three days ($P = 0.372$).

We next sought to further characterize the mechanosensory basis of motor recovery. Since proprioception allows the fly to learn about the stretch and location of its limbs and thus control them and the forces they exert, we hypothesized that disrupting proprioceptive feedback would hinder a fly's ability to adapt its locomotor behavior post-injury. The TRPV ion channels Inactive and Nanchung are co-expressed in the proprioceptive organs of the fly, including the chordotonal organs of the femur, tarsi, and antenna. It is known that they are required for wild type locomotion and hearing (Kim et al., 2003; Gong et al., 2004). As with wild type animals, flies mutant for *inactive* (*iav*^{3621}) exhibited little clockwise/counterclockwise bias while exploring the arena pre-amputation (μ = -0.026), and biased walking immediately following injury (μ = -0.247). However, unlike wild type animals, *iav*^{3621} flies failed to recover close to their baseline (μ = -0.129 after three days). In *nanchung* mutants (*nan*^{36a}), the recovery failure is even more pronounced (μ = -0.250 after three days). For both *inactive* and *nanchung* mutants, the distribution on day 3 was still significantly different from pre-amputation ($P < 0.001$). In *nanchung* mutants, we observed a larger bias in the days following amputation than immediately post-amputation, with the bias on day 3 post-amputation being the

same as that on day 1 post-amputation. Overall, while mutants did not exhibit as large a bias as wild type flies immediately post-amputation, their turning bias persisted for the entire duration observed, in contrast to wild type. Together, the behavior of the wild type and mutants before and after amputation implicate proprioception as important for recovery. How this happens requires a detailed analysis at the level of individual legs.

Injury alters gait permanently

In order to gain insight into the biomechanical processes underlying recovery, we turned to a finer-grained analysis of leg motion. We recorded video of flies walking before amputation, followed by recordings 1 hour, 24 hours, 48 hours, and 72 hours after amputation. For the gait data, we analyzed $50 \leq N \leq 56$ (wild type), $16 \leq N \leq 17$ (*inactive*), and $13 \leq N \leq 15$ (*nanchung*) walking bouts at each time-point, with the slight variation due to fly death or post-processing rejection of runs. Instead of measuring locomotion across entire circular arenas, we captured bouts of fast, straight walking through the middle of arenas at 60 Hz. Using custom semi-automated leg-detection software we recorded the position of all 6 (or 5 post-amputation) legs frame-by-frame. Figure 4.6A shows annotated frames of a fly moving in a typical (135)(246) tripod gait. Figure 4.6B shows stride-stance plots to visualize leg positions on the ground (white) and off the ground (blue) as a function of time. The pre-amputation stride-stance plot is an example of a typical tripod pattern. Immediately post-amputation, we see a non-canonical gait with what may be residual hints of tripod or tetrapod gait. On day 3 post-amputation we see an apparent tetrapod-like gait.

One way of characterizing gaits is on a frame-by-frame basis by considering the number of legs that are concurrently in swing phase (Kain et al., 2013; Mendes et al., 2013). However, this approach is not always satisfactory for several reasons (Wosnitza et al., 2013). Experimentally, since in reality swinging different legs is not done exactly simultaneously, there can be error introduced by imaging, which requires smoothing (Mendes et al., 2013). Second, from the point of view of more

faithfully capturing the underlying processes, a frame-by-frame analysis does not treat gaits as persistent behavioral states. It is also not immediately apparent how to apply the standard gait categorization rules to flies with 5 legs, as many prior states become highly degenerate. Therefore, to estimate the frequency of internal gait states, we assigned a gait label to each movie frame that is not based on the observed pattern set of legs in swing phase in that exact frame, but is instead the state of a Hidden Markov Model (HMM). This captures the spirit of gaits as persistent internal states which have respective probabilities of showing 1, 2 or 3 legs moving simultaneously, and is an algorithmic alternative to hand-tuning windows. To avoid ambiguity, we refer to these hidden states as "3-leg", "2-leg", and "1-leg" gaits, without distinguishing between which groups of legs move (though as seen in Figure 4.9, the predominant 3-leg motions pre-amputation correspond to canonical tripod and 2-leg motions to canonical tetrapod).

This allowed us to consider the relative frequencies of 1-leg, 2-leg, and 3-leg gaits (Figure 4.15A). In all three genotypes – wild type, *inactive*, and *nanchung* – we observed that the 3-leg gait frequency drops dramatically from pre-amputation to immediately post-amputation. In all three genotypes 3-leg gait frequency did not have a significant change over the post-amputation period ($P > 0.060$, F-test), remaining near 0. Interestingly, wild type flies showed some gait plasticity (there was a significant increase post-amputation in the frequency of 2-leg gait, $P = 0.003$), whereas both 2-leg gait and 1-leg gait did not change discernibly over the three-day period in either *inactive* or *nanchung* ($P > 0.647$ for all conditions). For all strains, speed immediately post-amputation decreased relative to the pre-amputation value (by 34% for wild type, 14% for *inactive*, and 56% for *nanchung*). While there was an upward trend in wild type and *nanchung* flies over three days, speed did not return close to baseline at the end of the three-day period for any of the strains (Figure 4.15B).

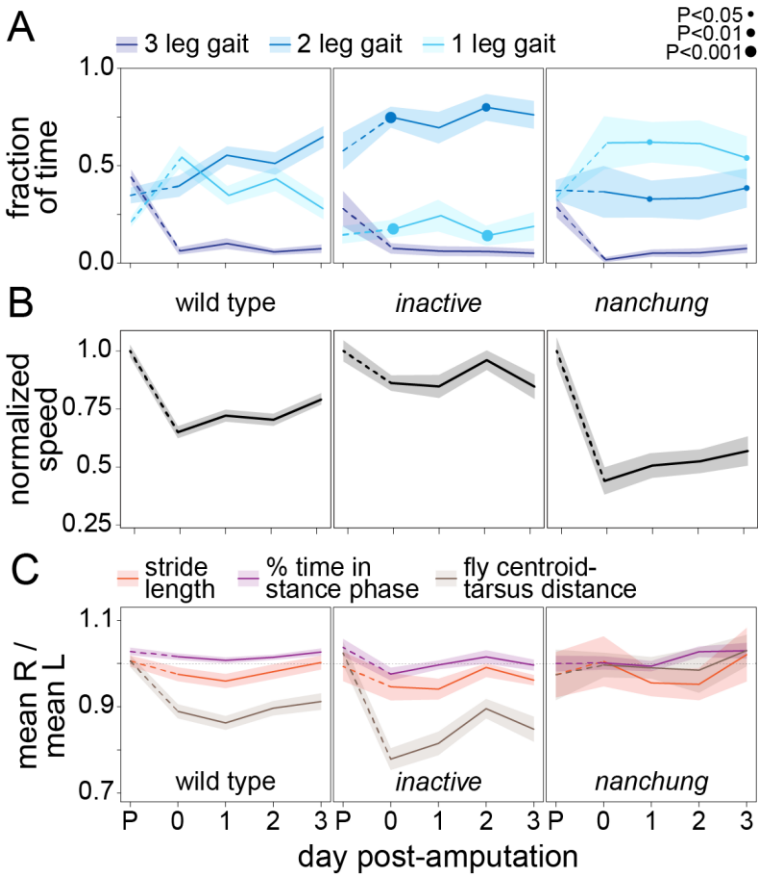

Figure 4.15. Analysis of parameters that may lead to recovery of unbiased turning. A. Frequencies of gaits for wild type, *inactive*, and *nanchung* mutants found from the HMM. N ≥ 50 (wild type), N ≥ 16 (*inactive*), and N ≥ 13 (*nanchung*) across all time-points. The frequency of 3-leg gait decreases from pre-amputation and does not recover significantly for any strain post-amputation (P > 0.060, F-test). Wild type flies exhibit some post-amputation gait plasticity, in contrast to the mutant strains. These results are correlated with walking speed. Individual dots on the post-amputation day lines for *inactive* and *nanchung* show significance levels of comparing gait frequency to the corresponding wild type time-point (unpaired t-test with Welch's correction). **B.** Speed by strain by day relative to pre-amputation value. Speed never recovers to the pre-amputation value, but wild type has a statistically significant upward slope (P = 0.001, F-test) which is not present in *inactive* (P = 0.741) or *nanchung* (P = 0.116). **C.** Ratio of right side average to left side average of various measures (leg distance moved per stride, proportion of time legs spend in on the ground, and leg distance from centroid at placement). The mean distance of the legs from the body centroid, at step placement, becomes lopsided (splayed to the left) after amputation for wild type and *inactive*, though it recovers significantly post-amputation for *inactive* (P = 0.017, F-test) and not wild type (P = 0.182). In *nanchung* mutants, it does not change from baseline. The mean distance moved by the tarsi on each side per stride and the proportion of time a leg is down in stance on the right side vs. the left side does not show significant recovery for any strain (P > 0.188).

Leg coordination pattern is correlated with walking speed, and hence the lack of recovery seen in both walking speed and gait may be related. Wild type flies walking at higher speeds tend to

use more legs (Wosnitza et al., 2013) and we found that this general pattern persists post-amputation. Interestingly, we found that the proportions of the number of legs swinging versus speed do not change significantly over the three-day period following amputation (the 95% confidence intervals overlap at nearly all points) (Figure 4.16).

Overall, as with 2-leg gait, walking speed shows a significant upward trend for wild type (P = 0.001, F-test) and not *inactive* (P = 0.741) or *nanchung* (P = 0.116). However, neither of these phenomena recapitulates the pattern seen in turn bias recovery. For instance, both wild type and *inactive* mutants exhibited a predominant proportion of 2-leg gait three days post-amputation (even though *inactive* mutants continue to have a turn bias), and the speed-coordination relationship remains largely the same at all days post-amputation for wild type. Thus, the mechanism of turning bias recovery could also lie elsewhere.

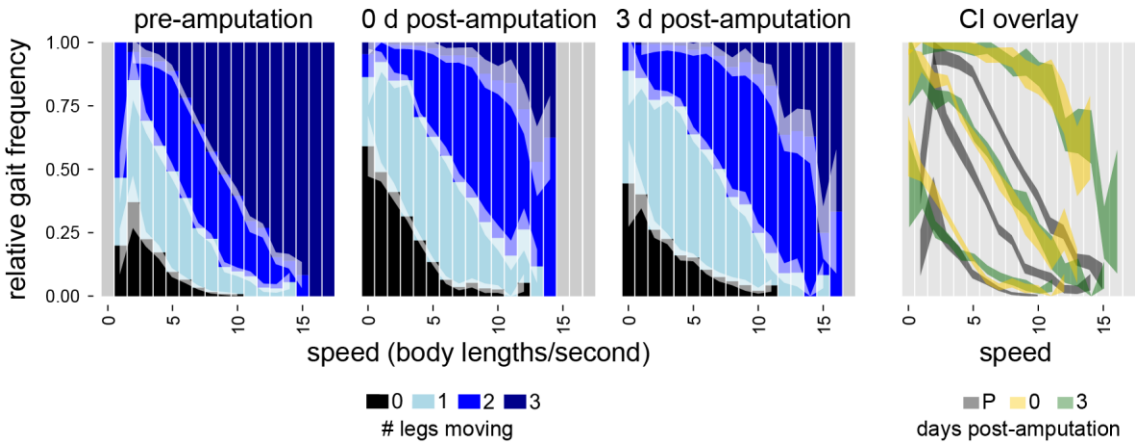

Figure 4.16. Relative gait frequency by speed range. Proportion of each gait within a speed bin (bin width = 1 body length/second) for wild type flies pre-amputation, 0 hours, and 72 hours post-amputation (three left panels; solid colors = mean, transparent white overlay = 95% confidence interval), and a separate overlay of 95% confidence intervals on the boundaries of coordination patterns for the three days (right panel; gray = pre-amputation, yellow = 0 hours, green = 72 hours post-amputation). Bins were only analyzed that had ≥ 5 data points. Pre-amputation, higher speeds correspond to a dominant 3-leg gait, and lower speeds to larger proportions of 2-leg and 1-leg gaits. The speed-coordination pattern does not change significantly over the three-day period post-amputation (confidence intervals overlap at nearly all points throughout the speed range).

Many leg parameters vary little through amputation and recovery

Turning back to the frame-by-frame analysis, we searched for other leg parameters (Figure 4.15C) that displayed dynamics matching those of turning bias, i.e. those responding to amputation in all genotypes (with a larger effect in wild type), and largely recovering by day 3 (relative to 1 hour post-amputation) in the direction of pre-amputation levels only in wild type animals.

We found that the mean distance of the legs from the body centroid, at step placement, becomes lopsided (splayed to the left) after amputation and fails to change discernibly for wild type ($P = 0.182$, F-test) but not *inactive* ($P = 0.017$). In the case of *nanchung* mutants, it does not change significantly from pre-amputation. The average distance moved by the tarsi on each side per stride does not change discernibly for any strain ($P > 0.188$). Similarly, the proportion of time a leg is down in stance on the right side vs. the left side stays essentially constant for all strains ($P > 0.221$) through the three-day period. Thus, none of these parameters on its own follows the qualitative pattern of turn bias recovery. However, it remains a possibility that the changes in a combination of these parameters in the action of walking could collectively explain turn bias recovery. To examine this, we now turn to a neuromechanical model of fly walking. If the virtual fly's turn bias follows the experimental turn bias pattern after fitting parameters from experiment, they may be sufficient to explain turn bias recovery.

Neuromechanical modeling implicates force modulation in recovery

The model framework (Figure 4.10A) featured three essential components: the brain which controls the legs, the body, and interaction with the environment mediated by the forces applied by the legs. We treated the neural component as a central pattern generator with 6 neuronal modules (one per leg) – see (Ijspeert, 2008) for an overview. We fit geometric and kinematic parameters with values reported in the literature (e.g. Mendes et al., 2013) and observed in our experiments (i.e.

Figure 4.15C), leaving only leg forces as free fitting parameters. The neuronal modules send a signal to the legs after reaching a threshold, which causes the legs to respond by exerting forces on the ground, so that the body moves according to the forces and torques it feels from the legs. This motion drives sensorimotor proprioceptive feedback to the neuronal module and the cycle repeats.

In order to compare the angular velocity output of the model (angular bias) to the arena locomotion turning bias, we determined a calibration curve (Figure 4.12). To do so, we used a simple arena path simulation that had angular bias as a tunable parameter and also captured realistic arena-scale behavior (Figure 4.11).

When we ran the model using all empirical parameters but held the force applied by each leg constant, all three genotypes exhibited no recovery in turn bias post-amputation (Figure 4.17A). If anything, all three lines exhibited increased bias with time, implying that the parameters we have measured so far are insufficient to explain recovery. Therefore, we examined whether force modulation would allow the model to exhibit turn bias recovery.

A Monte Carlo optimization approach allowed us to find a ratio of leg forces between right and left legs which yielded the appropriate angular bias. We did this for each strain and day using averages of the measured parameters. For example, to determine the leg forces needed to generate the locomotor turning bias observed in *nanchung* flies on day 2 post-amputation, we supplied the average time between strides, proportion of time each leg is down per stride, and leg phases, as measured directly from the corresponding video recordings. Leg forces that yielded μ values matching experimental values were then determined over the course of recovery.

This approach was able to recapitulate the overall trajectory of fly locomotor behavior in response to injury (mean discrepancy < 1%) (Figure 4.17A). By tuning the leg forces, the strong turning bias induced immediately after amputation could be undone. Thus, modulation of force

appears to be sufficient to recapitulate turn bias recovery. As there was no recovery in its absence, the model suggests that force modulation may be necessary for turn bias recovery.

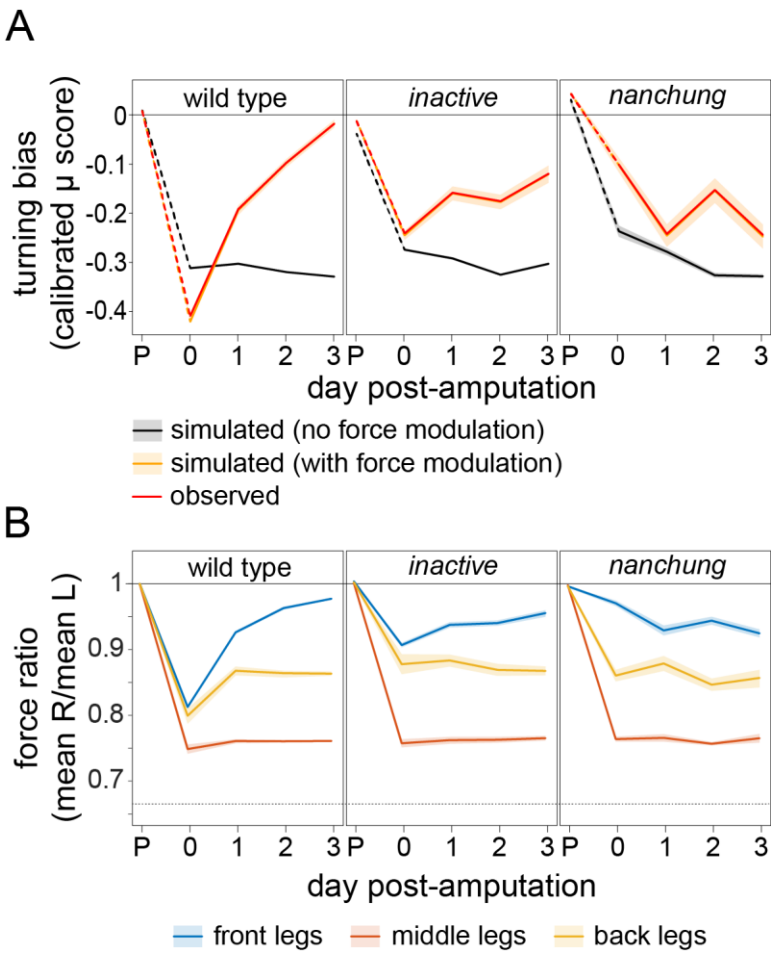

Figure 4.17. Model results show that tuning leg forces can capture the recovery profile. A. Turning bias (μ) determined from angular bias of simulated flies without and with force modulation using the turn bias-angular bias calibration curve. With force held constant (black lines) the model predicts that there would be no recovery of turn bias towards baseline from immediately post-amputation to day 3 post-amputation for all strains. With force tuning optimized to match observed recovery on a fly simulated with averaged parameters for each experimental group (orange lines), the experimental recovery profile (red lines as visual guide) can be recapitulated for all strains. Error bars are ±1 SEM. **B.** Force change required in each leg separately to fit the recovery profile. The dashed line is a visual guide indicating no force change. Means describe average total force exerted on the right side compared to the left side. The middle leg (red) requires the least change to fit a target recovery profile, followed by the hind (yellow) and front leg (blue). The recovery towards baseline is statistically significant for all legs in wild type ($P < 0.038$, F-test). In mutants, the slope is significantly positive only for the front leg of *inactive* ($P < 0.001$). Based on single-leg comparison, wild type flies tune force more than mutants over time to obtain a larger recovery effect.

Figure 4.17B shows the ratio of total leg forces on the right and left sides that yield simulated locomotion matching observed turn bias values, when the force parameter ($c_{i\in\{1,2,3\}}$) is swept linearly exclusively in one leg at a time (across all 5 legs, there are many combinations of force modulation that succeed.) The model suggests that the middle and hind legs require smaller force changes than the front leg to achieve a particular directional bias profile, and that this change is larger in wild type flies over the course of recovery. The recovery was significant for all legs in wild type ($P < 0.03$, F-test), though the modulation required in the middle legs to achieve turn bias recovery was markedly smaller than the front and hind legs, while among mutants, the recovery was only significant and positive for the front leg in *inactive* ($P < 0.001$), suggesting that wild type flies modulate force more significantly to achieve recovery.

Discussion

We see that wild type flies initially spend equal portions of time exploring in clockwise and counter-clockwise directions (Figure 4.13B). After amputation of the right foreleg, they exhibit a strong counter-clockwise bias. However, after three days, their behavior is largely unbiased. By contrast, *inactive* mutants recover approximately half-way from the maximum bias post-amputation and *nanchung* mutants do not recover at all. Moving from a behavioral assay to a gait analysis, we also considered the motion of individual legs. While the 3-leg gait never completely vanishes (either in the wild type or in the mutants), it is nonetheless marginalized starting immediately post-amputation and does not recover over three days even as turning bias does. Intriguingly, wild type flies appear to exhibit changing gait behavior post-amputation, with a significant increase in 2-leg gait and decrease in 1-leg gait, while the probability of using 3-leg gait remains unchanged. The probabilities of using 1-leg, 2-leg, and 3-leg gaits do not change significantly over time in mutants, though *inactive* appears

to favor a 2-leg gait immediately post-amputation while *nanchung* favors the 1-leg gait. This suggested that gait learning and recovery may be tied to proprioception.

The simple fact of a predominant 2-leg gait at the end of the experiment is unlikely to explain the observed recovery. After all, both wild type and *inactive* exhibit a similar predominant proportion of 2-leg gait three days after recovery (the difference is not significant; $P = 0.217$), and indeed the *inactive* mutants reach it first. Further, as we saw in the model, unbiased walking could not be achieved without force modulation; phase modulation alone was insufficient.

We then considered measures of several other parameters on a leg-by-leg basis, but found that they were insufficient to explain turn bias recovery. In some insects, it is known that different legs may play different roles in locomotion. For instance, in cockroaches, front legs are used more for steering and hind legs more for propulsion (Mu and Ritzmann, 2005). However, even for wild type flies, the difference in the remaining front leg distance from the centroid between 1 hour and 72 hours post amputation was <1%, suggesting that leg placement alone cannot account for the observed changes in walking direction during recovery. While several potential parameters of interests such as the average distance legs moved per stride and the proportion of time spent in the air during a stride are likely to be relevant, examining both individual leg and comparisons between the right and left side averages for all of these parameters did not yield a satisfactory explanation for the observed recovery of turning bias in walking; for instance, none showed the same qualitative pattern as Figure 4.13B. However, it remained a possibility that the small differences in these parameters could, in combination, explain turn bias recovery.

To test this, we developed a minimal Newtonian physical model for leg and body motion. After fitting all parameters from experiments, we were left with one tuning parameter: force. Holding force constant at pre-amputation levels yielded no turn bias recovery (Figure 4.17A). Tuning the forces exerted on each leg through a Monte Carlo optimization procedure to match the

average angular bias of flies within each experimental group, we found that we were able to recapitulate observed turning bias scores. Tuning force in the middle legs had the largest effect on turning bias. Therefore, the model suggests that force modulation appears to be necessary and sufficient to explain turn bias recovery, given the measured values of all other biophysical parameters and gait patterns. Our findings imply a space of leg force modulation solutions. Many combinations of force modulation across all 5 legs can balance average forces between the left and right sides of the animal, and flies likely change forces in all their legs as part of recovery. The front leg might be a special case due to having no contralateral leg to act against. It may be possible for a real fly to modulate the force from that leg with little constraint (e.g. by largely unloading it to become more "four-legged," thereby restoring symmetry).

This suggests that the coordination of forces exerted by each leg is a general mechanism which an animal can control to achieve unbiased walking. Second, this observation posits a fine-grained behavioral manifestation of proprioceptive defects. A number of studies have shown that deciphering forces and proprioceptive feedback are important in generating stable patterns/gaits (Pearson, 1972; Ridgel et al., 2000; Zill et al., 2004; Fuchs et al., 2012) (for a review from a modeling perspective see (Holmes et al., 2006); for a review with a more biological perspective, see (Delcomyn, 2004)) and may "directly influence [central pattern generators] and motoneurons to maintain phase relations in a decentralized, peripheral manner" (Holmes et al., 2006) through feedback. Proprioception has also been implicated in walking direction in e.g. stick insects (Akay et al., 2007). Equally important has been an exploration of the interplay of proprioception and recovery in motor control in various insects. For example, Page and Matheson considered locusts and found a shift in limb movements intended for scratching after a surgery-induced decrease in proprioception, followed by recovery to pre-surgery values over the course of a week (Page and Matheson, 2009); Buschges and Pearson discovered that the removal of wing proprioceptors in

locusts led to a decreased recovery of the flight motor pattern after wing injury (Buschges and Pearson, 1991; Buschges et al., 1992). Our study points to proprioception as a critical player in mediating orientation profile plasticity by determining how well an animal can control the individual forces it exerts. In other words, perhaps a proprioception-defective fly "wants" to exert more force on the right-hand side to counteract the effect of an amputated leg, but it cannot sense exactly how much force it is actually applying and is therefore doomed to continue making the same ineffectual exertions.

We note that the TRP channel mutants we considered have defects in various sensory structures, including all chordotonal organs across the body. The most relevant ones to this study are likely the legs, but it is possible other organs are involved, such as those between the abdominal segments. These possibilities could potentially be resolved using the *Drosophila* transgenic toolkit by, for example, using intersectional genetics to target *iav*- or *nan*-expressing neurons only in the leg. Inducible promoters could be used to compare the injury response of animals with inhibited chordotonal neurons to those with normal neuronal activity, while holding genotype constant. This would provide an advantage over the mutant approach, which might be confounded by other differences in genetic background.

Other proprioceptive organs than the chordotonal organs could be involved as well. For example, the campaniform sensilla (Zill et al., 2004) are known to measure force within the cuticle and could be part of post-injury force modulation. Chordotonal organs, by contrast, are generally considered to be stretch rather than force sensors, but if the nervous system encodes the mass of the animal, the information encoded by a dedicated stretch sensor could be used to compute force. Specifically, stretch-sensitive neurons that encode position, could stimulate a sequence of high-pass filtered (i.e. rapidly adapting) downstream neurons, which can readily compute signal derivatives. Multiplying the activity of these downstream neurons by the encoded mass value would produce a

neural code for force downstream of chordotonal organs. More directly, it has been experimentally observed in stick insects that the afferent projections of different proprioceptive organs (including the femoral chordotonal organs and campaniform sensilla) can interact by exerting presynaptic inhibition on each other (Stein and Schmitz, 1999). This phenomenon appears to be conserved in the Pancrustacea as the chordotonal neurons of crayfish impart presynaptic afferent depolarization on sensory neurons innervating touch-sensitive bristles on swimming limbs, but only at speeds matching those of locomotion (Newland et al., 1996). Thus, even if the chordotonal dendrites encode only position, the chordotonal neurons could encode force by virtue of their interaction with campaniform neurons.

Conclusion

This study points to a number of avenues for future work. One important direction is a closer experimental analysis of the effects of the proprioceptive defects on locomotion. A natural question is "how much does each part of the neuronal circuit lead to recovery failure?" In this context, one could consider the effect of *stum*, which is critical for transduction of mechanical stimuli in a subpopulation of proprioceptive neurons responsible for sensing joint angles (Desai et al., 2014) and *nompC*, which is required to have virtually any mechanosensory signaling such as a response to changing joint angles (Chadha et al., 2015). Another direction is performing similar experiments e.g. in larger hexapods (where forces are easier to measure) or tetrapods. It may also be interesting to better characterize coordination patterns (gaits) in animals after surgery.

From a modeling perspective, an interesting extension would be to define a neural network with dynamic connections between neuronal modules in place of a fixed phase, duration, and force. That is, one could treat the neuronal modules as oscillators which are explicitly coupled to each other and self-tune their phase or amplitude. Then, one could ask what simple rules could allow the

system to self-tune for recovery after injury (for the case of adaptive networks in a coupled oscillator system, see e.g. (Aoki and Aoyagi, 2011; Isakov and Mahadevan, 2014)). Another extension would be to incorporate "reflexes and preflexes" (Kukillaya et al., 2009; Proctor and Holmes, 2010) to understand what role these play in recovery. Finally, we can ask whether a simple force-balance rule can be used in robots such as those suggested in (Schilling et al., 2013; Cully et al., 2015), thereby encouraging "robotic recovery from injury" and allowing better performance in the field.

In the next chapter, we transition to the human scale and study collective behavior mediated by a social, rather than a physical, network. There, we study data from a society of agro-pastoralists in Africa to understand how social ties and individual characteristics such as leadership determine group formation for stealth raids (a form of collective action).

"For a friend with an understanding heart is worth no less than a brother."
- Homer, *Odyssey*

5

Leadership and Friendship in Intergroup Violence

Introduction

Intergroup violence is widespread in social animals, including humans. Yet, most studies of intergroup violence in small-scale populations have focused on the mortality rate and demographic effects of warfare, not the social precursors (Beckerman et al., 2009; Chagnon, 1988; Hill et al., 2007; Wiessner and Pupu, 2012). However, some prior research indicates that social organization may be an important factor in individual participation. Among the Yanomamö horticulturalists of Venezuela and Brazil, men who participated together in a killing during a raid were likely to live together and exchange marriage partners later in life (Macfarlan et al., 2014). This suggests that humans can use intergroup conflict strategically to advance subsequent relationships among participants. However, because this study did not analyze data on raiding party composition or social networks, inferences about the social processes contributing to conflict events are limited.

Here, we explore the role of individual differences and social networks in the emergence of collective violence among the Nyangatom, a group of nomadic pastoralists inhabiting a remote region along the border of South Sudan and Ethiopia (Glowacki and Wrangham, 2015). Intergroup

conflict among the Nyangatom is decentralized, occurring outside of formal institutions such as military units or governments, and resembles the dynamics of simple warfare in other small-scale societies.

Many Nyangatom live in mobile cattle camps containing between 10 and 100 persons, and the population and number of these camps are not fixed. Depending in part on the season, camps may disband with residents forming new camps, or they may aggregate and form larger groupings. The Nyangatom also have semi-permanent villages with dynamic membership, and movement between camps and villages is common. Livestock have a central place in the culture and diet of the Nyangatom and are necessary for many social exchanges, including marriage. In order to marry, a male is required to provide the family of the bride with bridewealth, often 30 to 60 cattle but sometimes as many as 100 cattle. Therefore, livestock are highly sought after, and violent conflict with other groups in order to obtain them is common (Mathew and Boyd, 2011; Glowacki and Wrangham, 2015).

The most common type of intergroup conflict event for the Nyangatom is the "stealth raid" (sing. *Emojirimónu)*, which is usually conducted by a small group of young to middle-aged men who usually attempt to seize livestock and attack enemies from other nearby ethnic groups and then escape unharmed. Similar to warfare among hunter-gatherers, stealth raids sometimes yield fatalities for members of enemy groups, but casualties among the raiding party are unusual since raiders seek low-risk opportunities to attack. In addition to livestock, successful raiders receive also social benefits, such as status and public praise, and men who have killed an enemy commonly receive a warrior name and honorific scars. The dynamics of Nyangatom violence against outgroups is similar to the processes of raids in other small-scale societies (Mathew and Boyd, 2011; Wrangham and Glowacki, 2012).

Stealth raids generally begin with one or two individuals who take a leadership role and attempt to recruit other participants, a process that can take several days. Raiding parties can also emerge when large groups of young men are congregated, such as during a ceremony. Individuals are not compelled to join a raiding party; many young men elect not to join and there are no formal sanctions for cowardice, desertion, or failure to participate (Glowacki and Wrangham, 2015). Leaders often take a more active role in the raid, such as acting as a scout, selecting the site of the raid, or developing tactics; they may also direct the division of loot after a successful raid.

We collected conflict histories and data on raiding co-participation over a three-year period among all age-appropriate Nyangatom men residing in the study area (N = 91). We also mapped the social networks of this population at one point in time using a gift allocation task where participants were asked to make anonymous allocations to other men. Together, these data allow us to explore the role of individual variation and the influence of social networks in the formation of raiding parties engaged in risky collective action.

Materials and methods

Our data were collected as part of an ongoing ethnographic study of the Nyangatom in which one of the researchers (Luke Glowacki) resided in the study area intermittently between 2009 and 2012. We used semi-structured interviews to collect information regarding intergroup conflict events that occurred between the Nyangatom and their neighbors, including the Turkana, Dassanetch, and Suri, and we established that 39 stealth raids were initiated by the study population in the area of research between June 2009 and December 2012. We identified 91 men residing in the study area who were of the appropriate age to participate in stealth raids (approximately 18 to 45 years old). These 91 men form the study population and we conducted interviews with each of these individuals collecting data on their conflict history and identifying co-participants.

Co-participants were identified by name or photos and composition was validated through peer reports. The presence of a raider on a raiding party was determined by an individual's participation in the raiding party for any portion of it; we did not measure desertion, and some individuals may have ceased their participation during the actual raid because they were afraid or for other reasons. Leadership was ascertained by cross-validated personal accounts elicited by questions about whether any person was a leader of the raid using two Nyangatom terms for leader (sing. *Ekarikon*; sing. *Eketamunan*).

We also performed a comprehensive, sociocentric network study of the entire population of raiding-age Nyangatom males (N = 91). In order to measure friendship ties within this group, we employed a gift task modeled on prior work with the Hadza hunter-gathers of Tanzania (Apicella et al., 2012) in which the Nyangatom subjects were asked to identify other study participants to whom they would like to give a gift of candy. Giving a gift is an important measure of friendship in most societies. We chose candy as the allocation currency because of its practical ease and because Nyangatom value it. Subjects were presented with three pieces of candy and shown photo sheets containing the facial portraits of study participants to whom an *anonymous* allocation could be given. They were asked to indicate the three persons that they would like to receive the gift of candy and told they would not be identified as the donor (and recipients did not know who gave them the gift either). All 91 subjects (100%) participated, yielding a total of 273 social ties within this group and distributions occurred only after all participants completed the task. We also measured a variety of attributes of the study participants including height, weight, and estimates of paternal wealth.

In order to explore associations between raid characteristics and raid participation, we evaluated linear regression models that estimated the association between an individual's decision to join a particular raid and various raid characteristics. The basic model is

$$E[Y_{ir}] = \theta_i + \gamma_r + \beta x_{ir}, \quad (5.1)$$

where the dependent variable $Y_{ir} = 1$ if person i joins raids r, and $Y_{ir} = 0$ otherwise, x_{ir} is a vector of characteristics for participant i and raid r, β is a vector of coefficients that indicate the degree of association with each characteristic, and θ_i and γ_r are individual and raid fixed effects, respectively. We report results of the linear model for more intuitive interpretation. The results are consistent in both sign and magnitude when compared to Generalized Linear Models (GLMs). See Appendix A for a description of supplemental methods.

Results

A total of 39 raids occurred during the study period, averaging approximately one every 5 weeks. Each raiding foray typically lasted several days and raids were generally non-overlapping in time. Raiding parties averaged 6.8 participants (SD = 3.4) (Figure 5.1A) and 86% of the men (78/91) participated in at least one raid (mean = 2.9 raids, SD = 3.3), but 14% of the population did not participate in any raids. Roughly 80% of the raids resulted in the capture of livestock, and raiders averaged 4 head per successful raid (including cattle, goats, and donkeys).

We measured a variety of attributes of the study population (N = 91) including height (mean = 175.4 cm, SD = 6.4), weight (mean = 58.5 kg, SD = 6.5), paternal wealth, and siblings in the sample (mean = 1 male sibling, SD = 1.3). Measured individual-level characteristics were associated with raid participation in bivariate models without controls – including the number of siblings, height, weight, and the number of livestock owned by the raider's father (paternal wealth). However, none of these egocentric characteristics remain significant in multivariate models and they do not appear independently relevant to whether a person went on a raid.

Five of the 91 men in the study were identifed as being a raid leader, and all raids included at least one of these leaders (see Figure 5.1B which shows a "bipartite" network of the 39 raids, with

the five individuals identified as leaders on any raid marked in red). We conducted a minimal set analysis to independently identify leaders using raid participation data alone. This analysis attempted to identify the smallest group of individuals at least one of whom participated in every raid. The procedure successfully identified all five individuals who were named as leaders on any raid by participants themselves. These five individuals also participated in significantly more raids than expected due to chance (all $P \ll 0.001$), as determined by a procedure where we randomly re-assign the identities of those who participated in each raid while keeping the distribution of raid participation fixed. These results are consistent with emic accounts indicating that leaders have an important role in the initiation and formation of raiding parties.

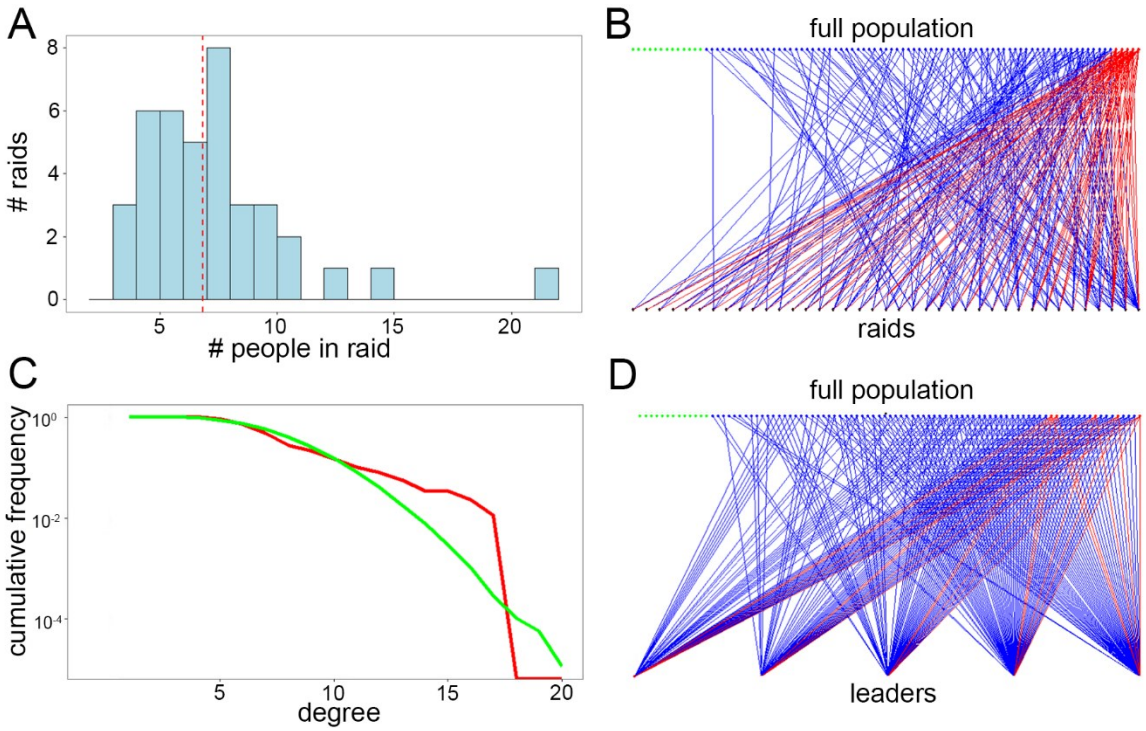

Figure 5.1. Overview of data on raid participation and the social network. A. Distribution of stealth raid sizes over a total of 39 raids. The dashed red line shows the mean. **B.** Bipartite participant-raid network. Top nodes represent people; bottom nodes represent raids. Red nodes denote leaders. Each raid has at least one leader. Red lines indicate connections of leaders to raids, and blue lines indicate connections of everyone else. Green dots are individuals who do not raid. **C.** Degree distribution (cumulative frequency) of the friendship network (red) and the average distribution of 10^3 random networks with the same number of nodes and edges (green). The real distribution does not differ significantly from a random graph. **D.** Participant-leader raid relation

Figure 5.1 (Continued) network. A leader (bottom) is connected to an alter if the alter went on a raid with the leader. Red dots are leaders and green dots are individuals who do not raid. Red lines denote leaders co-raiding with other leader alters.

We mapped the social network of all raiding-age Nyangatom men using a task where participants made anonymous gift allocations to other subjects (Figure 5.2). The mean number of incoming friendship nominations (in-degree) was 3 (SD = 2.7), and the range was 0 to 13. Although in-degree is associated with both paternal wealth and number of siblings, the strongest predictor of the number of friendship nominations is leadership status. Leaders have significantly more friends than non-leaders (5.2 vs. 2.4, P = 0.01). Leaders also score significantly higher on a measure of network centrality, even when controlling for in-degree (P = 0.04). This suggests that leaders not only have more friends, but also that their friends tend to be more popular, meaning leaders also have more friends of friends as well (as is visible in Figure 5.2). Figure 5.3 combines both gift-giving and co-raiding ties into one diagram.

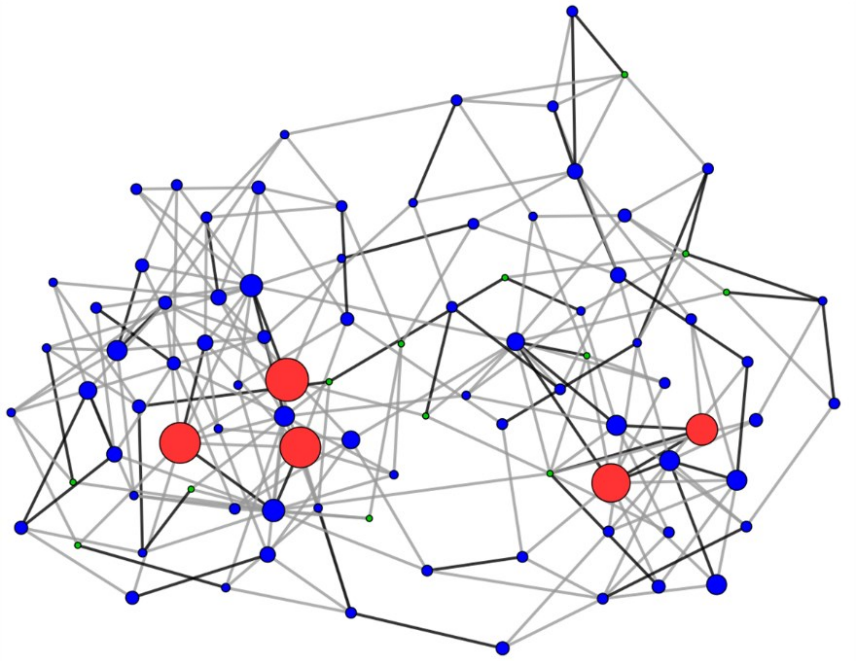

Figure 5.2. Nyangatom social network. Network of friendship ties in Nyangatom society determined using gift allocation task. Those who did not participate in any raids (non-participants) are shown in green, those who participated in at least one raid (participants) are shown in blue, and

Figure 5.2 (Continued) identified leaders are shown in red. Node size is proportional to raid participation (number of raids in which an individual participated). Dark gray arrows indicate reciprocal, two-way friendship ties, and light gray arrows are one-way ties.

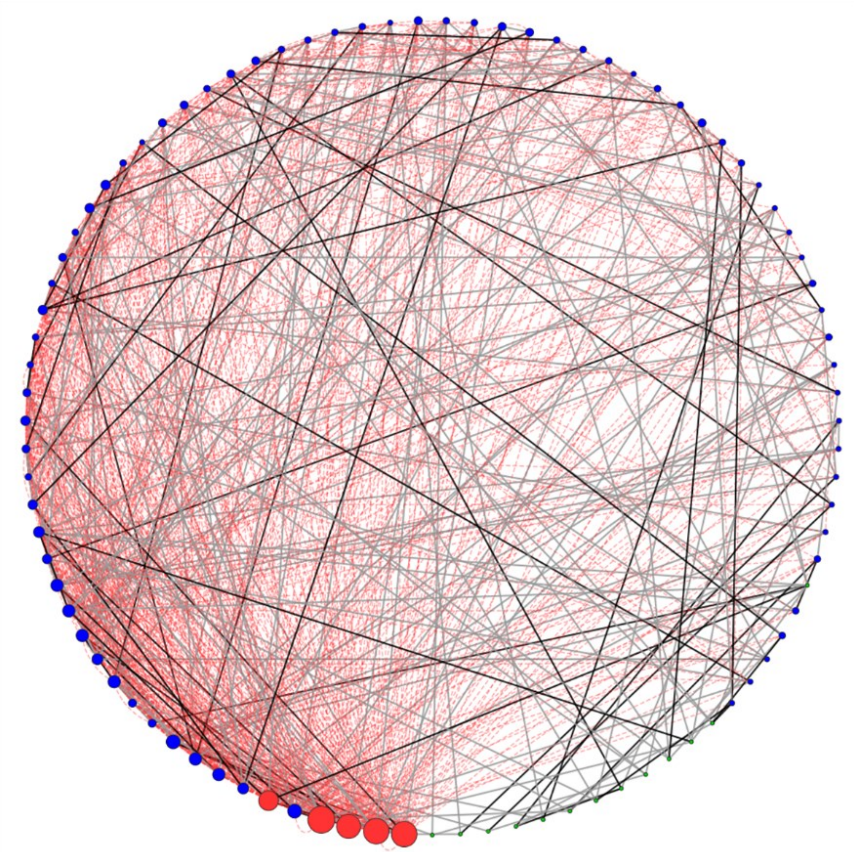

Figure 5.3. Nyangatom social network with co-raiding ties. Network of friendship ties in Nyangatom society determined using gift allocation task. Those who did not participate in any raids (non-participants) are shown in green, those who participated in at least one raid (participants) are shown in blue, and identified leaders are shown in red. Node size is proportional to raid participation (number of raids in which an individual participated). Dark gray arrows indicate reciprocal, two-way friendship ties, and light gray arrows are one-way ties. Dashed red lines indicate co-raiding.

We explored the ways the Nyangatom social network is similar to certain other social networks by measuring a comprehensive set of statistics (Apicella et al., 2012). While the cumulative degree distribution (Figure 5.1C) does not differ significantly from an Erdos-Renyi random network with the same number of vertices and edges ($P = 0.76$), a number of other important properties are shared with non-random social networks. Reciprocity (the probability that participant B names

participant A as a friend given that participant A names participant B) is significantly higher in the Nyangatom network (reciprocity = 0.37) than in a random network ($P \ll 0.001$); that is, there are significantly more two-way friendship ties in the real network than a random network. Transitivity (the probability that two of a participant's friends are friends with one another) is also significantly higher in the Nyangatom network than a random network (0.17, $P < 0.001$). Finally, there is also stronger homophily (the tendency of people with similar characteristics to have social ties with one another) by age (0.88, $P < 0.001$) and by degree (0.10, $P = 0.04$) compared to a random network.

Social network structure is relevant to the composition of raiding parties. Individuals with more social connections (i.e., higher degree) tend to go on more raids, even when we exclude leaders from the analysis (Figure 5.4A). Each additional social connection is associated with an increase of 0.45 raids (SE = 0.17, $P = 0.01$) in the expected number of raids in which a subject participates. Regression models that include in-degree, height, weight, wealth, and number of siblings show that social relationship "capital" is more strongly associated with raid participation than physical or material capital. Our measure of network in-degree is the only variable that survives various model specifications.

However, the emergence and composition of particular raiding parties is more nuanced than leaders simply being linked by friendship ties to other raid participants. We used regression analysis to evaluate the decision to join an individual raid, examining how this decision is associated with the total number of other people who join the raid, the number of one's friends who participate, and the number of other leaders who participate. In these models, we treated each individual's decision to participate in each raid as the dependent variable, and we assessed how the presence of other participants was associated with the probability that a person would join the raid. To control for unobserved characteristics of individuals (e.g., their attitudes towards violence or risk, as well as

other personality factors) and of raids (e.g., the distance to the raid target or the anticipated value of the raided items), we included fixed effects for both individuals and raids in the model.

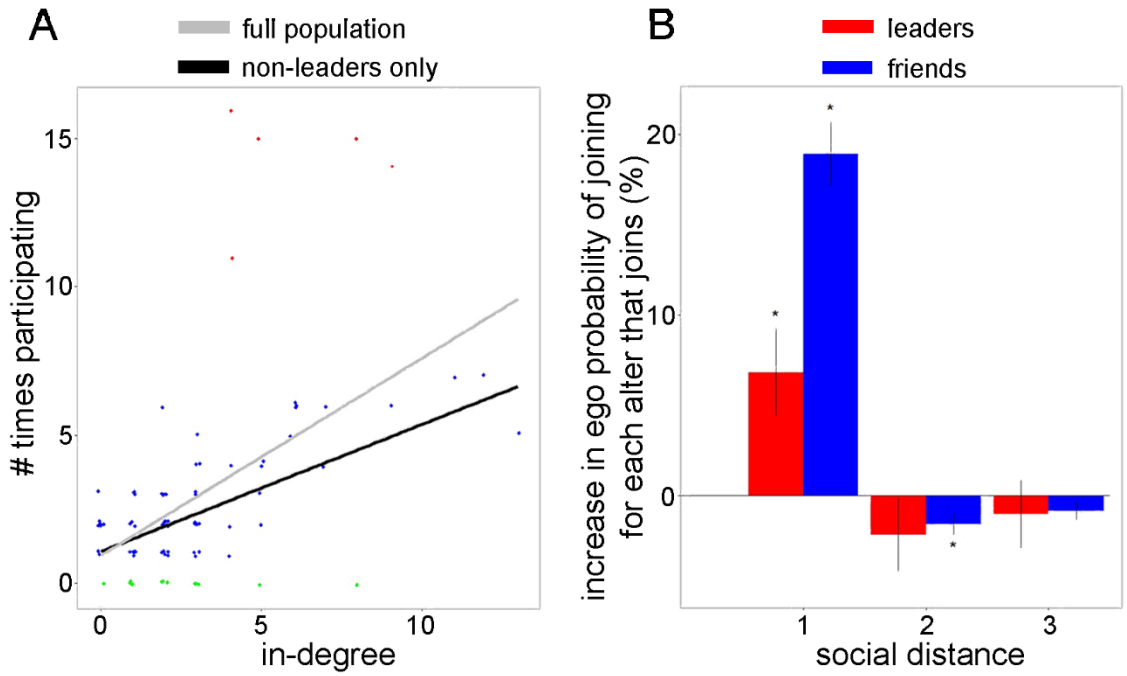

Figure 5.4. Probability of joining a raid depends on social ties. A. Number of times people joined raids as a function of social in-degree. Regression lines are shown for the full population (gray) ($R^2 = 0.32$) and excluding the leaders (black) ($R^2 = 0.42$). In all panels, people who participated in no raids are green, non-leader participants are blue, and leaders are red. **B.** Increase in probability of joining a raid based on geodesic social distance to leaders and to non-leader friends. Lines denote one standard error of the estimate. The large positive coefficients on first-degree connections show that direct non-leader friends are more motivating than leader friends, and both are significant. The negative coefficient on second-degree connections provides evidence against cascade formation beyond 1 degree in raiding party formation. Motivation did not spread significantly to third degree friends.

The presence of both leaders and friends was significant for the composition of the raiding party. Subjects were more likely to join a raid if they had a direct friendship to the leader of that raid (6.8%, SE = 2.4%). There was no effect for being friends of friends with the leader (social distance 2) or friends of friends of friends (social distance 3) on the probability of joining a raid (Figure 5.4B). This suggests that leaders mobilize their direct friendship contacts to join raids. However, leaders are less effective at mobilizing their direct friends than non-leaders. Each *non-leader* friend who participated in a raid increased the likelihood that the subject joined the raiding party (19.2%, SE =

1.4%). The effect of being friends with a co-raider who is not a leader on joining a particular raid is significantly higher than the boost in probability associated with being friends of a leader of a particular raid ($P \ll 0.001$). We found no evidence that kinship influences raiding party composition because siblings were not more (or less) likely than chance to raid with each other ($P = 0.23$).

We tested whether leaders, who are better connected to the network, matter more than non-leaders for overall raid composition. Using a model where we regress total participation by a person's friends on a person's decision to join, their leadership status, and an interaction variable that indicates the effect of leadership on total mobilization, we find that significantly more people join the raid when a non-leader joins than compared to a leader ($P = 0.008$). In other words, the key motivating factor to join a raid once a raid is initiated is not the presence of specific leaders but rather the presence of a non-leader friend who joined the raid.

Discussion

Here, we have analyzed co-participation in group-level conflict in a small-scale society that is relevant to understanding how risky collective action problems involving violence in humans are solved. We find that there is significant inter-individual variation in raiding participation. Most men participated in at least one raid, but 14% did not participate in any raids and five individuals participated in more than ten raids. The initiation of raiding parties appears to depend on the presence of key individuals (leaders) who function as nucleation sites for raids and attract other participants. Leaders may alter the costs and benefits for others either by reducing the costs of the raid (recruitment and scouting) to other participants or by exerting social pressure on others to join (King et al., 2009; Glowacki and von Rueden, 2015).

We found no effect of kinship in raiding party composition, suggesting alternative mechanisms for participation in conflict besides kin selection. This is similar to behavioral data on

collective action in chimpanzees engaged in collective hunting and border patrols where there is no indication that kinship influences the likelihood or effectiveness of such collective action (Langergraber et al., 2007).

The most important determinate of raiding party composition was social network structure, even when controlling for individual differences in the tendency to join raids. Leaders recruit co-raiders from both within and outside their social networks. Because leaders have larger social networks than non-leaders, they have a larger possible pool of co-raiders to draw from. However, we found the surprising result that co-raiders who are not leaders are more effective than leaders at recruiting on a per-individual basis: individuals are more likely to join raids when they are direct friends of a raider who is not the leader. This suggests that once a leader recruits co-raiders, these co-raiders in turn have a critical role in the composition of raids by encouraging members of their social network to join that the leader may not be connected to.

Our findings are especially notable because we did not uncover reports of formal sanctioning for non-participants although the withdrawal of a friendship tie may be a form of sanction (Rand et al., 2011). This is consistent with our finding that, overall in our population, and excluding leaders, the number of friendship ties is associated with raiding participation. Work in another setting suggests that co-participation in violence may result in subsequent formation of social bonds (Macfarlan et al., 2014), and men have also been observed to enlist in the army with their friends, or even been encouraged to do so, in state-sponsored wars (Viterna, 2006).

We are unable to exclude the possibility that friendships formed as a result of participating in a raid together, rather than friendships driving co-participation. However, we think friendship is a primary mechanism that contributes to co-participation rather than being a consequence of co-participation. Among East African pastoralist societies, young men engage in many collective activities together, such as herding and ceremonies, creating opportunities to meet other members of

their age group. As a result, they form very tightly bound cliques early in adolescence that are an important part of the social life of Nyangatom society. Because raids are risky and commonly accompanied by nervousness and fear, we expect that individuals would choose to raid with their friends, rather than people they are not well acquainted with. Rather than acting as a mechanism to generate friendships, raids may instead act to solidify and deepen friendships. This expectation is supported by our data on differential recruitment between leaders and non-leaders. If friendships that resulted from co-raiding were the primary determinate of our network data, leaders would be expected to recruit primarily from their social network. Instead, we find leaders co-raid with many individuals outside their immediate social network. Yet, non-leaders are more likely to raid with persons they are friends with, even if they are also friends with a leader of a raid. Nonetheless, we cannot exclude the possibility that some friendships formed as result of co-raiding; future research should include repeated measures of network structure to ascertain the extent to which collective violence directed at out-groups shapes the social network of the in-group.

Our results are consistent with research in other populations on the importance of networks in the promotion of violence. Friendships are a well-known predictor of joining state-sponsored warfare (Viterna, 2006). A recent analysis of participants in the Rwandan genocide showed that perpetrators had larger networks and were more likely to have connections to other perpetrators than non-perpetrators (McDoom, 2013). Similarly, in the Paris Commune, preexisting networks were important in supporting revolutionary mobilization (Gould, 1991).

The importance of key individuals in the initiation of raiding parties is consistent with a growing body of theoretical and empirical research on the role of individual variation in collective action. Modeling work shows that within-group heterogeneity may affect group performance (Gavrilets and Fortunato, 2014; Molleman et al., 2014, Gavrilets, 2015; Henrich et al., 2015), which suggests an important role for key individuals who may alter the costs and benefits of others so that

costly cooperation is favored. Empirical work in non-humans demonstrates the role of individual differences in promoting collective action. For example, among wild chimpanzees, a few individuals act as "impact males" whose presence catalyzes the group to engage in high-risk activity such as hunting or conducting boundary patrols (Gilby et al., 2008; Gilby et al., 2015). Impact males are more likely to initiate group hunts and the removal of an impact male results in lower hunting rates for the group (Gilby et al., 2015). Among captive chimpanzees, the success of solving more complex collective action problems is also sensitive to the presence of particular individuals in the group (Schneider et al., 2012).

Other social and group-living species have similar dynamics. Ring-tailed lemurs, for instance, have extensive inter-individual variation in their tendency to participate in territorial conflicts, with a few individuals initiating the majority of conflict (Nunn and Deaner, 2004), while among black howler monkeys, social groups are usually led by one "habitual leader" who initiates group movements and has the highest centrality among group members (van Belle et al., 2013). Similarly, collective group movement among both bottlenose dolphins (Lewis et al., 2011) and killer whales (Brent et al., 2015) commonly occurs with specific individuals functioning as leaders of the group during movement. Finally, experimentally manipulating spider colonies through the introduction of just a single bold individual produced pronounced group-level effects by increasing aggressive group foraging (Pruitt and Pinter-Wollman, 2015). These studies, combined with the data we present here, suggest that risky collective action may be sensitive to the presence of individuals who initiate group activity – though the mechanism by which this occurs is still not fully understood (Gavrilets, 2015; McAuliffe et al., 2015), and the social dynamics by which such actions spread within a group have not been explored using detailed network models.

Conclusion

In sum, we find that key individuals functioning as leaders are important for raid initiation. However, social network structure matters substantially to the actual composition of raiding party. While leaders recruit from within and outside their social network, the key factor for group composition appears to the presence of a co-raider that a raider is connected to via a direct friendship tie. To the extent that Nyangatom raiding behavior is relevant to other domains of collective action in humans and other primates, we have identified an important amplifying effect: a handful of motivated individuals, with distinctive network positions, coupled with a wider group of reinforcing individuals embedded within a network, can lead to population-level risky and violent behavior directed at other groups.

One direction for future research is studying the relevance of these results in other types of non-institutionalized violence, such as urban gangs (Papachristos, 2009), localized insurgencies (Viterna, 2006; Johnson et al., 2011), or revolutionary protests (Gould, 1991). Many types of violence do not depend solely on the desires and actions of individuals or even dyads, and instead may at least partially emerge and be supported by the very social structure in which all individuals are embedded.

These observations suggest two things with respect to the prospect of managing violence. On the positive side, attenuating the impact of a leader may prevent the original nucleation of the violence. Yet, on the negative side, once violence starts, people are likely to join from throughout the whole population, and so, once instigated, violence has a wide-reaching effect on the society. Further studies could be combined with modeling to address the question of whether it is possible to encourage peace through targeted interventions (Isakov et al., 2016). A related area that would be amenable to both modeling and observation (e.g. through extended longitudinal study) is understanding the emergence of leaders over time. That is, as a new generation is growing up, what

role does social structure play in determining who will become an impact individual, and how does that itself affect the formation of the social network?

"Knowledge of the fact differs from knowledge of the reason for the fact."
- Aristotle, *Posterior Analytics*

6

Conclusion

In this work, we have examined collective dynamics as they arise in a spectrum of systems across varying levels of agent complexity and agent-network interactions. We began with looking at one of the simplest physical models of interacting units, coupled phase oscillators, on a relatively simple set of networks (1-dimension rings, 2-dimensional lattices, and a modified scale-free model). Adding in a mechanism to make the underlying network dynamic (that is, allowing the strength of inter-agent connections to vary) by coupling it to the actions of the oscillators themselves, we characterized distinct regimes as a function of noise at the oscillator level and the time-scale, including those that can sustain oscillator flipping between aligned and anti-aligned states. We also characterized how these regimes depend on the structure of the underlying topology, namely on the average degree of the network.

Then, we built a model of a system of evolving agents where the agents' strategies directly define the network structure. The rules themselves are very simple: an agent can either take resources, give resources, or be disconnected from another agent, an individuals' fitness is proportional to the net resource intake, and the fittest individuals reproduce their relationships (up to mutation). Nonetheless, even with this small set of parameters, the network evolution was

remarkably complex – depending on the ratio of benefits to costs and on reproduction and mutation rates, the long-term population structure could be noise-dominated (random edges), collapse completely, or form a number of stable communities. Interestingly, starting with stable communities, we found that an intervention that disrupts cooperation in a small proportion of agents can lead to wide-spread population collapse, and the magnitude of this effect was highly non-linear in the disruptive power of the intervention. Even with just 2% of agents being susceptible to disruption, and with approximately half of their outgoing ties blocked, the probability of the whole population being unable to spontaneously recover was over 50%. This provides additional insight into the possibility of using evolution to e.g. help treat "cooperative bacteria" that are currently hard to treat with antibiotics.

After studying simple agents, we moved towards systems where the individual rules are significantly more complex. First, we looked at the locomotor system of fruit flies. Initially, adult flies spend approximately equal time walking clockwise and counterclockwise. Immediately after the amputation of the front right leg, flies display a striking turn bias in the counterclockwise direction. But, while wild type flies can recover to an unbiased turning state within three days after the injury, those with genetic mutations that disrupt proprioception do not show this significant recovery. Through a combination of experiments and modeling, we examined fly behavior both on the level of the whole fly and on the level of individual legs (6 units initially, then 5 after amputation). Of course, the units (legs) work together to produce the observed behavior of the fly body. We described how many walking parameters change over time in both wild type and mutant flies, and found that proprioception (feedback about leg stretch) plays an important role in allowing the force tuning necessary to achieve unbiased behavior.

Finally, we considered a larger network with very complex agents: a social network in a traditional population in Africa, the Nyangatom. We analyzed the problem of collective decision-

making in the context of raiding: what influences the composition of a group that will go on a stealth raid? There, we uncovered that both a leadership quality and friendship ties play significant and distinct roles. Leaders are necessary to initiate raids, but the final group structure is primarily determined by friendship ties. Practically, further research can lead to developing and studying the effect of interventions in promoting positive behavioral change at scale. It would likewise be interesting to understand the dynamics of how leadership arises over time (that is, have the network and the agents be dynamic).

Overall, the major theme of these studies and the proposed future work is both simple and elegant. It is well-known that having a collection of individuals that talk to each other, from simple oscillators changing their phases to be in tune with their neighbors to humans gathering to participate in collective action, leads to a range of phenomena. At the same time, it is important to understand that the effects of talking to each other can in turn change the way the agents interact, creating a feedback loop. That is, the underlying network structure of interactions influences, and is in turn influenced by, the interactions themselves. This way of thinking leads to substantially more realistic models that hope to explain the fascinating wealth of behaviors observed in real life, from the formation of the neural connections to the social organizations that define us.

Bibliography

Abrams, D. M. and Strogatz, S. H. (2004). Chimera states for coupled oscillators. *Phys. Rev. Lett.* **93**(17), 174102. DOI: 10.1103/PhysRevLett.93.174102

Abrams, D. M., Mirollo, R., Strogatz, S. H. and Wiley, D. A. (2008). Solvable model for chimera states of coupled oscillators. *Phys. Rev. Lett.* **101**(8), 084103. DOI: 10.1103/PhysRevLett.101.084103

Acebrón, J. A., Bonilla, L. L., Vicente, C. J. P., Ritort, F. and Spigler, R. (2005). The Kuramoto model: a simple paradigm for synchronization phenomena. *Rev. Mod. Phys.* **77**(1), 137-185. DOI: 10.1103/RevModPhys.77.137

Akay, T., Ludwar, B. C., Goritz, M. L., Schmitz, J. and Buschges, A. (2007). Segment specificity of load signal processing depends on walking direction in the stick insect leg muscle control system. *J. Neurosci.* **27**(12), 3285-3294. DOI: 10.1523/jneurosci.5202-06.2007

Akay, T., Tourtellotte, W. G., Arber, S. and Jessell, T. M. (2014). Degradation of mouse locomotor pattern in the absence of proprioceptive sensory feedback. *Proc. Natl. Acad. Sci.* **111**(47), 16877-16882. DOI: 10.1073/pnas.1419045111

Alexander, R. M. (1984). The gaits of bipedal and quadrupedal animals. *Int. J. Robot. Res.* **3**(2), 49-59. DOI: 10.1177/027836498400300205

Aoki, T. and Aoyagi, T. (2009). Co-evolution of phases and connection strengths in a network of phase oscillators. *Phys. Rev. Lett.* **102**(3), 034101. DOI: 10.1103/PhysRevLett.102.034101

Aoki, T. and Aoyagi, T. (2011). Self-organized network of phase oscillators coupled by activity-dependent interactions. *Phys. Rev. E* **84**(6), 066109. DOI: 10.1103/PhysRevE.84.066109

Apicella, C. L., Marlowe, F. W., Fowler, J. H. and Christakis, N. A. (2012). Social networks and cooperation in hunter-gatherers. *Nature* **481**(7382), 497-501. DOI: 10.1038/nature10736

Archetti, M and Scheuring, I. (2012). Review: game theory of public goods in one-shot social dilemmas without assortment. *J. Theor. Biol.* **299**, 9-20. DOI: 10.1016/j.jtbi.2011.06.018

Axelrod, R., Axelrod. D. and Pienta, K. (2006). Evolution of cooperation among tumor cells. *Proc. Natl. Acad. Sci. USA* **103**(36), 13474-13479. DOI: 10.1073/pnas.0606053103

Ballermann, M. and Fouad, K. (2006). Spontaneous locomotor recovery in spinal cord injured rats is accompanied by anatomical plasticity of reticulospinal fibers. *Eur. J. Neurosci.* **23**(8), 1988-1996. DOI: 10.1111/j.1460-9568.2006.04726.x

Barabási, A.-L. and Albert, R. (1999). Emergence of scaling in random networks. *Science* **286**(5439), 509-512. DOI: 10.1126/science.286.5439.509

Barta, Z. (2016). Individual variation behind the evolution of cooperation. *Philos. Trans. R. Soc. B Biol. Sci.* **371**(1687), 20150087. DOI: 10.1098/rstb.2015.0087

Bässler, U. (1977). Sensory control of leg movement in the stick insect *Carausius morosus*. *Biol. Cybern.* **25**(2), 61-72. DOI: 10.1007/BF00337264

Bässler, U., Wolf, H. and Stein, W. (2007). Functional recovery following manipulation of muscle and sense organs in the stick insect leg. *J. Comp. Physiol. A* **193**(11), 1151-1168. DOI: 10.1007/s00359-007-0268-0

Baum, L. E. and Petrie, T. (1966). Statistical inference for probabilistic functions of finite state Markov chains. *Ann. Math. Statist.* **37**(6), 1554-1563. DOI: 10.1214/aoms/1177699147

Beckerman, S., Erickson, P. I., Yost, J., Regalado, J., Jaramillo, L., Sparks, C., Iromenga, M. et al. (2009). Life histories, blood revenge, and reproductive success among the Waorani of Ecuador. *Proc. Natl. Acad. Sci. USA* **106**(20), 8134-8139. DOI: 10.1073/pnas.0901431106

Berman, G. J., Choi, D. M., Bialek, W. and Shaevitz, J. W. (2014). Mapping the stereotyped behaviour of freely moving fruit flies. *J. R. Soc. Interface* **11**(99), 20140672. DOI: 10.1098/rsif.2014.0672

Birn-Jeffery, A. V., Hubicki, C. M., Blum, Y., Renjewski, D., Hurst, J. W. and Daley, M. A. (2014). Don't break a leg: running birds from quail to ostrich prioritise leg safety and economy on uneven terrain. *J. Exp. Biol.* **217**(21), 3786-3796. DOI: 10.1242/jeb.102640

Blichkan, R. and Full, R. J. (1987). Locomotion energetics of the ghost crab: II. Mechanics of the centre of mass during walking and running. *J. Exp. Biol.* **130**(1), 155-174.

Borgmann, A. and Buschges, A. (2015). Insect motor control: methodological advances, descending control and inter-leg coordination on the move. *Curr. Opin. Neurobiol.* **33**, 8-15. DOI: 10.1016/j.conb.2014.12.010

Borgmann, A., Hooper, S. L. and Buschges, A. (2009). Sensory feedback induced by front-leg stepping entrains the activity of central pattern generators in caudal segments of the stick insect walking system. *J. Neurosci.* **29**(9), 2972-2983. DOI: 10.1523/JNEUROSCI.3155-08.2009

Bosse, M. J., MacKenzie, E. J., Kellam, J. F., Burgess, A. R., Webb, L. X., Swiontkowski, M. F., Sanders, R. W., Jones, A. L., McAndrew, M. P., Patterson, B. M. et al. (2002). An analysis of outcomes of reconstruction or amputation after leg-threatening injuries. *New Engl. J. Med.* **347**(24), 1924-1931. DOI: 10.1056/NEJMoa012604

Bowles, S. (2009). Did warfare among ancestral hunter-gatherers affect the evolution of human social behaviors? *Science* **324**(5932), 1293-1298. DOI: 10.1126/science.1168112

Boyle, K. E., Heilmann, S., van Ditmarsch, D. and Xavier, J. B. (2013). Exploiting social evolution in biofilms. *Curr. Opin. Microbiol.* **16**(2), 207-212. DOI: 10.1016/j.mib.2013.01.003

Brent, L. J. N., Franks, D. W., Foster, E. A., Balcomb, K. C., Cant, M. A. and Croft, D. P. (2015). Ecological knowledge, leadership, and the evolution of menopause in killer whales. *Curr. Bio.* **25**(6), 746-750. DOI: 10.1016/j.cub.2015.01.037

Buchanan, S. M., Kain, J. S. and de Bivort, B. L. (2015). Neuronal control of locomotor handedness in *Drosophila*. *Proc. Natl. Acad. Sci.* **112**(21), 6700-6705. DOI: 10.1073/pnas.1500804112

Buschges, A. and Pearson, K. G. (1991). Adaptive modifications in the flight system of the locust after the removal of wing proprioceptors. *J. Exp. Biol.* **157**(1), 313-333.

Buschges, A., Ramirez, J.-M., Driesang, R. and Pearson, K. G. (1992). Connections of the forewing tegulae in the locust flight system and their modification following partial deafferentation. *J. Neurobiol.* **23**(1), 44-60. DOI: 10.1002/neu.480230106

Cameron, A. C., Gelbach, J. B. and Miller, D. L. (2011). Robust inference with multiway clustering. *J. Bus. Econ. Stat.* **29**(2), 238-249. DOI: 10.1198/jbes.2010.07136

Caporale, N. and Dan, Y. (2008). Spike timing-dependent plasticity: a Hebbian learning rule. *Annu. Rev. Neurosci.* **31**(1), 25–46. DOI: 10.1146/annurev.neuro.31.060407.125639

Chadha, A., Kaneko, M. and Cook, B. (2015). NOMPC-dependent mechanotransduction shapes the dendrite of proprioceptive neurons. *Neurosci. Lett.* **597**, 111-116. DOI: 10.1016/j.neulet.2015.04.033

Chagnon, N. A. (1988). Life histories, blood revenge, and warfare in a tribal population. *Science* **239**(4843), 985-992. DOI: 10.1126/science.239.4843.985

Choi, J.-K. and Bowles, S. (2007). The coevolution of parochial altruism and war. *Science* **318**(5850), 636-640. DOI: 10.1126/science.1144237

Christakis, N. A. and Fowler, J. H. (2007). The spread of obesity in a large social network over 32 years. *N. Engl. J. Med.* **357**(4), 370-379. DOI: 10.1056/NEJMsa066082

Christakis, N. A. and Fowler, J. H. (2008). The collective dynamics of smoking in a large social network. *N. Engl. J. Med.* **358**(21), 2249-2258. DOI: 10.1056/NEJMsa0706154

Christensen, D. J., Schultz, U. P. and Stoy, K. (2013). A distributed and morphology-independent strategy for adaptive locomotion in self-reconfigurable modular robots. *Robot. Auton. Syst.* **61**(9), 1021-1035. DOI: 10.1016/j.robot.2013.05.009

Connor, R. C. (2010). Cooperation beyond the dyad: on simple models and a complex society. *Philos. Trans. R. Soc. B Biol. Sci.* **365**(1553), 2687-2697. DOI: 10.1098/rstb.2010.0150

Costerton, J. W., Stewart, P. S. and Greenberg, E. P. (1999). Bacterial biofilms: a common cause of persistent infections. *Science* **284**(5418), 1318-1322.

Couzin-Fuchs, E. Kiemel, T., Gal, O., Ayali, A. and Holmes, P. (2015). Intersegmental coupling and recovery from perturbations in freely running cockroaches. *J. Exp. Biol.* **218**(2), 285-297. DOI: 10.1242/jeb.112805

Cruse, H. (1976). The function of the legs in the free walking stick insect, *Carausius morosus*. *J. Comp. Physiol.* **112**(2), 235-262. DOI: 10.1007/BF00606541

Cully, A., Clune, J., Tarapore, D. and Mouret, J.-B. (2015). Robots that can adapt like animals. *Nature* **521**(7553), 503-507. DOI:10.1038/nature14422

Davis, G. F. and Greve, H. R. (1997). Corporate elite networks and governance changes in the 1980s. *Am. J. Sociol.* **103**(1), 1-37. DOI: 10.1086/231170

Delcomyn, F. (2004). Insect walking and robotics. *Ann. Rev. Entomol.* **49**(1), 51-70. DOI: 10.1146/annurev.ento.49.061802.123257

Desai, B.S., Chadha, A. and Cook, B. (2014). The *stum* gene is essential for mechanical sensing in proprioceptive neurons. *Science* **343**(6176), 1256-1259. DOI: 10.1126/science.1247761

Dickinson, M. H., Farley, C. T., Full, R. J., Koehl, M. A. R., Kram, R. and Lehman, S. (2000). How animals move: an integrative view. *Science* **288**(5463), 100-106. DOI: 10.1126/science.288.5463.100

Dietz, V., Grillner, S., Trepp, A., Hubli, M. and Bolliger, M. (2009). Changes in spinal reflex and locomotor activity after a complete spinal cord injury: a common mechanism? *Brain* **132**(8), 2196-2205. DOI: 10.1093/brain/awp124

Diggle, S., Griffin, A., Campbell, G. and West, S. (2007). Cooperation and conflict in quorum-sensing bacterial populations. *Nature* **450**(7168), 411-414. DOI: 10.1038/nature06279

Eddy, S. R. (1996). Hidden Markov Models. *Curr. Opin. Struct. Biol.* **6**(3), 361-365. DOI: 10.1016/S0959-440X(96)80056-X

Fuchs, E., Holmes, P., David, I. and Ayali, A. (2012). Proprioceptive feedback reinforces centrally generated stepping patterns in the cockroach. *J. Exp. Biol.* **215**(11), 1884-1891. DOI: 10.1242/jeb.067488

Full, R. J. and Tu, M. S. (1991). Mechanics of a rapid running insect: two-, four-and six-legged locomotion. *J. Exp. Biol.* **156**(1), 215-231.

Garrity, P. A., Goodman, M. B., Samuel, A. D. and Sengupta, P. (2010). Running hot and cold: behavioral strategies, neural circuits, and the molecular machinery for thermotaxis in *C. elegans* and *Drosophila*. **24**(21), 2365-2382. DOI: 10.1101/gad.1953710

Gat, A. (2015). Proving communal warfare among hunter-gatherers: the quasi-Rousseauan error. *Evol. Anthropol.* **24**(3), 111-126. DOI: 10.1002/evan.21446

Gavrilets, S. (2015). Collective action in heterogenous groups. *Philos. Trans. R. Soc. B Biol. Sci.* **370**(1683), 20150016. DOI: 10.1098/rstb.2015.0016

Gavrilets, S. and Fortunato, L. (2014). A solution to the collective action problem in between-group conflict with within-group inequality. *Nat. Commun.* **5**, 3526. DOI: 10.1038/ncomms4526.

Gilby, I. C., Eberly, L. E. and Wrangham, R. W. (2008). Economic profitability of social predation among wild chimpanzees: individual variation promotes cooperation. *Anim. Behav.* **75**(2), 351-360. DOI: 10.1016/j.anbehav.2007.06.008

Gilby, I. C., Machanda, Z. P., Mjungu, D. C., Rosen, J., Muller, M. N., Pusey, A. E. and Wrangham, R. W. (2015). 'Impact hunters' catalzye cooperative hunting in two wild chimpanzee communities. *Philos. Trans. R. Soc. B Biol. Sci.* **370**(1683), 20150005. DOI: 10.1098/rstb.2015.0005

Glowacki, L. and von Rueden, C. (2015). Leadership solves collective action problems in small-scale societies. *Philos. Trans. R. Soc. B Biol. Sci.* **370**(1683), 20150010. DOI: 10.1098/rstb.2015.0010

Glowacki, L. and Wrangham, R. W. (2013). The role of rewards in motivating participation in simple warfare. *Hum. Nat.* **24**(4), 444-460. DOI: 10.1007/s12110-013-9178-8

Glowacki, L. and Wrangham, R. (2015). Warfare and reproductive success in a tribal population. *Proc. Natl. Acad. Sci. USA* **112**(2), 348-353. DOI: 10.1073/pnas.1412287112

Glowacki, L., Isakov, A., Wrangham, R. W., McDermott, R., Fowler, J. H. and Christakis, N. A. (In preparation for *Proc. Natl. Acad. Sci. USA*). Social networks and the emergence of violence.

Gokhale, C. S. and Traulsen, A. (2014). Evolutionary multiplayer games. *Dyn. Games Appl.* **4**(4), 468-488. DOI: 10.1007/s13235-014-0106-2

Golubitsky, M., Nicol, M. and Stewart, I. (2004). Some curious phenomena in coupled cell networks. *J. Nonlinear Sci.* **14**(2), 207–236. DOI: 10.1007/s00332-003-0593-6

Golubitsky, M., Stewart, I. and Török, A. (2005). Patterns of synchrony in coupled cell networks with multiple arrows. *SIAM J. Appl. Dyn. Syst.* **4**(1), 78–100. DOI: 10.1137/040612634

Gong, Z., Son, W., Chung, Y. D., Kim, J., Shin, D. W., McClung, C. A., Lee, Y. et al. (2004). Two interdependent TRPV channel subunits, Inactive and Nanchung, mediate hearing in *Drosophila*. *J. Neurosci.* **24**(41), 9059-9066. DOI: 10.1523/jneurosci.1645-04.2004

Gore, J., Youk, H. and van Oudenaarden, A. (2009). Snowdrift game dynamics and facultative cheating in yeast. *Nature* **459**(7244), 253-256. DOI: 10.1038/nature07921

Gould, R. (1991). Multiple networks and moblization in the Paris Commune, 1871. *Am. Sociol. Rev.* **56**(6), 716-729. DOI: 10.2307/2096251

Grabowska, M., Godlewska, E., Schmidt, J. and Daun-Gruhn, S. (2012). Quadrupedal gaits in hexapod animals – inter-leg coordination in free-walking adult stick insects. *J. Exp. Biol.* **215**(24), 4255-4266. DOI: 10.1242/jeb.073643

Gray, M., Lukes, J., Archibald, J., Keeling, P. and Doolittle, W. (2010). Irremediable complexity? *Science* **330**(6006), 920-921. DOI: 10.1126/science.1198594

Grosberg, R. and Strathmann, R. (2007). The evolution of multicellularity: a minor major transition? *Annu. Rev. Ecol. Evol. Syst.* **38**(1), 621-654. DOI: 10.1146/annurev.ecolsys.36.102403.114735

Gross, T. and Blasius, B. (2008). Adaptive coevolutionary networks: a review. *J. R. Soc. Interface* **5**(20), 259-271. DOI: 10.1098/rsif.2007.1229

Harkema, S. J. (2001). Neural plasticity after human spinal cord injury: application of locomotor training to the rehabilitation of walking. *Neuroscientist* **7**(5), 455-468. DOI: 10.1177/107385840100700514

Hebb, D. O. (1949). *The organization of behavior.* New York, USA: Wiley.

Henrich, J., Chudek, M. and Boyd, R. (2015). The Big Man Mechanism: how prestige fosters cooperation and creates prosocial leaders. *Philos. Trans. R. Soc. B Biol. Sci.* **370**(1683), 20150013. DOI: 10.1098/rstb.2015.0013

Herculano-Houzel, S. (2009). The human brain in numbers: a linearly scaled-up primate brain. *Front. Hum. Neurosci.* **3**, 31. DOI: 10.3389/neuro.09.031.2009

Hill, K., Hurtado, A. M. and Walker, R. S. (2007). High adult mortality among Hiwi hunter-gatherers: implications for human evolution. *J. Hum. Evol.* **52**(4), 443-454. DOI: 10.1016/j.jhevol.2006.11.003

Holmes, P., Full, R. J., Koditschek, D. and Guckenheimer, J. (2006). The dynamics of legged locomotion: models, analyses, and challenges. *Siam Rev.* **48**(2), 207-304. DOI: 10.1137/S0036144504445133

Hughes, G. M. (1952). The co-ordination of insect movements 1. The walking movements of insects. *J. Exp. Biol.* **29**(2), 267-285.

Ijspeert, A. J. (2008). Central pattern generators for locomotion control in animals and robots: a review. *Neural Networks* **21**(4), 642-653. DOI: 10.1016/j.neunet.2008.03.014

Isakov, A. and Mahadevan, L. (2014). Synchronization in a stochastic Hebbian network of phase oscillators. *ArXiV:* 1404.2328

Isakov, A. and Rand, D. G. (2012). The evolution of coercive institutional punishment. *Dyn. Games Appl.* **2**(1), 97-109. DOI: 10.1007/s13235-011-0020-9

Isakov, A., Murdaugh, K. M., Burke, W. C., Zimmerman, S., Roche, E., Holland, D., Einarsson, J. I. et al. (2014). A new laparoscopic morcellator using an actuated wire mesh and bag. *J. Med. Devices* **8**(1), 011009. DOI: 10.1115/1.4026294

Isakov, A., Holcomb, A., Glowacki, L. and Christakis, N. A. (2016). Modeling the role of networks and individual differences in inter-group violence. *PLoS One* **11**(2), e0148314. DOI: 10.1371/journal.pone.0148314

Isakov, A., Buchanan, S. M., Sullivan, B., Ramachandran, A., Chapman, J. K. S., Lu, E. S., Mahadevan, L. and de Bivort, B. (In press). Recovery of locomotion after injury in *Drosophila* depends on proprioception. *J. Exp. Biol.* DOI: 10.1242/jeb.133652. Available online at: http://jeb.biologists.org/content/early/2016/03/17/jeb.133652

Jahan-Parwar, B. and Fredman, S. M. (1978). Control of pedal and parapodial movements in *Aplysia*. I. Proprioceptive and tactile reflexes. *J. Neurophysiol.* **41**(3), 600-608.

Jain, S. and Krishna, S. (2001). A model for the emergence of cooperation, interdependence, and structure in evolving networks. *Proc. Natl. Acad. Sci. USA* 98(2), 543-547. DOI: 10.1073/pnas.021545098

Jiricny, N., Diggle, S. P., West, S. A., Evans, B. A., Ballantyne, G., Ross-Gillespie, A. and Griffin, A. S. (2010). Fitness correlates with the extent of cheating in a bacterium. *J. Evol. Biol.* **23**(4), 738-747. DOI: 10.1111/j.1420-9101.2010.01939.x

Johnson, N., Carran, S., Botner, J., Fontaine, K., Laxague, N., Nuetzel, P., Turnley, J. et al. (2011). Pattern in escalations in insurgent and terrorist activity. *Science* **333**(6038), 81-84. DOI: 10.1126/science.1205068

Kain, J., Stokes, C., Gaudry, Q., Song, X., Foley, J., Wilson, R. and de Bivort, B. (2013). Leg-tracking and automated behavioural classification in *Drosophila*. *Nat. Commun.* **4**, 1910. DOI: 10.1038/ncomms2908

Kelly, R. C. (2005). The evolution of lethal intergroup violence. *Proc. Natl. Acad. Sci. USA* **102**(43), 15294-15298. DOI: 10.1073/pnas.0505955102

Kim, J., Chung, Y. D., Park, D.-Y., Choi, S., Shin, D. W., Soh, H., Lee H. W. et al. (2003). A TRPV family ion channel required for hearing in *Drosophila*. *Nature* **424**(6944), 81-84. DOI: 10.1038/nature01733

King, A. J., Johnson, D. D. P. and Van Vugt, M. (2009). The origins and evolution of leadership. *Curr. Bio.* **19**(19), R911-R916. DOI: 10.1016/j.cub.2009.07.027

Kirkpatrick Jr, S., Delatt, C. D. and Vecchi, M. P. (1983). Optimization by simulated annealing. *Science.* **220**(4598), 671-680. DOI: 10.1126/science.220.4598.671

Kirpensteijn, J., van den Bos, R. and Endenburg, N. (1999). Adaptation of dogs to the amputation of limb and their owners' satisfaction with the procedure. *Vet. Rec.* **144**, 115-118. DOI: 10.1136/vr.144.5.115

Kohler, T., Perron, G. G., Buckling, A. and van Delden, C. (2010). Quorum sensing inhibition selects for virulence and cooperation in *Pseudomonas aeruginosa*. *PLoS Pathog.* **6**(5), e1000883. DOI: 10.1371/journal.ppat.1000883

Kukillaya, R., Proctor, J. and Holmes, P. (2009). Neuromechanical models for insect locomotion: stability, maneuverability, and proprioceptive feedback. *Chaos* **19**(2), 026107. DOI: 10.1063/1.3141306

Kuramoto, Y. (1984). Chemical oscillations, waves, and turbulence. New York, NY: Springer-Verlag.

Lam, T. and Pearson, K. G. (2001). Proprioceptive modulation of hip flexor activity during the swing phase of locomotion in decerebrate cats. *J. Neurophysiol.* **86**(3), 1321-1332.

Landa, P. and Duboshinsky, Y. B. (1989). Auto-oscillating systems with high-frequency energy sources. *Sov. Phys. Usp.* **158**(4), 729-742.

Langergraber, K. E., Mitani, J. C. and Vigilant, L. (2007). The limited impact of kinship on cooperation in wild chimpanzees. *Proc. Natl. Acad. Sci. USA* **104**(19), 7786-7790. DOI: 10.1073/pnas.0611449104

Lewis, J. S., Wartzok, D. and Heithaus, M. R. (2011). Highly dynamics fission-fusion species can exhibit leadership when traveling. *Behav. Ecol. Sociobiol.* **65**(5), 1061-1069. DOI: 10.1007/s00265-010-1113-y

Lozano, R., Naghavi, M., Foreman, K., Lim, S., Shibuya, K., Aboyans, V., Abraham, J. et al. (2012). Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet* **380**(9859), 2095-2128. DOI: 10.1016/S0140-6736(12)61728-0

Macfarlan, S. J., Walker, R. S., Flinn, M. V. and Chagnon, N. A. (2014). Lethal coalitionary aggression and long-term alliance formation among Yanomamö men. *Proc. Natl. Acad. Sci. USA* **111**(47), 16662-16669. DOI: 10.1073/pnas.1418639111

Martens, E. A., Thutupalli, S., Fourriere, A. and Hallatschek, O. (2013). Chimera states in mechanical oscillator networks. *Proc. Natl. Acad. Sci. USA* **110**(26), 10563–10567. DOI: 10.1073/pnas.1302880110

Mathew, S. and Boyd, R. (2011). Punishment sustains large-scale cooperation in prestate warfare. *Proc. Natl. Acad. Sci. USA* **108**(28), 11375-11380. DOI: 10.1073/pnas.1105604108

McAuliffe, K., Wrangham, R., Glowacki, L. and Russel, A. F. (2015). When cooperation begets cooperation: the role of key individuals in galvanizing support. *Philos. Trans. R. Soc. B Biol. Sci.* **370**(1683), 20150012. DOI: 10.1098/rstb.2015.0012

McDoom, O. S. (2014). Antisocial capital: a profile of Rwandan genocide perpetrators' social networks. *J. Conflict Resolut.* **58**(5), 865-893. DOI: 10.1177/0022002713484282

Mellbye, B. and Schuster, M. (2011). The sociomicrobiology of antivirulence drug resistance: a proof of concept. *mBio* **2**(5), e00131-11. DOI: 10.1128/mBio.00131-11

Mendes, C. S., Bartos, I., Akay, T., Marka, S. and Mann, R. S. (2013). Quantification of gait parameters in freely walking wild type and sensory deprived *Drosophila melanogaster*. *Elife* **2**, e00231. DOI: 10.7554/eLife.00231

Mendes, C. S., Rajendren, S. V., Bartos, I., Marka, S. and Mann, R. S. (2014). Kinematic responses to changes in walking orientation and gravitational load in *Drosophila melanogaster*. *PLos One* **9**(10), e109204. DOI: 10.1371/journal.pone.0109204

Mertens, D. and Weaver, R. (2011). Synchronization and stimulated emission in an array of mechanical phase oscillators on a resonant support. *Phys. Rev. E* **83**(4), 046221. DOI: 10.1103/PhysRevE.83.046221

Mi, Y., Liao, X., Huang, X., Zhang, L., Gu, W., Hu, G. and Wu, S. (2013). Long-period rhythmic synchronous firing in a scale-free network. *Proc. Natl. Acad. Sci. USA* **110**(50), 4931–4936. DOI: 10.1073/pnas.1304680110

Modlmeier, A. P., Keiser, C. N., Watters, J. V., Sih, A. and Pruitt, J. N. (2014). The keystone individual concept: an ecological and evolutionary overview. *Anim. Behav.* **89**, 53-62. DOI: 10.1016/j.anbehav.2013.12.020

Molleman, L., van den Berg, P. and Weissing F. (2014). Consistent individual differences in human social learning strategies. *Nat. Commun.* **5**, 3570. DOI: 10.1038/ncomms4570

Mu, L. and Ritzmann, R. E. (2005). Kinematics and motor activity during tethered walking and turning in the cockroach, *Blaberus discoidalis*. *J. Comp. Physiol. A* **191**(11), 1037-1054. DOI: 10.1007/s00359-005-0029-x

Nagy, J., Victor, E. and Cropper, J. (2007). Why don't all whales have cancer? A novel hypothesis resolving Peto's paradox. *Integr. Comp. Biol.* **47**(2), 317-328. DOI: 10.1093/icb/icm062

Néda, Z., Ravasz, E., Brechet, Y., Vicsek, T. and Barabási, A.-L. (2000). The sound of many hands clapping. *Nature* **403**(6772), 849–850. DOI: 10.1038/35002660

Newland, P. L., Aonuma, H., Sato, M. and Nagayama, T. O. (1996). Presynaptic inhibition of exteroceptive afferents by proprioceptive afferents in the terminal abdominal ganglion of the crayfish. *J. Neurophysiol.* **76**(2), 1047-1058.

Newman, M. E. J. (2001). The structure of scientific collaboration networks. *Proc. Natl. Acad. Sci. USA* **98**(2), 404-409. DOI: 10.1073/pnas.021544898

Newman, M. E. J. (2010). *Networks: an introduction.* Oxford, England: Oxford University Press.

Nowak, M. (2006). Five rules for the evolution of cooperation. *Science* **314**(5805), 1560-1563. DOI: 10.1126/science.1133755

Nunn, C. L. and Deaner, R. O. (2004). Patterns of participation and free riding in territorial conflicts among ringtailed lemurs (*Lemur catta*). *Behav. Ecol. Sociobiol.* **57**(1), 50-61. DOI: 10.1007/s00265-004-0830-5

Otsu, N. (1979). A threshold selection method from gray-level histograms. *IEEE. Trans. Syst. Man Cybern.* **9**(1), 62-66. DOI: 10.1109/TSMC.1979.4310076

Padgett, J. F. and Ansell, C. K. (1993). Robust action and the rise of the Medici, 1400-1434. *Am. J. Sociol.* **98**(6), 1259-1319. DOI: 10.1086/230190

Page, K. L. and Matheson, T. (2009). Functional recovery of aimed scratching movements after a graded proprioceptive manipulation. *J. Neurosci.* **29**(12), 3897-3907. DOI: 10.1523/jneurosci.0089-09.2009

Paoletti, P. and Mahadevan, L. (2014). A proprioceptive neuromechanical theory of crawling. *Proc. Roy. Soc. B Biol. Sci.* **281**(1790), 20141092. DOI: 10.1098/rspb.2014.1092

Papachristos, A. V. (2009). Murder by structure: dominance relations and the social structure of gang homicide. *Am. J. Sociol.* **115**(1), 74-128. DOI: 10.1086/597791

Pearson, K. G. (1972). Central programming and reflex control of walking in the cockroach. *J. Exp. Biol.* **56**(1), 173-193.

Perc, M. and Szolnoki, A. (2010). Coevolutionary games: a mini review. *Biosyst.* **99**(2), 109-125. DOI: 10.1016/j.biosystems.2009.10.003

Perc, M., Gomez-Gardenes, J., Szolnoki, A., Floria, L. and Moreno, Y. (2013). Evolutionary dynamics of group interactions on structured populations: a review. *J. R. Soc. Interface* **10**(80), 20120997-20121013. DOI: 10.1098/rsif.2012.0997

Proctor, J. and Holmes, P. (2010). Reflexes and preflexes: on the role of sensory feedback on rhythmic patterns in insect locomotion. *Biol. Cybern.* **102**(6), 513-531. DOI: 10.1007/s00422-010-0383-9

Pruitt, J. N. and Pinter-Wollman, N. (2015). The legacy effects of keyston indivdiuals on collective behavior scale to how long they remain within a group. *Proc. Roy. Soc. B Biol. Sci.* **282**(1814), 20151766. DOI: 10.1098/rspb.2015.1766

R Foundation for Statistical Computing. (2014). *R: A language and environment for statistical computing*. Vienna, Austria.

Rand, D., Arbesman, S. and Christakis, N. (2011). Dynamic social networks promote cooperation in experiments with humans. *Proc. Natl. Acad. Sci. USA* **108**(48), 19193-19198. DOI: 10.1073/pnas.1108243108

Rand, D. G. and Nowak, M. A. (2013). Human cooperation. *Trends Neurosci.* 17(8), 413-425. DOI: 10.1016/j.tics.2013.06.003

Rapoport, A. and Horvath, W. J. (1961). A study of a large sociogram. *Behav. Sci.* **6**(4), 279–291. DOI: 10.1002/bs.3830060402

Ratcliff, W., Denison, R., Borrello, M. and Travisano, M. (2012). Experimental evolution of multicellularity. *Proc. Natl. Acad. Sci. USA* **109**(5), 1595-1600. DOI: 10.1073/pnas.1115323109

Ridgel, A. L., Frazier, S. F., DiCaprio, R. A. and Zill, S. N. (2000). Encoding of forces by cockroach tibial campaniform sensilla: implications in dynamic control of posture and locomotion. *J. Comp. Physiol. A* **186**(4), 359-374. DOI: 10.1007/s003590050436

Roca, C., Cuesta, J. and Sanchez, A. (2009). Evolutionary game theory: temporal and spatial effects beyond replicator dynamics. *Phys. Life Rev.* **6**(4), 208-249. DOI: 10.1016/j.plrev.2009.08.001

Roithmayr, D., Isakov, A. and Rand, D. G. (2015). Should law keep pace with society? Relative update rates determine the co-evolution of institutional punishment and citizen contributions to public goods. *Games* **6**(2), 124-149. DOI: 10.3390/g6020124

Sandoz, K. M., Mitzimberg, S. M. and Schuster, M. (2007). Social cheating in *Pseudomonas aeruginosa* quorum sensing. *Proc. Natl. Acad. Sci. USA* **104**(40), 15876-15881. DOI: 10.1073/pnas.0705653104

Saper, C. B., Scammell, T. E. and Lu, J. (2005). Hypothalamic regulation of sleep and circadian rhythms. *Nature* **437**(7063), 1257–1263. DOI: 10.1038/nature04284

Schilling, M., Hoinville, T., Schmitz, J. and Cruse, H. (2013). Walknet, a bio-inspired controller for hexapod walking. *Biol. Cybern.* **107**(4), 397-419. DOI: 10.1007/s00422-013-0563-5

Schneider, A.-C., Melis, A. P. and Tomasello, M. (2012). How chimpanzees solve collective action problems. *Proc. Roy. Soc. B Biol. Sci.* **279**(1749), 4946-4954. DOI: 10.1098/rspb.2012.1948

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., Amin, N. et al. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Research*. **13**(11), 2498-2504. DOI:10.1101/gr.1239303

Skardal, P. S., Taylor, D. and Restrepo, J. G. (2014). Complex macroscopic behavior in systems of phase oscillators with adaptive coupling. *Physica D* **267**, 27–35. DOI: 10.1016/j.physd.2013.01.012

Stam, C. J., Hillebrand, A., Wang, H. and Van Mieghem, P. (2010). Emergence of modular structure in a large-scale brain network with interactions between dynamics and connectivity. *Front. Comput. Neurosci.* **4**, 1–13. DOI: 10.3389/fncom.2010.00133

Stein, W. and Schmitz, J. (1999). Multimodal convergence of presynaptic afferent inhibition in insect proprioceptors. *J. Neurophysiol.* **82**(1), 512-514.

Stoltzfus, A. (1999). On the possibility of constructive neutral evolution. *J. Mol. Evol.* **49**(2), 169-181. DOI: 10.1007/PL00006540

Strauss, R. and Heisenberg, M. (1990). Coordination of legs during straight walking and turning in *Drosophila melanogaster*. *J. Comp. Physiol. A* **167**(3), 403-412. DOI: 10.1007/BF00192575

Takeoka, A., Vollenweider, I., Courtine, G. and Arber, S. (2014). Muscle Spindle Feedback Directs Locomotor Recovery and Circuit Reorganization after Spinal Cord Injury. *Cell* **159**(7), 1626-1639. DOI: 10.1016/j.cell.2014.11.019

Sadilek, M. and Thurner, S. (2015). Physiologically motivated multiplex Kuramoto model describes phase diagram of cortical activity. *Sci. Rep.* **5**(10015), 1-8. DOI: 10.1038/srep10015

The MathWorks Inc. (2012) *Matlab and Statistics Toolbox and Image Processing Toolbox Release 2012a.* Natick, Massachusetts, United States.

Timms, L. and English, L. Q. (2014). Synchronization in phase-coupled Kuramoto oscillator networks with axonal delay and synaptic plasticity. *Phys. Rev. E* **89**(3), 032906. DOI: 10.1103/PhysRevE.89.032906

van Belle, S., Estrada, A. and Garber, P. A. (2013). Collective group movement and leadership in wild black howler monkeys (*Alouatta pigra*). *Behav. Ecol. Sociobiol.* **67**(1), 31-41. DOI: 10.1007/s00265-012-1421-5

van Veelen, M., Garcia, J., Rand, D. G. and Nowak, M. A. (2012). Direct reciprocity in structured populations. *Proc. Natl. Acad. Sci. USA* **109**(25), 9929-9934. DOI: 10.1073/pnas.1206694109

Vaughan, C. L. (2003). Theories of bipedal walking: an odyssey. *J. Biomech.* **36**(4), 513-523. DOI: 10.1016/S0021-9290(02)00419-0

Viterna, J. S. (2006). Pulled, pushed, and persuaded: explaining women's mobilization into the Salvadoran Guerrilla Army. *Am. J. Sociol.* **112**(1), 1-45. DOI: 10.1086/502690

von Rueden, C., Gavrilets, S. and Glowacki, L. (2015). Solving the puzzle of collective action through inter-individual differences. *Philos. Trans. R. Soc. B Biol. Sci.* **370**(1683), 20150002. DOI: 10.1098/rstb.2015.0002

Vural, D. C., Morrison, G. and Mahadevan L. (2014). Aging in complex interdependency networks. *Phys. Rev. E* **89**(2), 022811. DOI: 10.1103/PhysRevE.89.022811

Vural D. C., Isakov, A. and Mahadevan, L. (2015). The organization and control of an evolving interdependent population. *J. R. Soc. Interface* **12**(108), 20150044. DOI: 10.1098/rsif.2015.0044

Wang, Z., Szolnoki, A. and Perc, M. (2013). Interdependent network reciprocity in evolutionary games. *Sci. Rep.* **3**(1183), 1-7. DOI:10.1038/srep01183

Ward, A., Liu, J., Feng, Z. and Xu, X. Z. S. (2008). Light-sensitive neurons and channels mediate phototaxis in *C. elegans*. *Nat. Neurosci.* **11**(8), 916-922. DOI: 10.1038/nn.2155

Wen, Q., Po, M. D., Hulme, E., Chen, S., Liu, X., Kwok, S. W., Gershow, M. et al. (2012). Proprioceptive coupling within motor neurons drives *C. elegans* forward locomotion. *Neuron* **76**(4), 750-761. DOI: 10.1016/j.neuron.2012.08.039

West, S., Griffin, A., Gardner, A. and Diggle, S. (2006). Social evolution theory for microorganisms. *Nat. Rev. Microbiol.* **4**(8), 597-607. DOI: 10.1038/nrmicro1461

White, J. G., Southgate, E., Thomson, J. N. and Brenner, S. (1986). The structure of the nervous system of the nematode *Caenorhabditis elegans*: the mind of a worm. *Philos. Trans. R. Soc. B Biol. Sci.* **314**(1165), 1-340. DOI: 10.1098/rstb.1986.0056

Wiesenfeld, K., Colet, P. and Strogatz, S. (1998). Frequency locking in Josephson arrays: connection with the Kuramoto model. *Phys. Rev. E* **57(2)**, 1563–1569. DOI: 10.1103/PhysRevE.57.1563

Wiessner, P. and Pupu, N. (2012). Toward peace: foreign arms and indigenous institutions in a Papua New Guinea society. *Science* **337**(6102), 1651-1654. DOI: 10.1126/science.1221685

Wilson, M. L., Boesch, C., Fruth, B., Furuichi, T., Gilby, I. C., Hashimoto, C., Hobaiter et al. (2014). Lethal aggression in *Pan* is better explained by adaptive strategies than human impacts. *Nature* **513**(7518), 414-417. DOI: 10.1038/nature13727

Wosnitza, A., Bockemuhl, T., Dubbert, M., Scholz, H. and Buschges, A. (2013). Inter-leg coordination in the control of walking speed in *Drosophila*. *J. Exp. Biol.* **216**(3), 480-491. DOI: 10.1242/jeb.078139

Wrangham, R. W. and Glowacki, L. (2012). Intergroup aggression in chimpanzees and war in nomadic hunter-gatherers. *Hum. Nat.* **23**(1), 5-29. DOI: 10.1007/s12110-012-9132-1

Wu, B., Zhou, D., Fu, F., Luo, Q., Wang, L. and Traulsen, A. (2010). Evolution of cooperation on stochastic dynamical networks. *PLoS One* **5**(6), e11187. DOI: 10.1371/journal.pone.0011187

Yi, S. Do, Baek, S. K., Zhu, C. P. and Kim, B. J. (2013). Phase transition in a coevolving network of conformist and contrarian voters. *Phys. Rev. E* **87**(1), 012806. DOI: 10.1103/PhysRevE.87.012806

Yin, H., Mehta, P. G., Meyn, S. P. and Shanbhag, U. V. (2012). Synchronization of coupled oscillators is a game. *IEEE Trans. Automat. Contr.* **57**(4), 920–935. DOI: 10.1109/TAC.2011.2168082

Yuan, W. J. and Zhou, C. (2011). Interplay between structure and dynamics in adaptive complex networks: emergence and amplification of modularity by adaptive dynamics. *Phys. Rev. E* **84**(1), 016116. DOI: 10.1103/PhysRevE.84.016116

Zill, S., Schmitz, J. and Buschges, A. (2004). Load sensing and control of posture and locomotion. *Arth. Struct. Dev.* **33**(3), 273-286. DOI: 10.1016/j.asd.2004.05.005

Appendix A

Supplemental methods and results for Chapter 5

Study population

The Nyangatom number approximately 21,000 and live along the Ethiopian and South Sudanese border in and adjacent to the Lower Omo Valley of Ethiopia. Research was conducted in a border area along the Nyangatom-Turkana frontier to the north of the Kibish River and west of the Kuraz Mountains.

The population structure of the Nyangatom is dynamic. Many Nyangatom reside in mobile villages. These villages may exist for several weeks or longer before disbanding or relocating. Sometimes, members of multiple camps or villages may join together to form a larger village or larger villages may break up into smaller villages. A particular village composition usually results from considerations for providing suitable resources for livestock as well as to maximize security. There are also semi-permanent villages in settled areas, but the population is highly dynamic. Individuals commonly move between these villages and the mobile villages. Young men are generally not attached to any one village for their primary residence, but instead are attached to the livestock owned by their father or other paternal relatives. They change their residence based on the movements of these livestock.

The Nyangatom are organized into territorial sections (pl. *ngiteala*). Membership provides culturally recognized rights to resources in a certain areas of Nyangatom territory. Thus, men living in a specific area generally share membership in common territorial sections. Although individuals may change their residence multiple times a year, they usually do so within a constrained area of Nyangatom territory and revisit the same areas seasonally.

Because of the complex residence dynamics of the Nyangatom in the study area, conducting an analysis of village residence was impractical. Rather, we identified individuals who resided at least seasonally in the study area and were the appropriate age to participate in stealth raids. Men who are elders (sing. *Ekasukout*) do not participate in stealth raids. Although we lack data on their ages, elders are estimated to be above 45 years of age. We also exclude young men below the ages of approximately 18 because they usually do not participate in stealth raids. We identified our sample (N = 91) by their membership in culturally specific age groups where men are deemed old enough to engage in activities such as cattle herding and raiding but not having become elders.

These individuals are expected to be familiar with each other for several reasons. First, the dynamic residence structure ensures that individuals may have resided with each other in the same village. Second, when resources allow, multiple herds of livestock graze together in grazing areas. This usually occurs after the rainy season but can occur at other times as well. In these cases, men from many villages herd together providing collective defense. Finally, there are many ceremonial engagements in which men from throughout the area come together during which ceremonies are performed and animals slaughtered and communally consumed. These ceremonial activities allow men from the study population to spend time together on a semi-regular basis.

Male subjects had facial photographs taken for identification purposes. Each subject was assigned a unique identification number that was used to match the photograph to the subject. Photographs were compiled onto two photo sheets measuring approximately 30 cm x 35 cm each, containing 42 and 49 photographs. Each photograph measured approximately 3 by 5 cm. The photo sheets (supplemented by individual photographs) were used to allow participants to make visual identification of other study participants.

Study participants were compensated with local currency (approximately US$0.25 to US$1.00), tea, or sugar for their participation in study elements. A translator was used for most of the semi-structured interviews and interviews would proceed by the researcher presenting the question to the translator who would ask the question and then translate the answer back to the researcher. The researcher who resided with the Nyangatom (Luke Glowacki) is competent in the local language, allowing him to directly ask interview questions or follow-up questions if the subject's answers were not clear.

Because the field researcher was based at the fieldsite for an extended period of time, data collection occurred throughout the duration of the study and multiple interviews were conducted with study participants. Interviews were frequently conducted with the subject alone. However, due to the open nature of the society, subjects would sometimes be joined by their friends or relatives. Subject comprehension and accuracy was validated in the field by employing consistency checks within and between interviews. These involved asking participants to repeat their answers to questions to test for consistency with their previous answers.

Conflict landscape

The threat of conflict is a daily feature of life for those living in the study area. The Nyangatom have ongoing conflict with several of the neighboring populations. The conflicts involve the use of automatic weapons, including Kalashnikovs that were introduced in the late 1980s and are used throughout the region. Similar to other pastoralist groups in the region, the Nyangatom conduct stealth raids that involve a small number of young men who attempt to capture livestock by stealth. Our data focus on these small raids, ethnographically most similar to those among evolutionary relevant groups such as mobile foragers.

Stealth raids have a very low mortality rate because raiders seek to seize livestock using ambush and stealth and do so only when there is very little risk to themselves. If they cannot find an opportunity with low risk, they will generally abandon their plans. In some cases, dehydration may cause death on raids, but no such instances occurred during the study. During the research period, no subjects were killed or wounded from their participation in a stealth raid.

Stealth raids usually have informal leadership. This usually occurs when an individual decides to initiate a stealth raid and he may then spend several days recruiting other individuals to join him. In some cases, he may visit the village of his desired co-raiders over a period of several days to convince them to go. In other cases, stealth raids emerge after age group events in which large cohorts of men from the same age group are congregated.

During the study period, no commercial or political elements to the conflict were observed or described for the motivations of participating in conflict; thus, the circumstances under study seem appropriate as an ethnographic example of small-scale non-state warfare.

Conflict history data

Conflict is a regular feature of Nyangatom life and talked about openly. It is common for individuals to publicly recount their participation or the participation of their peers. Conflict data was obtained through interviews with study participants about their participation in intergroup conflict events. Co-participants were identified by name or by the use of the facial photograph sheets if the researcher was not aware of the co-participants. Composition was validated through peer reports with other participants allowing confirmation of membership for each raiding party (there were no disagreements about raid party membership). In a few cases, the researcher was unable to confirm co-participant identity, usually because they were not members of the study

population. These individuals were coded as non-subjects and no information was obtained about them.

Leaders were identified through interview questions with raid participants about whether any person was a leader on a particular raid the subject participated in using one of the two words for leader (sing. *Ekarikon*; sing. *Eketamunan*). Leadership was validated by reports from more than one raid participant not including the leader himself. This resulted in the identification of 5 leaders on 19 of the 39 raids. There were six additional cases where three of these five individuals were indicated as being more than a mere participant in raids in which they were not named as a "leader," either by contributing tactical advice or selecting the location of the raid. However, because they were not identified with the locally used referents for leaders, we excluded these six cases.

We do not present specific individual contributions to raids, incidents of defection, or the outcome of the raids. Instead, we focus here on the presence or absence of any individual from the subject population in each raid.

Other data

We collected a variety of other individual-level data on study participants, including height, weight, number of siblings, and paternal wealth. This information is summarized in Table A.1.

Anthropometry

Body weight (kg) and height (cm) were measured in the field for all available subjects. Weight was measured using an electronic scale and height with a stadiometer. Fifteen subjects were unavailable during the anthropometric data collection and are excluded from such analyses.

Sibling relationships

Sibling relationships were collected as part of the demographic and genealogical data collection for the study population. They were elicited through interviews with subjects, in which they were asked to identify their siblings and whether any of their siblings were among the subjects on the photo sheets.

Paternal wealth rankings

Among the Nyangatom, men who are not elders seldom own more than a few head of livestock themselves. Rather, livestock is generally owned by elder male family members. Thus, measures of individual wealth are not culturally appropriate for the men in our sample. We used measures of paternal wealth to explore the relationship between the wealth of a raider's father and raiding party composition. Paternal wealth rankings were obtained for a subset of the study participants (N = 42). These scores were generated from a ranking task in which elder men (raters) – who were not in this sample – were asked to sort facial photographs of subjects into three piles based on the relative wealth of the father. They were initially asked to identify any men featured in the photographs that they did not know. If they could identify all the individuals in the photographs, they were then asked to look at the photographs and determine whether they knew the father of the individual in the photograph. Six elders successfully identified all subjects and their fathers and provided the rankings of paternal wealth.

Raters were instructed that they would sort the photographs into three piles based on information about the fathers of the individual featured in the photograph. They were told that in one pile they were to place the photographs of the men who had the wealthiest fathers. They were then told that a second pile was for the men whose fathers had the least wealth. The final pile was

situated between the wealthiest and the least wealthy and raters were instructed to place the photographs of the men who were between the wealthiest and least wealthy into this pile.

Each individual was ranked 6 times. Each time a subject was placed in the wealthiest pile they received a score of 3; each time they were placed in the middle pile they received a score of 2; and each time they were placed in the least wealthy pile they received a score of 1. The maximum any subject could receive as a wealth ranking was a score of 18 and the minimum a score of 6.

Friendship network data

Subjects were asked to participate in a gift allocation task that was used to generate the friendship network. During this task, subjects were presented with three pieces of candy. Because the population has only minimal access to a market economy, novelty food items, such as candy, are valued. Study participants were presented with the two photo sheets containing the photographs of all 91 study participants. They were asked to make anonymous allocations to three individuals who they desired to receive candy and whose picture was featured on the photo sheets. They were told to do so by placing a piece of candy on the photograph of the subject. Self-allocations and multiple allocations to the same subject were not allowed. Subjects were informed that these allocations would be made at the conclusion of the study but that the identity of the donor would remain anonymous. These allocations were used to generate the friendship network data. Distribution of the candy based on the gift allocations was conducted after the completion of this study element.

We counted the number of times each person was nominated as a friend in the friendship network; this simple measure is called *in-degree*. We also used the nominations to map the full friendship network, and to calculate other social network measures.

Characterizing whether raid participation is due to chance

First, we identify whether raid participation is due purely to chance using two methods: a chi-square test on simulated data and a permutation test. For the chi-square test, we have the null hypothesis H_0 that people are drawn uniformly, i.e. that the probability of drawing person i is the same as the probability of drawing person j, for $i, j \in \{1, \ldots, 91\}$.

The alternate hypothesis H_1 is that not all people have the same probability of being drawn for a raid. We calculate the probability (under the null hypothesis of uniformity) of person i being drawn in raid r, where raid r has the observed number of participants n_r. We find that $P(\text{person } i \text{ being drawn in raid } r) = \frac{\binom{91-1}{n_r-1}}{\binom{91}{n_r}} = \frac{n_r}{91}$, and summing over all raids yields the expected number of observations

$$E_i = \sum_{r=1}^{39} \frac{n_r}{N} \tag{A.1}$$

and hence the chi-square statistic

$$\chi^2 = \sum_{i=1}^{91} \frac{(O_i - E_i)^2}{E_i} \tag{A.2}$$

Performing the chi-square test allows us to reject the null hypothesis ($P \ll 0.001$). We cross-validated this approach using a permutation test with synthetic data. A data set was generated to have the same number of participants in each raid (to account for different costs and benefits associated with each) but with people having equal probability of being chosen in a raid. Here, the null hypothesis H_0 is that the two data sets come from the same distribution, and the alternate hypothesis H_1 is that the two data sets do not come from the same distribution. We pool the data, permute the observations, split into two groups of the same sizes as the original, and calculate the

statistic of interest from this new data. The simulation was run 10^6 times to obtain the null distribution of the statistic. As a natural statistic of interest we choose

$$S = \sum_{i=1}^{91} \frac{|\Delta O_i|}{91}, \tag{A.3}$$

where ΔO_i is the difference of the total number of observations of person *i* among the two groups. Comparing with the original statistic using the observed group and the initial synthetic data, we find that $P \ll 0.001$, which is sufficient to reject the null hypothesis. Therefore, we conclude that raid participation is not simply due to chance.

Identifying leaders algorithmically from the raid participation data

We find a minimal set of raiders that account for participation in all raids. To facilitate this, we use a simple algorithm to establish an upper bound:

1. Calculate the number of raids in which each person participated
2. Find the person who participated in the most raids
3. Mark the raids in which he participated and remove those raids
4. Repeat from step 1 until all raids have been removed

We find that the five ethnographically identified leaders form precisely such a set (see Figure 5.1B, which shows that every raid has at least one leader identified in this manner).

To test that this is a minimal set, we reduce the search space and enumerate all possibilities in the search for a four-person (smaller) set, M. We used a counting argument to reduce the search space. Suppose no ethnographically identified leaders are present in M. Then, the four most active allowed participants only participated a total of 26 times (7 + 7 + 6 + 6). Even if they joined completely disjoint raids, this is not sufficient to account for all 39 raids. Therefore, at least one

leader must be present in M. If exactly one leader is present, the same argument shows that the maximum possible number of raids accounted for is 36. This simple argument shows that at least two ethnographically identified leaders must be present in the proposed set M. We enumerate all the allowed possibilities and do not find such a set, so we conclude that the ethnographically identified leaders indeed form a minimal set.

This analysis also suggests that raid participation data alone would have helped us to identify the leaders that were identified via questions about leadership.

Friendship network structure

While direct comparisons of network datasets are difficult due to different contexts, different ways of ascertaining social connections, and structural differences themselves (e.g., differing network sizes, differing numbers of edges), some comparison of statistics is still informative. In the main text, we present measures of reciprocity, transitivity, and homophily (assortativity) based on degree and age, and discuss the characteristic ways that the Nyangatom social network resembles and differs from simulated random networks and other network data published previously. We find that the degree distribution of the Nyangatom social network was not significantly different from an Erdos-Renyi random network with an identical number of nodes and edges, though the other properties were different in a way similar to several modern networks; in particular, reciprocity, transitivity, and degree assortativity were significantly larger than in random networks, consistent with measures given in prior work (Apicella et al., 2012). As there are multiple related definitions for these measures used in the literature, here we provide formulas for the way our calculations were performed.

Reciprocity (the probability that person i is nominated as a friend by person j given that person j is nominated by person i) was measured as the proportion of mutual connections. That is, given connectivity matrix A,

$$Reciprocity = \frac{\sum_{i,j} A_{ij} A_{ij}^T}{\sum_{i,j} A_{ij}} \tag{A.4}$$

Transitivity (the likelihood that two of a person's friends are themselves friends) is calculated as a global network parameter, i.e., the ratio of connected triples to the total number of possible connected triples in the graph.

To calculate *homophily* (the probability of nodes with similar characteristics being connected), we first assign values of interest to the nodes (e.g. degree, age group). Let e_{ij} be the fraction of edges connecting nodes of type i and j, let $q_i^{(1)} = \sum_j e_{ij}$, $q_i^{(2)} = \sum_j e_{ji}$, and let $\sigma^{(1)}, \sigma^{(2)}$ be the standard deviations of $q_i^{(1)}$, $q_i^{(2)}$, respectively. Then, we calculate homophily as

$$Homophily = \frac{\sum_{j,k} jk(e_{jk} - q_j^{(1)} q_k^{(2)})}{\sigma^{(1)} \sigma^{(2)}} \tag{A.5}$$

Eigenvector centrality assumes that the centrality of a given individual is an increasing function of the centralities of all the individuals to whom he or she is connected. While this is an intuitive way to think about which subjects might be better connected, it yields a practical problem: how do we simultaneously estimate the centrality of all subjects in the network? Let $a_{ij} = 1$ if subjects i and j have a social connection and $a_{ij} = 0$ if they do not. Furthermore, let x be a vector of centrality scores so that each subject's centrality x_j is proportional to the sum of the centralities of the subjects to whom they are connected: $\lambda x_i = a_{1i} x_1 + \cdots + a_{ni} x_n$. This yields n equations, which can be represented as $\lambda x = A^T x$. The vector of centralities x can now be computed since it is an

eigenvector with the eigenvalue λ. Although there are n nonzero solutions to this set of equations, in symmetric matrices, the eigenvector corresponding to the principal eigenvalue is used because it maximizes the accuracy with which the associated eigenvector can reproduce the original social network. To be sure of reaching a solution, we symmetrized all asymmetric relationships in the observed network (i.e., we assumed all friendship ties were mutual).

Social distance has an unusual relationship in the results for non-leader friends (Figure 5.4B). After controlling for friend participation, each friend of friend who participates in a raid actually *decreases* the likelihood a subject will join by 1.6% (SE = 0.6%, P = 0.006). This suggests that people just outside of a person's direct social network may slightly de-motivate participation in raids; weak ties are apparently not useful for recruiting, and may even be somewhat detrimental. The significance of these associations survives models with various controls.

Modeling individual characteristics

In this section, we describe methods to explore what individual characteristics are associated with raid participation, leadership, and network in-degree.

Table A.2-Table A.7 show linear regressions that measure the association between raid participation and various individual characteristics. The basic model is

$$E[Y_i] = \alpha + \beta x_i, \tag{A.6}$$

where the dependent variable Y_i is the total number of raids in which person i participated, x_i is a vector of individual characteristics for participant i, and β is a vector of coefficients that indicate the degree of association with each characteristic. The specific independent variables are height (cm), weight (kg), number of siblings (paternal or maternal), and paternal wealth (only measured for N = 42 people). Models are calculated for the full population (left side of tables) and for the subset of the

population who are not leaders (right side of tables). Only the significance of in-degree (net of sibling contributions, which are measured separately) remains for both the full population and the population that does not include leaders when all individual characteristics are allowed. These models suggest that social information is important for raid participation levels than individual characteristics. We used OLS regression to estimate these models, but count models yielded similar results.

Table A.8-Table A.12 show a similar set of models, but the dependent variable Y_i is the in-degree of person i (net of siblings). Siblings, paternal wealth, and leadership status individually appear to be significantly associated with raid participation in both the full population and the subset of the population that does not include leaders.

Finally, Table A.13 shows a similar model, but the dependent variable Y_i is the eigenvector centrality of person i (net of siblings). This regression shows that even when controlling for the number of direct contacts, leaders tend to have higher centrality, suggesting that leaders not only have more friends, but their friends are more popular and they have more friends of friends as well.

Models with raid and social information and fixed effects

In this section, we describe methods to explore associations between raid characteristics and raid participation.

Table A.14-Table A.20 show linear regressions that measure the association between an individual's decision to join a particular raid and various raid characteristics. The basic model is

$$E[Y_{ir}] = \alpha + \theta_i + \gamma_r + \beta x_{ir}, \tag{A.7}$$

where the dependent variable $Y_{ir} = 1$ if person i joins raids r, and $Y_{ir} = 0$ otherwise, α is a constant (dropped if fixed effects are included), x_{ir} is a vector of characteristics for participant i and

raid r, β is a vector of coefficients that indicate the degree of association with each characteristic, and θ_i and γ_r are individual and raid fixed effects, respectively.

Fixed effects are included to control for variation in stable characteristics across individuals (e.g., are some individuals inherently more likely to join raids?) and across raids (e.g., are some raids inherently more important?). This approach effectively controls for all possible stable individual and raid characteristics. For example, it ensures that personal differences that may impact the tendency of a person to engage in risky behavior or differences that may impact the importance of a raid are not driving the results. Additionally, since there are multiple (and probably correlated) observations for both raids and individuals, we adjust standard errors by clustering them on both raids and individuals using multiway clustering (Cameron et al., 2011).

Table A.14 and Table A.15 show regressions of raid participation on the total number of leaders and non-leaders who joined the raid. Although Table A.14 suggests that the number of non-leaders participating is significantly associated with the decision to join, when we control for individual fixed effects in Table A.15, the association ceases to be significant. Since neither the number of leaders nor the number of non-leaders survives both specifications, we turn to models based on participation by friends rather than total participation.

Table A.16-Table A.18 show regressions of raid participation on social aspects of raid composition. Table A.18 shows the strictest specification with both raid and individual fixed effects. As discussed in the main text, the number of first-degree leaders, first-degree friends, and second-degree friends on a raid are all significantly associated with raid participation, and these results survive multiple model specifications and strong controls for fixed individual and raid characteristics.

Table A.19 shows a regression of raid participation on the number of siblings also in the raid with and without individual or raid fixed effects. Only in the model without fixed effects is the number of siblings significant. For completeness, we consider the prior (full) model with siblings separated out.

Table A.20 shows a regression of raid participation on the social aspects of raid composition (as in Table A.18) with the further inclusion of the number of siblings who participated on raids. We again find that leaders of distance 1, friends of distance 1, and friends of distance 2 are significant, while the number of siblings on raids (net of leader and non-leader friends of distance 1, 2, and 3) is not significant. This again suggests that it is friends, not siblings that matter for the emergence of violence.

Finally, Table A.21 shows a model that regresses the number of non-leader friends who join a raid on a person's own decision to join the raid (1 = joined), their leadership status (1 = leader), and an interaction of the two. The results suggest that leaders who join raids actually mobilize significantly fewer individuals to join than non-leaders, and this is in spite of the fact that leaders tend to have more friends as shown in Table A.12.

Table A.22 shows a model that regresses a person's own decision to join a raid upon the number of leaders and non-leader friends, both net of siblings, who join a raid. These results support the hypothesis that a friendship relationship is more important than a family (sibling) relationship in deciding to join raids. Again, net of siblings, the first-degree friendship effect is by far the largest determinant for joining raids.

To further test the hypothesis that it is friendship ties, and not family ties, that are important, we performed a permutation test of which raids individuals join, keeping the number per raid and the total number of raids participated in by each individual constant, and we asked what percentage

of raids have any sibling pairs. We find that the observed value lies near the center of the distribution of the permuted values, suggesting that siblings do not raid in pairs more frequently than chance, in line with the regression analysis.

Table A.1. Individual level summary statistics.

Variable	*Mean*	*SD*	*Min*	*Max*
Height (cm)	175.4	6.36	160	196
Weight (kg)	58.5	6.46	42	71
Number of Siblings	1.0	1.34	0	5
Paternal Wealth	12.3	3.47	6	18
In-Degree (Number of Friendship Nominations)	3.0	2.67	0	13

Table A.2. Regression of total participation on height. OLS regression of total number of times a person participated in raids on height.

	Dependent Variable: Number of Times Participated in Raids (Full Population)			*Dependent Variable: Number of Times Participated in Raids (Non-Leaders Only)*		
	Coef.	*SE*	*P*	*Coef.*	*SE*	*P*
Height (cm)	0.036	0.059	0.537	-0.033	0.039	0.397
Intercept	-3.414	10.318	0.742	8.144	6.785	0.234
R^2		0.005			0.010	
N		76			72	

Table A.3. Regression of total participation on weight. OLS regression of total number of times a person participated in raids on weight.

	Dependent Variable: Number of Times Participated in Raids (Full Population)			*Dependent Variable: Number of Times Participated in Raids (Non-Leaders Only)*		
	Coef.	*SE*	*P*	*Coef.*	*SE*	*P*
Weight (kg)	**0.120**	**0.056**	**0.036**	0.010	0.036	0.773
Intercept	-4.039	3.313	0.227	1.761	2.084	0.401
R^2		0.058			0.001	
N		76			72	

Table A.4. Regression of total participation on number of siblings. OLS regression of total number of times a person participated in raids on the number of siblings. Both paternal and maternal siblings are counted.

	Dependent Variable: Number of Times Participated in Raids (Full Population)			*Dependent Variable: Number of Times Participated in Raids (Non-Leaders Only)*		
	Coef.	*SE*	*P*	*Coef.*	*SE*	*P*
Siblings (#)	**0.696**	**0.248**	**0.006**	**0.397**	**0.144**	**0.007**
Intercept	2.230	0.414	0.000	1.900	0.231	0.000
R^2		0.081			0.083	
N		91			86	

Table A.5. Regression of total participation on paternal wealth. OLS regression of total number of times a person participated in raids on paternal wealth.

	Dependent Variable: Number of Times Participated in Raids (Full Population)			*Dependent Variable: Number of Times Participated in Raids (Non-Leaders Only)*		
	Coef.	*SE*	*P*	*Coef.*	*SE*	*P*
Paternal Wealth	**0.350**	**0.116**	**0.004**	**0.163**	**0.079**	**0.046**
Intercept	-1.631	1.476	0.276	0.212	0.989	0.832
R^2		0.186			0.100	
N		42			40	

Table A.6. Regression of total participation on social network degree. OLS regression of total number of times a person participated in raids on in-degree (net of siblings) in the gift-giving network.

	Dependent Variable: Number of Times Participated in Raids (Full Population)			*Dependent Variable: Number of Times Participated in Raids (Non-Leaders Only)*		
	Coef.	*SE*	*P*	*Coef.*	*SE*	*P*
In-Degree (Non-Sibling Friends)	**0.766**	**0.120**	**0.000**	**0.495**	**0.064**	**0.000**
Intercept	0.989	0.417	0.020	1.099	0.213	0.000
R^2		0.315			0.416	
N		91			86	

Table A.7. Regression of Total Participation on Individual Characteristics. OLS regression of total number of times a person participated in raids on height, weight, number of siblings, paternal wealth, and in-degree (net of siblings).

	Dependent Variable: Number of Times Participated in Raids (Full Population)			*Dependent Variable: Number of Times Participated in Raids (Non-Leaders Only)*		
	Coef.	*SE*	*P*	*Coef.*	*SE*	*P*
Height (cm)	**0.181**	**0.077**	**0.026**	0.034	0.067	0.613
Weight (kg)	0.097	0.072	0.191	0.105	0.052	0.056
Siblings (#)	0.160	0.278	0.570	-0.228	0.214	0.296
Paternal Wealth	**0.282**	**0.122**	**0.028**	0.168	0.091	0.076
In-Degree (Non-Sibling Friends)	**0.453**	**0.170**	**0.012**	**0.338**	**0.124**	**0.011**
Intercept	-40.121	11.777	0.002	-12.690	10.729	0.247
R^2		0.577			0.434	
N		36			34	

Table A.8. Regression of network in-degree (net of siblings) on height.

	Dependent Variable: In-degree (Full Population)			*Dependent Variable: In-degree (Non-Leaders Only)*		
	Coef.	*SE*	*P*	*Coef.*	*SE*	*P*
Height (kg)	-0.041	0.045	0.366	-0.74	0.050	0.138
Intercept	9.864	7.927	0.217	15.519	8.677	0.078
R^2		0.011			0.031	
N		76			72	

Table A.9. Regression of network in-degree (net of siblings) on weight.

	Dependent Variable: In-degree (Full Population)			*Dependent Variable: In-degree (Non-Leaders Only)*		
	Coef.	*SE*	*P*	*Coef.*	*SE*	*P*
Weight (kg)	0.003	0.045	0.945	-0.026	0.046	0.576
Intercept	2.478	2.630	0.349	4.018	2.689	0.140
R^2		0.000			0.004	
N		76			72	

Table A.10. Regression of network in-degree (net of siblings) on number of siblings.

	Dependent Variable: In-degree (Full Population)			*Dependent Variable: In-degree (Non-Leaders Only)*		
	Coef.	*SE*	*P*	*Coef.*	*SE*	*P*
Siblings (#)	**0.711**	**0.174**	**0.000**	**0.664**	**0.182**	**0.000**
Intercept	1.820	0.290	0.000	1.750	0.293	0.000
R^2		0.158			0.137	
N		91			86	

Table A.11. Regression of network in-degree (net of siblings) on paternal wealth.

	Dependent Variable: In-degree (Full Population)			*Dependent Variable: In-degree (Non-Leaders Only)*		
	Coef.	*SE*	*P*	*Coef.*	*SE*	*P*
Paternal Wealth	**0.286**	**0.098**	**0.006**	**0.254**	**0.104**	**0.019**
Intercept	-0.997	1.256	0.432	-0.659	1.299	0.615
R^2		0.175			0.136	
N		42			40	

Table A.12. Regression of network in-degree (net of siblings) on leadership status.

	Dependent Variable:		
	In-Degree		
	Coef.	*SE*	*P*
Leadership Status	**2.816**	**1.065**	**0.010**
Intercept	2.384	0.250	0.000
R^2		0.073	
N		91	

Table A.13. Regression of eigenvector centrality on number of siblings, network in-degree (net of siblings), and leadership status.

	Dependent Variable:		
	Eigenvector Centrality		
	Coef.	*SE*	*P*
Siblings (#)	**0.011**	**0.003**	**0.001**
In-Degree	**0.014**	**0.014**	**0.000**
Leadership Status	**0.038**	**0.018**	**0.044**
Intercept	0.040	0.006	0.000
R^2		0.564	
N		91	

Table A.14. Regression of raid participation on general raid composition. OLS regression of raid participation on raid composition, with multiway clustering of standard errors on participant and raid. Measures of raid composition exclude the participant (in other words, how many *other* leaders and non-leaders joined the raid?).

	Dependent Variable:		
	Raid Participation		
	Coef.	*SE*	*P*
# Leaders on Raid	-0.002	0.006	0.690
# Non-Leaders on Raid	**0.006**	**0.003**	**0.035**
Intercept	0.048	0.015	0.001
Individual Fixed Effects		NO	
Raid Fixed Effects		NO	
R^2		0.004	
N		3549	

Table A.15. Regression of raid participation on general raid composition with individual fixed effects. OLS regression of raid participation on raid composition, with multiway clustering of standard errors on participant and raid. The model includes fixed effects (not shown) for individuals. Measures of raid composition exclude the participant (in other words, how many *other* leaders and non-leaders joined the raid?).

	Dependent Variable: *Raid Participation*		
	Coef.	*SE*	*P*
# Leaders on Raid	0.004	0.005	0.430
# Non-Leaders on Raid	0.006	0.003	0.051
Individual Fixed Effects		YES	
Raid Fixed Effects		NO	
R^2		0.104	
N		3549	

Table A.16. Regression of raid participation on number of non-leader friends who join the raid. OLS regression of raid participation on the total number of non-leader friends at social distance 1, 2, and 3 who join the raid, with multiway clustering of standard errors on participant and raid.

	Dependent Variable: *Raid Participation*			*Dependent Variable:* *Raid Participation*			*Dependent Variable:* *Raid Participation*		
	Coef.	*SE*	*P*	*Coef.*	*SE*	*P*	*Coef.*	*SE*	*P*
# Non-Leader Friends on Raid (Distance 1)	**0.192**	**0.014**	**0.000**	**0.192**	**0.015**	**0.000**	**0.191**	**0.015**	**0.000**
# Non-Leader Friends of Friends on Raid (Distance 2)				–0.001	0.006	0.888	0.001	0.006	0.859
# Non-Leader Friends of Friends of Friends on Raid (Distance 3)							-0.005	0.003	0.060
Intercept	0.014	0.006	0.023	0.014	0.004	0.001	0.021	0.004	0.000
Individual Fixed Effects		NO			NO			NO	
Raid Fixed Effects		NO			NO			NO	
R^2		0.218			0.218			0.219	
N		3549			3549			3549	

Table A.17. Regression of raid participation on social aspects of raid composition (leaders up to distance 3). OLS regression of raid participation on social aspects of raid composition – the total number of leaders at social distance 1, distance 1 and 2, and distance 1, 2, and 3 on raid, with multiway clustering of standard errors on participant and raid.

	Dependent Variable: *Raid Participation*			*Dependent Variable:* *Raid Participation*			*Dependent Variable:* *Raid Participation*		
	Coef.	*SE*	*P*	*Coef.*	*SE*	*P*	*Coef.*	*SE*	*P*
# Leader Friends on Raid (Distance 1)	**0.140**	**0.022**	**0.000**	**0.140**	**0.022**	**0.000**	**0.143**	**0.022**	**0.000**
# Leader Friends on Raid (Distance 2)	---	---	---	0.008	0.009	0.403	0.009	0.009	0.337
# Leader Friends on Raid (Distance 3)	---	---	---	---	---	---	0.013	0.008	0.118
Intercept	0.055	0.007	0.000	0.051	0.009	0.000	0.042	0.009	0.000
Individual Fixed Effects		NO			NO			NO	
Raid Fixed Effects		NO			NO			NO	
R^2		0.040			0.040			0.041	
N		3549			3549			3549	

Table A.18. Regression of raid participation on social aspects of raid composition (leaders and friends up to distance 3) with individual and raid fixed effects. OLS regression of raid participation on social aspects of raid composition – the total number of leaders and non-leader friends at social distance 1, 2, and 3 on raid, with multiway clustering of standard errors on participant and raid. The model includes fixed effects (not shown) for both individuals and raids.

	Dependent Variable: *Raid Participation*		
	Coef.	*SE*	*P*
# Leader Friends on Raid (Distance 1)	**0.068**	**0.024**	**0.005**
# Leader Friends on Raid (Distance 2)	-0.022	0.020	0.286
# Leader Friends on Raid (Distance 3)	-0.010	0.019	0.576
# Non-Leader Friends on Raid (Distance 1)	**0.189**	**0.017**	**0.000**
# Non-Leader Friends of Friends on Raid (Distance 2)	**-0.016**	**0.006**	**0.007**
# Non-Leader Friends of Friends of Friends on Raid (Distance 3)	-0.009	0.005	0.064
Individual Fixed Effects		YES	
Raid Fixed Effects		YES	
R^2		0.309	
N		3549	

Table A.19. Regression of raid participation on number of siblings in raid with and without fixed effects. OLS regression of raid participation on the number of siblings participating in a raid, with multiway clustering of standard errors on participant and raid. Both paternal and maternal siblings are counted. The models in the middle and the right of the table include fixed effects (not shown) for individuals and for raids, respectively.

	Dependent Variable: *Raid Participation*			*Dependent Variable:* *Raid Participation*			*Dependent Variable:* *Raid Participation*		
	Coef.	*SE*	*P*	*Coef.*	*SE*	*P*	*Coef.*	*SE*	*P*
# Siblings on Raid	**0.045**	**0.021**	**0.032**	0.028	0.022	0.207	0.033	0.021	0.122
Intercept	0.070	0.009	0.000	---	---	---	---	---	---
Individual Fixed Effects		NO			YES			NO	
Raid Fixed Effects		NO			NO			YES	
R^2		0.004			0.101			0.021	
N		3549			3549			3549	

Table A.20. Regression of raid participation on social aspects of raid composition and siblings with individual and raid fixed effects. OLS regression of raid participation on social aspects of raid composition – the total number of leaders and non-leader friends at social distance 1, 2, and 3 on raid, as well as siblings (net leaders and non-leader friends at social distance 1, 2, and 3), with multiway clustering of standard errors on participant and raid. The model includes fixed effects (not shown) for both individuals and raids.

	Dependent Variable: *Raid Participation*		
	Coef.	*SE*	*P*
# Leader Friends on Raid (Distance 1)	**0.067**	**0.024**	**0.005**
# Leader Friends on Raid (Distance 2)	-0.022	0.020	0.282
# Leader Friends on Raid (Distance 3)	-0.011	0.019	0.564
# Non-Leader Friends on Raid (Distance 1)	**0.189**	**0.017**	**0.000**
# Non-Leader Friends on Raid (Distance 2)	**-0.016**	**0.006**	**0.007**
# Non-Leader Friends on Raid (Distance 3)	-0.009	0.005	0.056
# Siblings on Raid (Net Non-Leader Friends and Leaders at Distance 1,2,3)	-0.067	0.056	0.229
Individual Fixed Effects		YES	
Raid Fixed Effects		YES	
R^2		0.310	
N		3549	

Table A.21. Regression of number of non-leader friends on raid on whether a person joins the raid and their leadership status. OLS regression of number of non-leader friends on a raid on an indicator variable of whether a person joins the raid, leadership status, and the interaction of the indicators with raid fixed effects (not shown) and multiway clustering on individuals and raids.

	Dependent Variable: *Number of Non-Leader Friends On Raid*		
	Coef.	*SE*	*P*
Person Joins Raid	**1.216**	**0.130**	**0.000**
Leadership Status	0.056	0.056	0.316
Person Joining Raid × Leadership Status	**-0.585**	**0.221**	**0.008**
Individual Fixed Effects		NO	
Raid Fixed Effects		YES	
R^2		0.275	
N		3549	

Table A.22. Regression of raid participation on social aspects of raid composition (net of siblings) with individual and raid fixed effects. OLS regression of raid participation on social aspects of raid composition – the total number of leaders and non-leader friends at social distance 1, 2, and 3 on raid (net leaders and non-leader friends at social distance 1, 2, and 3), both social relationships net of siblings, clustering standard errors on each participant and raid. The model includes fixed effects (not shown) for both individuals and raids.

	Dependent Variable:		
	Raid Participation		
	Coef.	*SE*	*P*
Number of Siblings on Raid	0.009	0.020	0.642
Number of Non-Sibling Leaders on Raid (Distance 1)	**0.084**	**0.026**	**0.002**
Number of Non-Sibling Leaders on Raid (Distance 2)	-0.019	0.020	0.331
Number of Non-Sibling Leaders on Raid (Distance 3)	-0.009	0.018	0.631
Number of Non-Sibling, Non-Leader Friends on Raid (Distance 1)	**0.202**	**0.017**	**0.000**
Number of Non-Sibling, Non-Leader Friends on Raid (Distance 2)	**-0.014**	**0.006**	**0.013**
Number of Non-Sibling, Non-Leader Friends on Raid (Distance 3)	**-0.010**	**0.004**	**0.022**
Individual Fixed Effects		YES	
Raid Fixed Effects		YES	
Residual		0.316	
N		3549	