THE UNIVERSITY OF CHICAGO

# STRUCTURE, FLUCTUATIONS AND TRANSPORT IN ACTIVE MATTER SYSTEMS: INSIGHTS FROM STOCHASTIC THERMODYNAMICS AND MACHINE LEARNING

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE DIVISION OF THE PHYSICAL SCIENCES
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

DEPARTMENT OF CHEMISTRY

BY
LAURA ROXANA TOCIU

CHICAGO, ILLINOIS
JUNE 2022

In memory of Constantin-Leon Tociu

"Better we raise our skill than lower the climb."

– Royal Robbins

# TABLE OF CONTENTS

# LIST OF FIGURES

# ACKNOWLEDGMENTS

I want to shout out a special thank you to Suri for being an unwaveringly supporting mentor, with an optimism that is unrivalled and a pool of cool new ideas that is never depleted. I could have never ended up where I am in life now, in the city of my dreams with a PhD as well as computational and data science skills without your understanding, flexibility and selflessness!

Thank you to the Vaikuntanathan group members and alumni, for always having time to answer my questions and work with me on every single little bug. You guys are such a fun group: curious, hard-working, insightful and with plenty of interests that sparked so much engaging conversation.

Thank you to my committee member, Benoit Roux, for incredibly useful feedback during my candidacy exam and an engaging biophysics course. Thank you to both Benoit and Vincenzo Vitelli, my other committee member, for inspiring research and conversations.

Thank you to Tessa, Atreyi, Charles and Phil for showing me another side of Chemistry research and for countless evenings of laughter and good times.

Thank you to Jonathan and Chris for working on problem sets and putting up with my know-it-all attitude during the courses of our first year of graduate school.

Thank you to all my climbing pals such as Laimei, Yuqing, Sebastien and the URock club for giving me belays and inspiration to keep pursuing my climbing passion throughout graduate school. Also thank you loads to all my climbing buddies in Las Vegas who did the same thing during the pandemic: Jesse, Pat, Natasja, Nancy and many more.

Last but not least, special thank you to my boyfriend James for being the sweetest companion and most supportive fan of mine in everything I do, including this PhD.

In the Introduction and Future Directions chapters of this dissertation, I resort to the pronoun “I” in order to illustrate that it is my own view of what is important to mention in order to familiarize the reader with the work and to get him/her excited for the future

of this research. In Chapters 2-4, I resort to the pronoun "we" when I describe the research done, in order to emphasize that, even when the work was done predominantly by myself, collaboration with my wonderful colleagues and adviser was vital and led to the finished work reproduced here.

The research described in this dissertation was financially supported by funds from the University of Chicago; by the University of Chicago John C. Light Fellowship in Theoretical Chemistry; by the University of Chicago Eckhardt Fellowship; by the MRSEC Fellowship for Collaborative Research; by the Department of Energy, Basic Energy Sciences (DOE BES) through Grant DE-SC0019765; by the Army Research Office under grant number W911NF-16-1-0415; by the University of Chicago Materials Research Science and Engineering Center, which is funded by National Science Foundation under award number DMR-1420709 and by the National Science Foundation under award number DMR-1848306.

# ABSTRACT

Living systems, as well as most physical systems with promise in industry, technology and medicine, operate outside of thermodynamic equilibrium. While there exists a precise theoretical framework called statistical mechanics to treat equilibrium systems, there is no equivalent theory for out-of-equilibrium systems, making the discovery of principles that dictate nonequilibrium phenomena of crucial importance.

The search for principles to better understand non-equilibrium dynamics often starts with a quest to understand energy dissipation, the amount of energy consumed due to non-conservative forces acting on the system.

In this dissertation I look at several systems that display a variety of non-equilibrium behaviors in order to understand how energy dissipated by nonconservative forces applied at the microscopic level affects macroscopic properties. This dissertation contains three major chapters, each one containing derivations of new nonequilibrium statistical mechanics results, followed by proof of principle experiments that show how the analytical work could be used to garner novel understanding.

First, concepts from stochastic thermodynamics and liquid state theories are used to study the relationship between energy dissipation and changes in the structural properties of a minimal model of a liquid driven by external rotating forces. As a proof of principle, a model system of a particle diffusing in a fluctuating energy landscape under a force that has a non-conservative component is considered. This model is used to draw insight into how nonequilibrium forcing as well as a changing conservative force landscape felt by an active particle can affect its diffusion and lead to phase transitions.

Second, explicit coarse-graining and stochastic calculus are used to obtain more precise, quantitative relationships between energy dissipation, density correlations and diffusion in generic nonequilibrium liquids, a class that the fluid in part one belongs to. As a proof of principle, large-deviation biased ensembles are considered, where trajectories are steered in

a way that mimics the effects of an external drive. The choice of the biasing function is informed by the aforementioned relationship between dissipation and density correlations. Using analytical and computational techniques, it is shown that the choice of biasing effectively renormalizes interactions in a controlled manner.

Third, a mean-field framework to connect energy dissipation and structure is constructed. Unlike most existing approaches, this framework is applicable even for nonequilibrium liquids with strong interactions. The theory reveals a robust analytical relation between dissipation and density correlations. As a proof of principle, these results are further used to construct a neural network which maps static configurations of particles to their dissipation rate without any prior knowledge of the underlying dynamics.

Altogether, these results add significant contributions to the small pool of existing analytical and theoretical expressions to describe nonequilibrium systems. Furthermore, through proof of principle experiments inspired by these analytical results, this work provides novel perspectives on the interplay between dissipation and the structure, transport and function of liquids and soft materials.

# CHAPTER 1

# INTRODUCTION

Equilibrium statistical mechanics is a powerful theory that connects observable, macroscopic properties such as enthalpy and entropy to its microscopic constituents. It has a long history and is a well established theoretical framework in which many simple systems can be solved exactly. However, the most common and interesting systems that occur in nature, or are used in industry and technology, are predominately stochastic systems that are dissipative in nature and in which many constituents interact with each other. Such systems do not achieve equilibrium states such as those usually found in thermodynamics, and are thus known as nonequilbrium systems. Examples of these systems span many length- and time-scales, from bacterial swarms [47], assemblies of self-propelled colloids [134], self-healing materials [192], to animal groups [23] and human crowds [3]. As all of these systems dissipate heat to their environment when they operate, it is natural to ask how the energy dissipated is connected to their structural and functional properties. The close connection between the net dissipation of energy, powered by non-conservative forces, and spatial organization is especially apparent in living systems [183, 58, 7, 70]. As an example, the flagella motors of *E. Coli* exhibit a unique phenomenology combining ultra-sensitive response, adaptation, and motor restructuring as a function of the applied load [102, 100, 191]. Moreover, *in vivo* studies of the cellular cytoskeleton, as well as *in vitro* experiments on reconstituted systems, have also shown that motor-induced forces control a large variety of functionality in the cell [159, 151, 13, 121, 37].

While one can intuitively appreciate these fascinating phenomena, in order to harness their functionality to create useful materials and machines, a meticulous understanding of how energy dissipation maps to function output is needed.

In this dissertation, I contribute to this problem by focusing on how energy consumption controls the structural and functional properties of a class of systems that model real-life

nonequilibrium phenomena.

The class of systems I am concerned with are active matter systems, in which non-conservative forces, either external or self-propulsion forces, act at the individual level of each constituent. The constituents can either interact with each other or not. The energy dissipated by nonconservative forces, sometimes at interplay with the interactions between particles, leads to interesting nonequilibrium phenomena. Examples of natural systems that can be modeled using this framework are flocks of starlings [4], swarms of bacteria [42], or the molecular motors that make up the cell cytoskeleton [63].

Specifically, in Chapter 2, I consider a model system of Brownian particles interacting through repulsive forces. To drive this system out of equilibrium, external forces are applied to the particles. In order to keep the nature of the external forces as simple as possible so as to be able to focus on the larger picture of the nonequilbrium nature of the system, circular forces that are all in phase, with constant period and amplitude are applied to the system. However, applying such a force to all particles would result in an equilibrium system by switching to a point of view that rotates with the particles. To circumvent this, the external rotating forces are applied to only half of the particles, achieving the desired nonequilibrium effect. This system is partly inspired by and exhibits a similar nonequilibrium phase transition to the experimental colloidal system considered in [76]. In Chapter 3 and 4 the complexity is increased by considering more generic active matter dynamics such as active Ornstein-Uhlenbeck processes [43]. The particles I study in these chapters interact through repulsive potentials too, and some or all of them experience self-propulsion forces in the form of colored noise. In the proof of concept experiments, I consider even more complex systems which illustrate the promise of the results in this work to explore new pathways to use dissipation to tune functionality. Specifically, in Chapter 3, I look at Vicsek particles, which are interacting dipoles subject to self-aligning propulsion forces, to illustrate how a dynamical bias inspired by the analytical results can steer the system between isotropic and

polar phases. In Chapter 4, I consider active rotors similar to those in [3], a class of active matter systems in which constituents are actively rotating but experience no self-propulsion, to show how a machine learning framework can infer novel analytical results similar to the ones derived in this chapter, without any prior knowledge.

The focus of this dissertation is primarily to connect energy dissipation, arising from microscopic nonequilibrium forces, to emerging structure, transport and functionality. Secondarily, I also strive to understand how interactions between the microscopic constituents affect this connection.

To address the first question, I rely heavily on tools from stochastic thermodynamics. Stochastic thermodynamics is a framework in which the energetics of stochastic processes are understood in terms of the usual energy terms in thermodynamics [156]. To visualize this, let us consider one of the simplest stochastic systems, that of a single 1D particle experiencing Langevin dynamics with damping coefficient $\gamma$ and subject to a potential $U$:

$$\gamma \frac{dx}{dt} = -\frac{dU}{dx} + \epsilon(t) \tag{1.1}$$

In the above the fluctuating thermal noise term obeys $\langle \epsilon(t) \rangle = 0$ and $\langle \epsilon(t)\epsilon(t') \rangle = 2\gamma T \sigma(t - t')$.

By rearranging the above equation and multiplying by $dx$, it is derived that:

$$0 = -\left( -\gamma \frac{dx}{dt} + \epsilon(t) \right) dx + \frac{dU}{dx} dx \tag{1.2}$$

The first term on the right-hand side represents the reaction force that the particle exerts on the bath, multiplied by the displacement. This term is thus the work done by the particle on the bath, or, in other words, the heat discarded by the system into the heat bath. Note that this definition of heat may have a different sign from other commonly used conventions. The second term on the right-hand side is the change in internal energy of the system,

yielding:

$$0 = d\mathcal{Q} + dU \tag{1.3}$$

By making the potential in Eq. 1.1 depend on an external tunable parameter $a$, it can be derived that [156]:

$$\frac{\partial U(a,x)}{\partial a} da = d\mathcal{W} = d\mathcal{Q} + dU \tag{1.4}$$

Altogether, it becomes apparent that in a Langevin system, the balance of forces ensures that the energy of the system plus bath remains constant, which recovers the first law of thermodynamics above.

In the presence of a nonequilibrium external or self-propulsion force, let's call it $F$, the following Langevin equation of motion can be written down:

$$\gamma \frac{dx}{dt} = -\frac{dU}{dx} + F(t) + \epsilon(t) \tag{1.5}$$

Analogously to the prior case, the energy dissipated by the system into the bath will be equal to $-\left(-\gamma \frac{dx}{dt} + \epsilon(t)\right) dx$, which can be re-expressed in terms of the nonequilibrium force. In this dissertation, for reasons explained in detail in Chapter 3, Section 3.4, I focus on the flux rate of the non-constant part of this quantity, which I refer to as the rate of work done by the nonconservative force, $\dot{w} = \left\langle \frac{dU}{dx} F \right\rangle$.

In Chapters 3 and 4 of this work, time-averaged energy balance equations in nonequilibrium steady states are written down, using the concepts introduced here, to connect the rate of work $\dot{w}$ to liquid properties. Specifically, in Chapter 3, a connection between the rate of work and force fluctuations in a simple system is derived. In Chapter 4, I show an exact expression that illustrates how rate of work changes the two-body and three-body correlation functions of generic active liquids.

My second main focus in this dissertation is on how the strength and nature of the interactions between the constituents affects the derived relationships between rate of work and function. In Chapter 3 a handful of quantitative connections are derived, for example, between rate of work and diffusion of a tracer particle. These connections are derived in the context of a mean-field theory in which the coarse-grained density field is solved for approximately, under the assumption of weak interactions.

To introduce this coarse-graining, let's consider the Langevin dynamics of a system of particles in a thermal white noise heat bath. The particles interact via a pairwise potential $V(x)$. Each particle experiences a thermal white noise and moves under the force generated from the potential due to its neighbours. Consequently the $i$th particle obeys the Langevin

$$\frac{dX_i(t)}{dt} = \eta_i(t) - \sum_{j=1}^{N} \nabla V(X_i(t) - X_j(t)), \tag{1.6}$$

The noise is uncorrelated in time and the noise acting on a single particle is not correlated with the noise acting on the others, so that $\langle \eta_i^{\mu}(t)\eta_j^{\nu}(t')\rangle = 2T\delta_{ij}\delta^{\mu\nu}\delta(t-t')$. In 1996, Dean wrote down a mean-field equation [36] for the evolution of the density field $\rho(x,t) = \sum_{i=1}^{N}\delta(X_i(t)-x)$ under the assumption of high density or weak interactions, such that any fluctuations around the mean density can be neglected at second order. The mean-field equation reads:

$$\frac{\partial\rho(x,t)}{\partial t} = \nabla\cdot[\eta(x,t)\rho^{1/2}(x,t)] + \nabla\cdot\left(\rho(x,t)\int dy\rho(y,t)\nabla V(x-y)\right) + T\nabla^2\rho(x,t) \tag{1.7}$$

In Chapter 4, I use a version of Eq. 1.7 adapted for an active tracer and ideas from solid state theories to make a substitution that extends the validity of the mean field theory to strong pair-wise interactions as well. Using this theory, it is possible to make accurate predictions on how rate of work affects structure in a more comprehensive range of systems than in Chapters 2 and 3.

The unifying theme in this dissertation is the focus on dissipation and, to a lesser extent, interactions as the concepts with the most promise to help understand how to harness the desirable properties of active matter. The tools and the results of this work draw important lessons about the relationship between microscopic forces and emerging nonequilibrium phenomena.

# CHAPTER 2

# DISSIPATION, PHASE DIAGRAM, STRUCTURE AND TRANSPORT IN A MODEL ACTIVE LIQUID

The material in this chapter is adapted with permission from the following publication: [93] del Junco, C.; Tociu, L.; Vaikuntanathan, S. Energy dissipation and fluctuations in a driven liquid. *Proc. Natl. Acad. Sci. USA*, 115.14 (2018).

## 2.1 Introduction

This chapter studies how energy dissipation affects the phase diagram, structure and transport of a minimal model of an active liquid. Minimal models of active matter have provided an analytically and computationally tractable test bed to study non-equilibrium systems. Phase transitions in some classes of model systems composed of self-propelled particles are beginning to be characterized [173, 167, 175]. Recent work has also studied nucleation phenomena [147, 145] and obtained expressions for pressure and other mechanical properties of active media [160, 11]. In spite of these advances, understanding the connections between energy consumption and the structural properties of these systems remains a challenging problem [7].

In our work, we explore these connections in a model non-equilibrium liquid driven by time-periodic forces. The rotational dynamics that result from the driving have similarities with a range of systems, including colloids in a periodically changing magnetic or electric field [197, 76, 196, 190], shaken plastic particles and chiral wires [74, 184], and chemical and biological microswimmers with active rotational degrees of freedom [98, 135]. Despite their relevance to this wide range of experimental and biological systems, model systems with rotational dynamics have only recently begun to be studied [200, 130, 33].

The central points of this paper are as follows: First, we describe the class of driven

liquids considered in this paper and their associated phase diagrams. Second, we show that the density fluctuations in our non-equilibrium system are surprisingly well-described by Gaussian statistics [24]. Within this effective description, we are able to write down simple scaling relations for the amount of work performed on the system due to the non-equilibrium driving forces. These analytical predictions are validated by simulation data from many-particle systems. Third, we derive a relation that shows how the rate of work done on the system changes the fluctuations in the conservative forces experienced by the particles. In other words, our relation describes how energy dissipation changes the structural properties of the non-equilibrium material. This relation between work and force fluctuations can be viewed as an instantiation of the Harada-Sasa relation that connects work performed in a non-equilibrium system to a breakdown of the fluctuation dissipation relation [81, 80]. We demonstrate this relation numerically for a variety of driven systems to firmly establish its general nature. Fourth, we use a minimal model to demonstrate how a breakdown of the fluctuation dissipation relation can change the diffusion constant of a system. While our driven liquid is substantially more complex than the minimal model, the model provides intuition for the interplay between dissipation and transport. In particular, it shows how a violation of the fluctuation dissipation relation can lead to an increase in the diffusion constant. Finally, we show that such an enhancement of the diffusion constant can drive phase transitions in the liquid. Taken together, our results elucidate how a violation of time reversal symmetry can be used to alter the structural, transport, and phase transition properties of a liquid.

## 2.2 Model and Methods

Our model system is composed of purely repulsive 2-dimensional disks whose positions evolve in time according to driven Brownian dynamics:

$$\dot{\mathbf{r}}_\mathbf{i}(t) = D_0\beta\left(\mathbf{F}_{\mathrm{c,i}}(t) + \mathbf{F}_\mathrm{d}(t)\right) + \boldsymbol{\eta}_i(t) \tag{2.1}$$

where $D_0$ is the bare diffusion constant and $\mathbf{F}_{\mathrm{c,i}}$ is the force on particle $i$ which results from a purely repulsive Weeks-Chandler-Anderson interaction potential [193]:

$$u(r_{ij}) = \begin{cases} 4\epsilon\left[\left(\frac{r_0}{r_{ij}}\right)^{12} - \left(\frac{r_0}{r_{ij}}\right)^{6}\right] + \epsilon, & r \leq 2^{1/6}r_0 \\ 0, & r > 2^{1/6}r_0 \end{cases} \tag{2.2}$$

where $r_0$ is the particle radius and $\epsilon$ is the interaction strength. The random noise terms $\boldsymbol{\eta}_i(t) = (\eta_{i,x}(t), \eta_{i,y}(t))$ are Gaussian-distributed random variables with $\langle\boldsymbol{\eta}_i(t)\rangle = (0,0)$ and $\left\langle\eta_{i,\mu}(t)\eta_{j,\nu}(t')\right\rangle = 2D_0\delta_{i,j}\delta_{\mu,\nu}\delta(t-t')$. In all of our simulations and calculations, we set $\beta = 1/k_BT = 1$, $D_0 = 1$ and $r_0 = 1$. All results reported are for a driven particle fraction $f = 0.5$ and number density $\rho = N/L^2 = 0.45$.

In addition to the conservative forces, half of the particles are driven by an external force acting on the center of mass of the particle whose direction changes periodically in time according to:

$$\mathbf{F}_\mathrm{d} = A\sin\theta\hat{e}_x + A\cos\theta\hat{e}_y \tag{2.3}$$

$$\theta = 2\pi t/\tau. \tag{2.4}$$

For the other half of the particles, $\mathbf{F}_\mathrm{d} = 0$. Thus, in the zero-temperature limit, a single driven particle will trace a circle in the plane. $\mathbf{F}_\mathrm{d}$ is the same for all driven particles, so that a pure system of driven particles, if we move to a frame of reference that is rotating with

$\mathbf{F}_{\mathrm{d}}$, will look the same as the equilibrium system. This model was motivated in part by a recent experimental active matter system developed by Luijten, Granick and coworkers [76].

The driving force in Eq. 2.3 is characterized by a period $\tau$ and an amplitude $A$. The latter is quantified by the Peclet number, a dimensionless measure of the ratio of advective to diffusive velocity in the system. Here we define it as $Pe = \frac{A/\gamma}{D_0/r_0}$, where $A$ is the amplitude of the driving force and $\gamma$ is the damping coefficient. In all of the simulations, $\gamma = 100$. The period of the driving force is measured in units of time set by $t_0 = r_0^2/D_0$. LAMMPS [139] was used to integrate the equations of motion.

To construct the phase diagram in Sec. 2.3, square simulation boxes of size $50r_0 \times 50r_0$ with periodic boundary conditions were filled with 1125 particles with random initial positions to get $\rho = 0.45$. We scanned values of $Pe$ from 10-160 and $\tau$ from 0.05 to $0.3t_0$. Simulations were run with timestep $\Delta t = 10^{-6}t_0$ up to $t = 1000t_0$. To identify different phases in the system, we defined an order parameter $\Omega$ which measures the degree of segregation of driven and passive particles. To compute $\Omega$, we randomly selected a number of small probe areas with radius $1.75r_0$ in the steady-state system and we measured $|N_R - N_B|/N$, where $N_{R,B}$ is the number of red/blue particles and $N$ is the total number of particles, in each probe volume, to give a value between 0 and 1. We then averaged over probe areas and many frames.

To calculate the density fluctuations in a probe volume in Sec. 2.4, the number of particles $N$ in a randomly selected probe volume of radius $1.75r_0$ was measured in 20 probe volumes per frame in 200 frames of a simulation after the system had reached the steady state. Individual data were binned to construct the histogram.

To calculate the work that we refer to in Sec. 2.5, we used the same simulations that we used to construct the phase diagram, and we integrated $\dot{w} = \left\langle \mathbf{F}_{\mathrm{c,i}} \delta \mathbf{r}_{\mathrm{i}} \right\rangle$, where the $\delta \mathbf{r}_{\mathrm{i}}$ is the displacement of particle $i$ due to the non-conservative forces only in 1 timestep and $\mathbf{F}_{\mathrm{c,i}}$ is the conservative force. The average is over all particles in the system. Measurements were

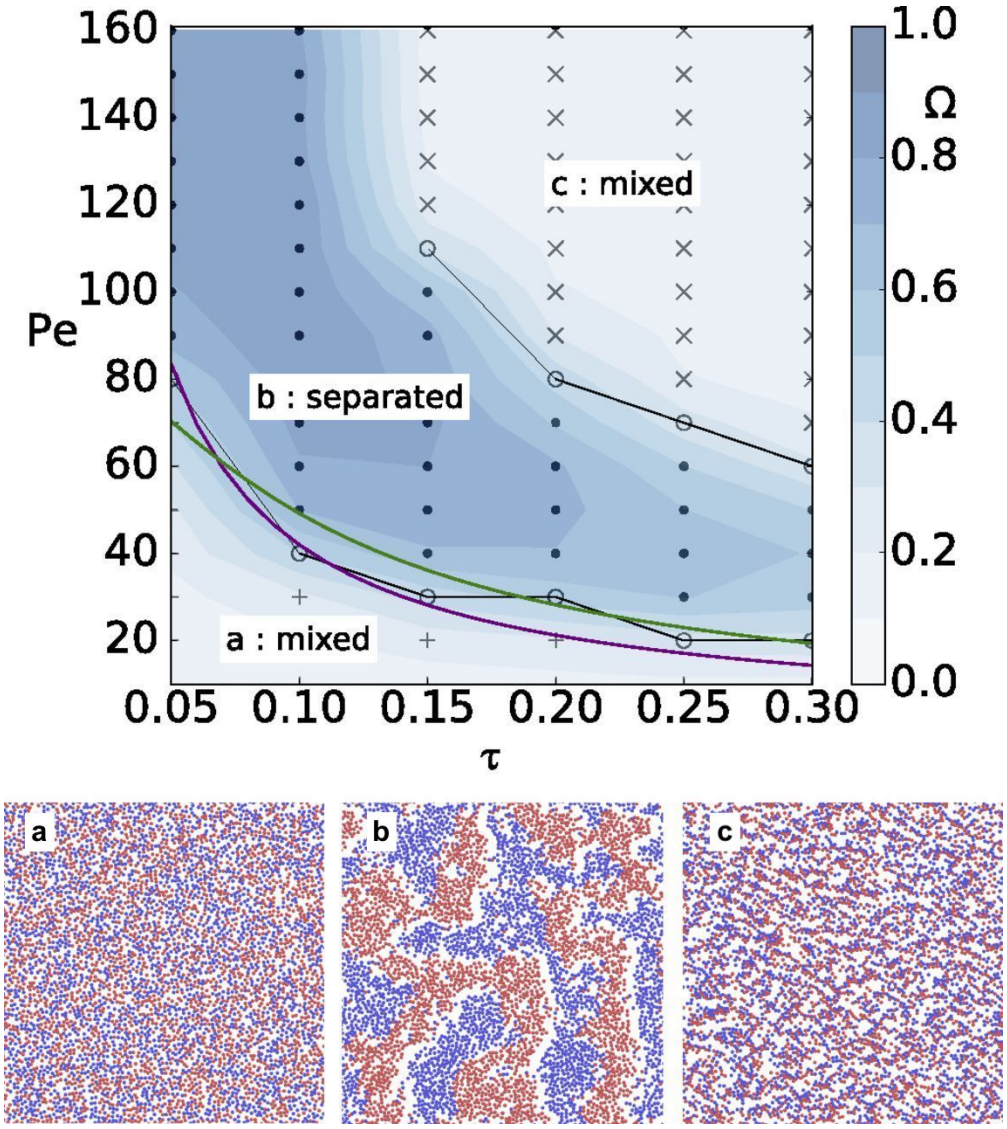



Figure 2.1: (Top) Phase diagram of the system with circular driving. At low $Pe$, the system is mixed (plus signs). As $Pe$ is increased, the system undergoes a transition to a phase separated state (points). As $Pe$ is increased further, there is a transition to a mixed state characterized by large variations in local density (crosses). Letters refer to representative snapshots below. The locations of the transitions are defined as the maxima of $\partial\Omega/\partial Pe$ (black lines with circles). The green line is proportional to $Pe^2 b(\tau)$, where $b(\tau)$ is defined in Eq. 2.16. The purple line is proportional to the work as given by Eq. 2.8, where we have used the mean values of $\gamma_G$ and $K_G$ found from the fits in Fig. 2.5. Contours were drawn using interpolated values. (Bottom) Snapshots of the final frame of a simulation in different regions of the phase diagram. Driven particles are colored red, and undriven particles are colored blue. (a) At $Pe = 15$, $\tau = 0.1$. (b) At $Pe = 70$, $\tau = 0.15$. (c) At $Pe = 150$, $\tau = 0.2$.

made for $100\tau$ after the system reached the steady state.

To calculate diffusion of the fluid in Sec. 2.7, square simulation boxes of size $200r_0 \times 200r_0$ with PBC were filled with 18000 particles. Trajectories were run for at least $1000\tau$ with a timestep of $\Delta t = 10^{-5}t_0$. The mean-squared-displacement (MSD) of all particles in the simulations was measured after each cycle $\tau$, so that there is no net displacement due to the driving forces. The diffusion coefficient was extracted by fitting the MSD as a function of time according to $\left\langle x^2(t)\right\rangle = 4Dt$.

## 2.3 Phase Behavior

To construct the phase diagram of the system, we used the parameter $\Omega$ described in the Methods section. For a completely mixed system $\Omega$ is 0, while for a completely separated system $\Omega$ is slightly lower than 1 due to the presence of interfaces. The choice of a volume of radius $1.75r_0$ mentioned in the Methods section was large enough to ensure $\Omega$ takes on a range of values, but also small enough that inhomogeneities are not all washed out. Since the value of $\Omega$ is sensitive to the size of probe volume and the overall density of the system, we defined a phase transition as the location where $\partial\Omega/\partial Pe$ at constant $\tau$ is steepest, as illustrated in Fig. 2.2.

We found 3 distinct phases in the region of $Pe, \tau$ space that we studied here (Fig. 2.1). At low $Pe$, driven and passive particles remain mixed and the system is homogeneous. As $Pe$ is increased, they segregate into regions of purely driven or passive particles. Similar to other strongly damped active systems with rotating dynamics, the interfaces have no particular orientation, and there is a particle current along the interface that decays rapidly into the bulk [150]. The steady state is time-periodic with a period $\tau$. As $Pe$ is increased further, the system undergoes a transition to a mixed phase characterized by large variations in the local particle number density. Through both transitions, the value of $\Omega$ changes smoothly. As $\tau$ is increased, phase separation persists over a smaller range of $Pe$. The curve separating the

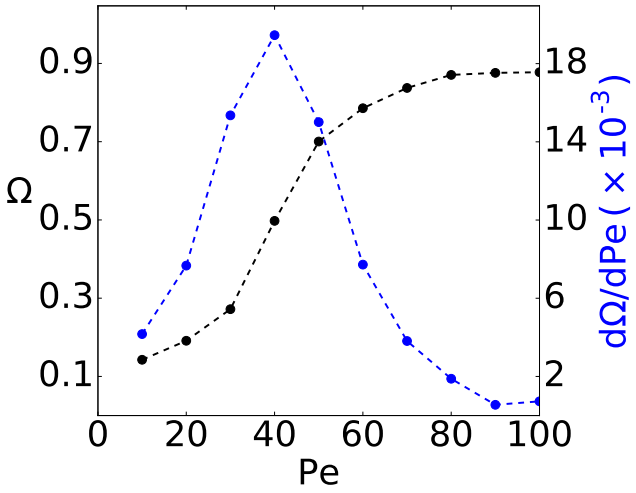



Figure 2.2: A representative plot of $\Omega$ and $\partial\Omega/\partial Pe$ near the location of the lower phase transition, taken at a cross-section of the phase diagram where $\tau = 10$. The value of $\Omega$ changes rapidly near the location of the transition, resulting in a sharp peak in $\partial\Omega/\partial Pe$. Finite size scaling analysis would be required to determine whether the value of $\Omega$ changes smoothly or jumps in the limit of a large simulation size. Dashed lines are guides for the eye.

regions $a$ and $b$ in the phase diagram in Fig. 2.1 are well described by the functional form $Pe^2 f(\tau) = c$, where $c$ is a constant.

At this point, we are not able to unambiguously determine the order of this phase transition. We note, however, that interfaces in region $b$ seem to possess a finite surface tension and merge to form a single stable system-spanning interface in the steady state. This result, together with the behavior of the order parameter shown in Fig. 2.2, suggest a first order phase transition separating regions $a$ and $b$. In Ref. [76], the authors describe a phase transition in an equivalent system from a mixed state to an unmixed state when the value of $\tau$ is decreased while the driving force on the particles is held constant, corresponding to passing from region $c$ to region $b$ in our phase diagram. They demonstrate that it is a bona fide phase transition. The system in this reference has no random noise and so would not mix again at finite $Pe$; here, as a result of the finite temperature, we observe the second change from region $b$ to $a$. The work in this paper and Ref. [76] is therefore distinct and complementary.

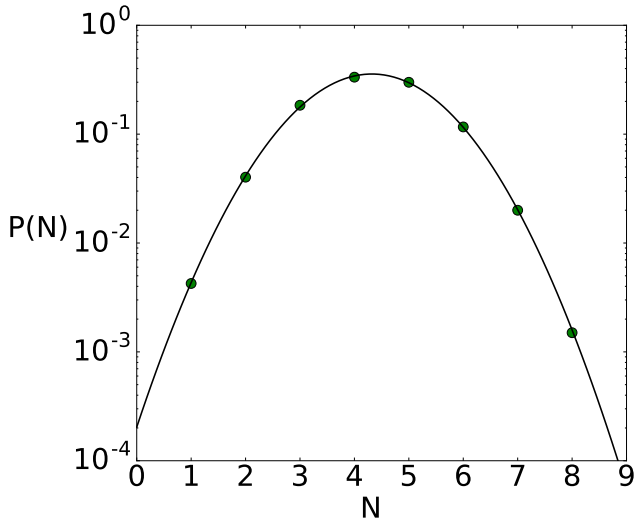


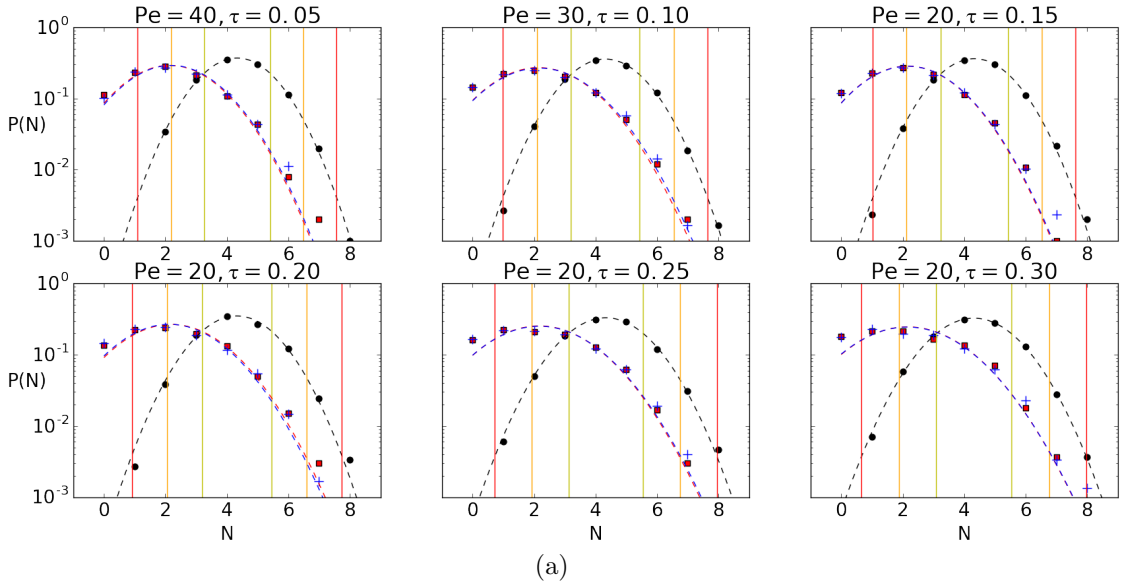

Figure 2.3: Number density probability distribution for a system with $Pe = 20, \tau = 0.1$. Points are the measured probability of finding N particles in a randomly selected probe volume of radius $1.75r_0$. The line is a Gaussian distribution with mean $\langle N \rangle = 4.32$ and variance $\sigma^2 = 1.25$. Error bars are smaller than the symbols. The density statistics are Gaussian to approximately $\pm 3\sigma$ from the mean. Number density statistics are similarly Gaussian for all of the points represented in Fig. 2.5-2.8.

## 2.4 Gaussian Density Fluctuations

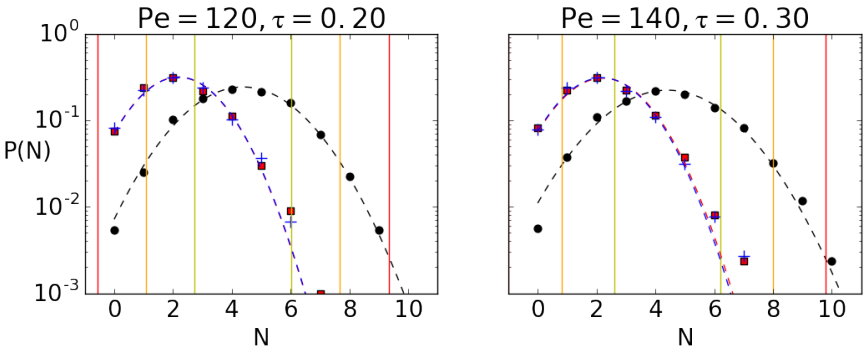
In many equilibrium liquids ranging from simple hard spheres to water, the statistics of fluctuations in the number of particles inside a small probe volume are surprisingly Gaussian [24, 85]. This property has enabled the development of quantitatively accurate theories for certain thermodynamic properties such as solvation energies [187, 188, 85]. A Gaussian description for density fluctuations also implies that changes in density due to external forces can be effectively captured within a linear response theory. In our context, linear response could help obtain relations between the driving forces and the energy supplied to the system.

To determine whether a Gaussian theory can describe the fluctuations out of equilibrium, we measured the statistics of number density fluctuations in the driven liquid in a small probe volume [24, 85]. We find that they are indeed Gaussian to a good approximation at many points on the phase diagram (below the line where the system first phase separates). One example histogram is shown in Fig. 2.3. In Fig. 2.4 , we show additional data to demonstrate

that the statistics remain Gaussian even for points that come close to the phase transition line, and for points above the second phase transition.




(a)




(b)

Figure 2.4: Number density statistics for (a) a range of points near the top of the lower mixed region (region $a$ in the phase diagram in Fig. 2.1) and (b) in the high-drive mixed region (region $c$ in the phase diagram). Black points are the total number of particles, red squares are the number of driven particles, and blue plus signs are the number of inactive particles. Curves are Gaussian fits to data of the same color. Yellow, orange, and red lines are located at $\pm 1, 2, 3$ standard deviations of the Gaussian fit from the mean. (a) demonstrates that the number density statistics remain Gaussian for values of $Pe$ and $\tau$ that are closer to the transition from mixed to separated than the values of $Pe$ and $\tau$ at which we computed the work, force correlations and diffusion in Fig. 2.3

This finding is particularly unexpected since intuition and experience give us little reason to expect that the statistics would remain Gaussian out of equilibrium – for example,

'giant' (i.e. non-Gaussian) number fluctuations have been observed in some active matter systems [125, 117, 134]. Yet, we find Gaussian fluctuations for systems where on the order of $4k_BT$ of work is done per particle per cycle.

## 2.5 Work Done by Nonconservative Forces

The Gaussian nature of density fluctuations allows us to predict how the work performed on the non-equilibrium liquid scales with $Pe$ and $\tau$. Here, and in the rest of the paper, we use the term *work* to denote the change in the energy of the system due to the action of non-conservative forces. This definition of work differs from commonly used conventions [81]. Specifically, we define the rate at which work is done on the system by the non-conservative forces as:

$$\langle \dot{w} \rangle = -\frac{1}{N} \sum_{i=1}^{N} \frac{1}{\tau} \int_0^\tau \frac{\langle \mathbf{F}_{\mathrm{c},i}(t) \rangle \cdot \mathbf{F}_{\mathrm{d},i}(t)}{\gamma} dt \tag{2.5}$$

As we will show in the subsequent sections, this definition of work proves convenient to quantify the influence of the driving forces on the structural and transport properties of the system.

Given the Gaussian nature of density fluctuations, it is reasonable to speculate that the average restoring force, $\langle \mathbf{F}_{\mathrm{c,i}}(t) \rangle$ in Eq. 2.5 above, exerted by the system in response to the driving forces, is a linear function of $Pe$. In the context of equilibrium liquids, Gaussian density fluctuations imply, within certain approximations [149, 36, 24, 96], that the coarse-grained field $\delta\rho(\mathbf{r}, t) = \rho(\mathbf{r}, t) - \bar{\rho}$, where $\bar{\rho}$ is the bulk density of the liquid, satisfies the following Gaussian equation of motion:

$$\dot{\delta\rho}(\mathbf{r}, t) = -\frac{K_G}{\gamma_G} \delta\rho(\mathbf{r}, t) + \eta_G(\mathbf{r}, t) \tag{2.6}$$

Here $K_G$ is an effective spring constant, $\gamma_G$ is an effective friction, and $\eta_G$ is an effective noise with statistics $\langle \eta_G(t) \rangle = 0$ and $\langle \eta_G(t) \eta_G(t') \rangle = 2\delta(t - t')\delta(\mathbf{r} - \mathbf{r}')D_G$, with $D_G$ an

effective diffusion. We provide an expression for $K_G$ in Appendix A1.1. We note that Eq. 2.6 does not conserve the density as it strictly should. The equivalent density-conserving EOM are available elsewhere [96]; as we will demonstrate, Eq. 2.6 is more than adequate for our purposes. Eq. 2.6 can be driven out of equilibrium by adding an extra driving term $PeF_{G,\mathrm{d}}(t)$ to it:

$$\dot{\delta\rho}(\mathbf{r},t) = -\frac{K_G}{\gamma_G}\delta\rho(\mathbf{r},t) + \frac{PeF_{G,\mathrm{d}}(t)}{\gamma_G} + \eta_G(\mathbf{r},t)\,. \tag{2.7}$$

Such a driving force has the effect of changing the local density and can mimic events where an active particle is driven into surrounding passive particles. The response of the system to this driving, $\langle\delta\rho\rangle$, scales linearly with $Pe$ [84]. In this way, due to the Gaussian density fluctuations, we anticipate that the work performed on the system will be a quadratic function of $Pe$.

The scaling of $\langle\dot{w}\rangle$ with $\tau$ can be similarly inferred using a simple scaling argument that assumes events are exponentially decorrelated with a characteristic time scale. Specifically, for values of $\tau$ much larger than the decorrelation time for fluctuations in the system, we simply anticipate that the work performed per cycle grows linearly with $\tau$. For values of $\tau$ much smaller than the correlation time, the work done per cycle can be a non-linear function of $\tau$.

These scaling trends can be combined into an expression for $\langle\dot{w}\rangle$ in terms of $Pe$ and $\tau$:

$$\langle\dot{w}\rangle \propto Pe^2\left[1 - \frac{\tau_G}{\tau}\left[1 - e^{-\tau/\tau_G}\right]\right]. \tag{2.8}$$

The scaling of the work performed in Eq. 2.8 with $Pe$ and $\tau$ follows the expected trends outlined above, with $\gamma_G/K_G \equiv \tau_G$ acting as a correlation time.

In Fig. 2.5, we show that $\langle\dot{w}\rangle$ in the atomistic simulations does indeed scale as predicted in Eq. 2.8, with $\tau_G$ as a fitting parameter. We have checked that the scaling also holds for different driving protocols including cases where the particles are not phase locked or

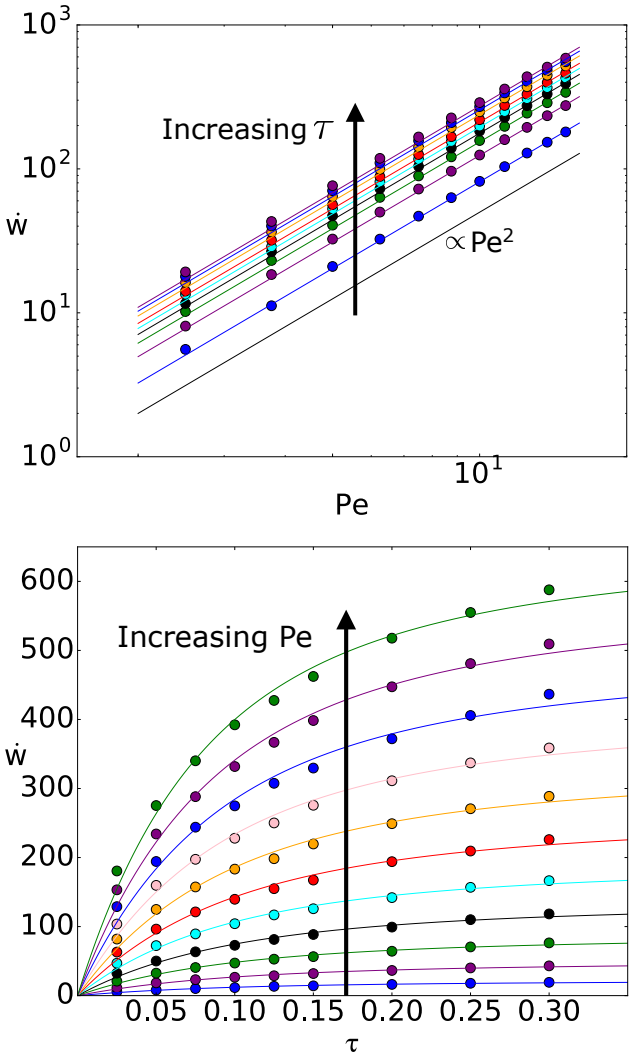



Figure 2.5: Rate of work done on the system as a function of $Pe$ (top) and $\tau$ (bottom) by the driving forces in the full many-particle simulation, for $Pe$ ranging from 2.5-15 and $\tau$ ranging from 2.5-30. Lines are fits to Eq. 2.8. Parameters $\gamma_G$ and $K_G$ were fit separately to each line; the variation in $\gamma_G$ is less than 20% between curves; the variation in $K_G$ is less than 5%. Error bars are smaller than the points.

have random phases, and in systems with unequal number fraction of driven and undriven particles (Fig. 2.7).

## 2.6 Force Fluctuations

We now demonstrate how the work done on the system by the non-equilibrium forces affects its microscopic statistics. We begin by noting that according to Floquet theory [78], the non-equilibrium steady state induced by time-periodic driving forces is also time-periodic. For the system to achieve such a time-periodic steady state, the increase in the internal energy of the system due to the total work, $w$, over each cycle has to be dissipated as heat. The average rate of heat emitted by the driven system (per particle) can be conveniently expressed in terms of the following stochastic integral, interpreted in the Stratonovich sense [156]:

$$\langle \dot{q} \rangle = -\frac{1}{N} \sum_{i=1}^{N} \frac{1}{\tau} \int_0^\tau \langle \mathbf{F}_{\mathrm{c,i}} \cdot \dot{\mathbf{r}}_{\mathrm{c,i}} \rangle dt \tag{2.9}$$

where $\mathbf{F}_{\mathrm{c,i}}$ denotes the conservative force vector for particle $i$, $\dot{\mathbf{r}}_{\mathrm{c,i}}$ is the rate of change of the position vector of particle $i$ due to the conservative and random forces only, the sum is over all of the particles in the system, and $\tau$ is the period of the driving. The average rate of heat emission $\langle \dot{q} \rangle$ should equal the negative of the average rate at which work is performed on the system, $-\langle \dot{w} \rangle$, given by Eq. 2.5.

Indeed, by adding together Eq. 2.5 and Eq. 2.9, we obtain an energy balance expression,

$$\frac{1}{\tau} \int_0^\tau \left\langle \frac{dU}{dt} \right\rangle dt = -\frac{1}{N} \sum_{i=1}^{N} \frac{1}{\tau} \int_0^\tau \langle \mathbf{F}_{\mathrm{c,i}} \cdot \dot{\mathbf{r}}_i \rangle dt = \langle \dot{q} \rangle + \langle \dot{w} \rangle = 0 \tag{2.10}$$

where $\dot{\mathbf{r}}_{\mathrm{i}}$ is the rate of change of the position vector of particle $i$ defined in Eq. 2.1. Expanding this energy balance equation, we find that $\langle \dot{w} \rangle$ can be expressed in terms of the statistics of force fluctuations,

$$\langle \dot{w} \rangle \propto \langle \mathbf{F}^2 \rangle - \langle \mathbf{F}^2 \rangle_0 \tag{2.11}$$

where

$$\langle \mathbf{F}^2 \rangle = \frac{1}{N} \sum_{i=1}^{N} \langle \mathbf{F}_{\mathrm{c,i}}^2 \rangle \tag{2.12}$$

and

$$\langle \mathbf{F}^2 \rangle_0 = -\frac{1}{2N} \sum_{i=1}^{N} \lim_{t \to 0} \langle \mathbf{F}_{\mathrm{c,i}}(t) \cdot \eta_i(0) \rangle \tag{2.13}$$

In Appendix A1.2, we show that in equilibrium the two terms on the right hand side of Eq. 2.11 are equal as a consequence of the fluctuation dissipation theorem. In the above equations, $\left\langle \mathbf{F}^2 \right\rangle_0$ captures the response of the non-equilibrium liquid following a random perturbation while $\left\langle \mathbf{F}^2 \right\rangle$ describes the force fluctuations on a tagged particle in its non-equilibrium steady state. Out of equilibrium, the work performed is positive and the relation between fluctuations and response breaks down. In this case, the difference between fluctuations and response is predicted exactly by flux of heat, or alternately the rate at which work is performed, as illustrated in Eq. 2.11. We note that this expression is completely insensitive to the choice of non-conservative driving forces in Eq. 2.1. Eq. 2.11 is an instantiation of the Harada-Sasa relation [81, 80]. Here we have obtained it for our non-equilibrium system following simple thermodynamic arguments.

Latly, we note that the term $\langle \mathbf{F}^2 \rangle_0$, defined above in Eq. 2.13, appears naturally in the Stratonovich interpretation of energy fluxes and is related to $\langle \dot{w} \rangle$. In an Itô formalism, we analogously can obtain an expression for rate of work:

$$0 = \frac{1}{\tau} \int \left\langle \frac{dU}{dt} \right\rangle = -\langle \mathbf{F}^2 \rangle + K_B T \langle \nabla^2 U \rangle + \langle \dot{w} \rangle \tag{2.14}$$

We can see that $\langle \mathbf{F}^2 \rangle_0 = K_B T \langle \nabla^2 U \rangle$, which is the result we got in Appendix A1.2 as well, showing that our relationship between work and force fluctuations is equivalent in the

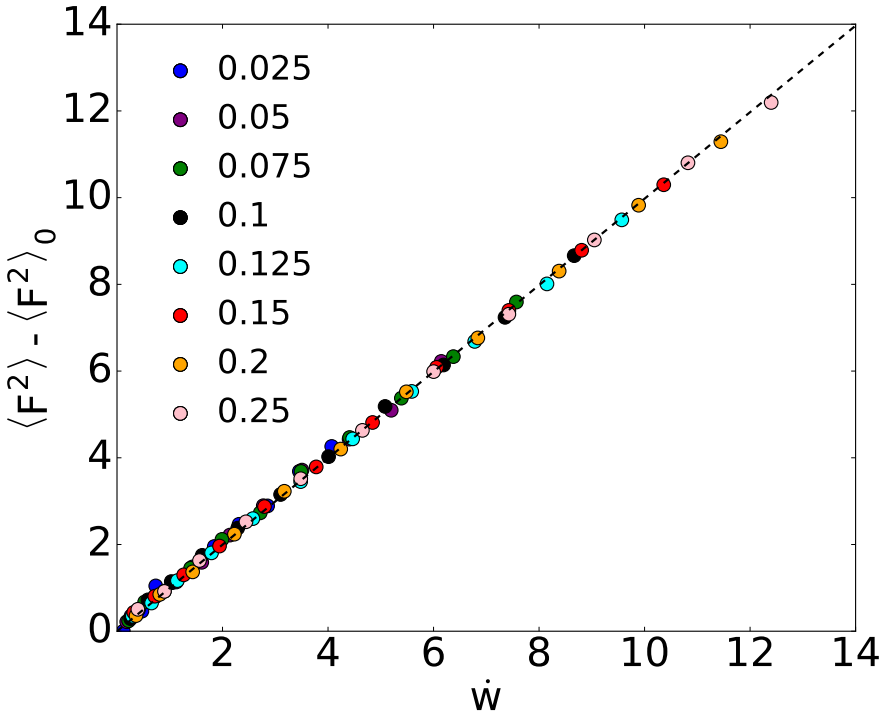



Figure 2.6: Deviation from fluctuation-dissipation theorem as a function of the rate of work done per particle. Colors indicate values of $\tau$. The data collapses onto the line predicted by Eq. 2.11. Error bars are smaller than the points.

Itô and the Stratonovich sense.

We find numerically that the work performed in the many-particle system is indeed related to the breakdown of the equilibrium relation between fluctuations and response (Fig. 2.6). We verified Eq. 2.11 for a wide range of amplitudes and time periods of the driving force in the homogeneous part of the phase diagram. We have also verified that this result holds for driving forces in which the particles are not phase locked, have random phases, and in systems with unequal number fraction of driven and undriven particles (Fig. 2.7). The relation also holds when applied separately to driven and undriven particles in all of these cases – in other words, the work performed on average due to the driven particles predicts the change in the force fluctuations of the undriven particles.

We now study the implications of this result for the diffusion constant of the non-equilibrium system.

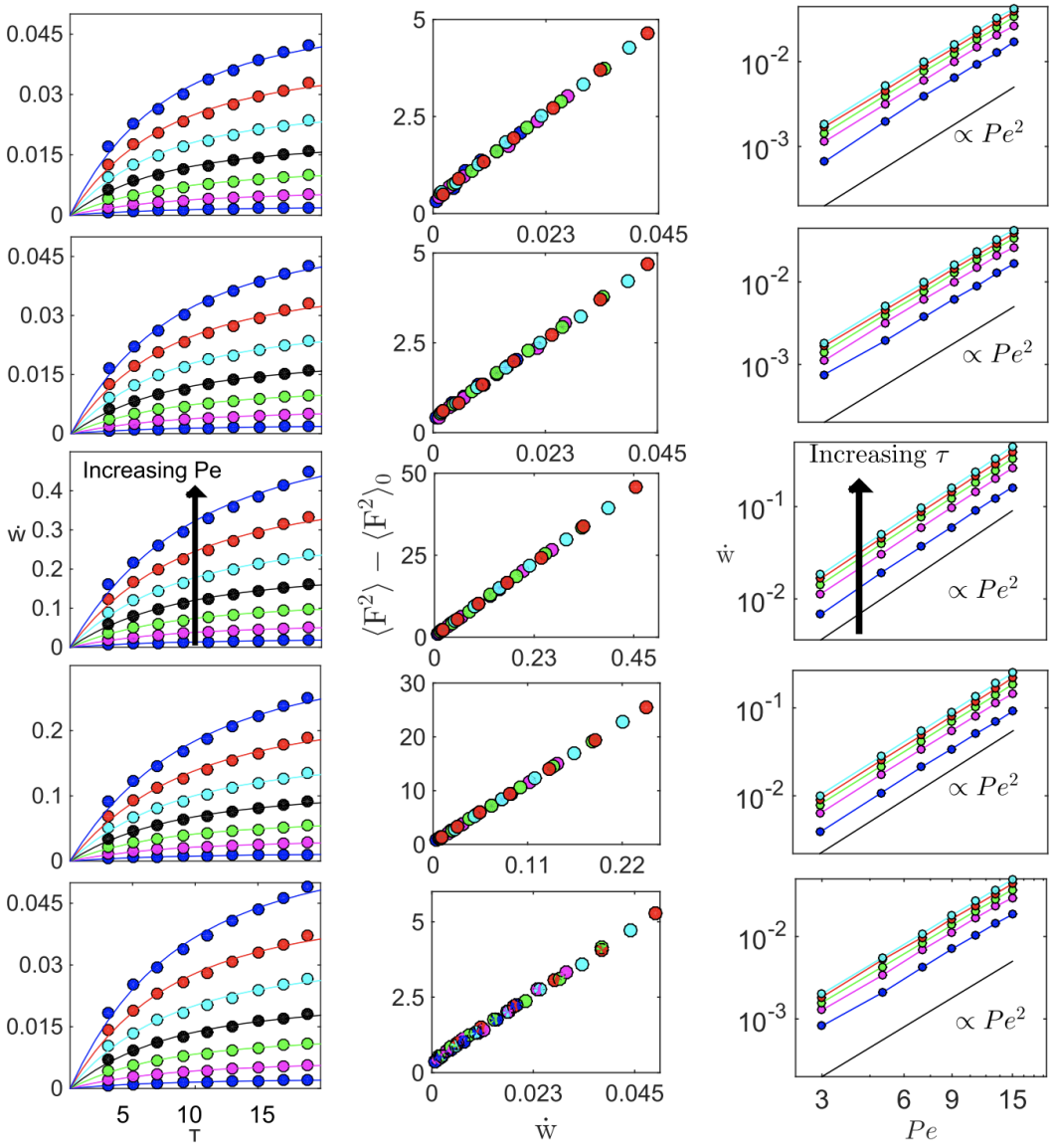



Figure 2.7: We show the scaling of the work with $\tau$ (left), force fluctuations (center) and $Pe^2$ (right). Top two rows correspond to in-phase driving, but the fraction of active particles is, in this order, 0.9 and 0.1. The bottom three rows correspond to active particles with initial phases drawn from a uniform distribution from 0 to $2\pi$ and fractions of active particles 0.9, 0.5 and 0.1, in this order. Error bars are smaller than the points. In the scaling of $\langle F\rangle^2 - \langle F^2\rangle_0$ with the work, the $\tau$ values sampled were 3, 7, 11, 15 and 19. We note that in Fig. 2.5, renormalized units ($\gamma = 1$, which renormalizes $\tau$ by 1/100) are used, whereas in this plot the real simulation units ($\gamma = 100$) are used. In the left column, the $\tau_G$ parameter in Eq. 2.8 can vary between the five investigated systems, but for each system the variation in this parameter is small. Maximum variation in $\tau_G$ was around 10%.

## 2.7 Diffusion Coefficient

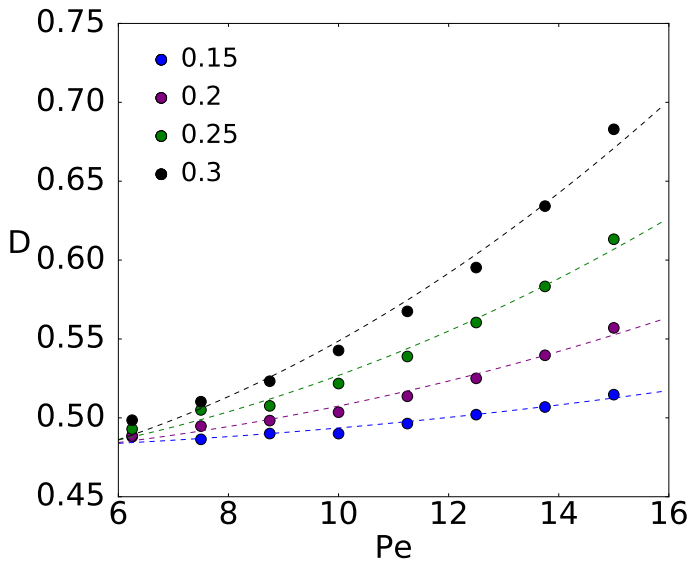



Figure 2.8: Scaling of the diffusion coefficient of a tagged particle in the liquid, $D$, with the amplitude of the driving force, $Pe$. The fit is of the form $D = D_{eq} - aPe + bPe^2$, where $a$ and $b$ depend on $\tau$. $D_{eq}$ is the equilibrium diffusion constant. Colors indicate values of $\tau$. Error bars are smaller than the points.

In Brownian dynamics, the diffusion can be written as follows in terms of force autocorrelation functions:

$$D - D_0 = \frac{1}{d}\int_0^\infty \left[\left\langle \frac{\mathbf{F}(0)\cdot\mathbf{F}(t)}{\gamma^2}\right\rangle + \left\langle \frac{\boldsymbol{\eta}(0)\cdot\mathbf{F}(t)}{\gamma}\right\rangle\right] dt\,, \tag{2.15}$$

where $d$ denotes the dimensionality of the system. It is reasonable to ask whether the change in the force correlations that accompanies a breakdown of the fluctuation dissipation relation, Eq. 2.11, affects the diffusion coefficient of the driven liquid.

From simulations, we indeed find that the diffusion constant gets renormalized due to the driving forces and increases with the amplitude of the driving force after a transient region.

The renormalization of the diffusion constant is very well-described by the functional form

$$D - D_{eq} = -a(\tau)Pe + b(\tau)Pe^2 \tag{2.16}$$

(Fig. 2.8). Understanding the basis for the renormalization of $D$ due to the non-conservative forces is important [17, 77] – as we will show, in the present context it can help explain how the energy input due to the non-conservative forces can drive phase separation.

In order to qualitatively understand the observed dependence of $D$ on the non-conservative forces, we consider a minimal model of a tracer particle diffusing in a fluid according to the equation of motion

$$\dot{\mathbf{r}} = h\tilde{\mathbf{F}}(\mathbf{r}) + \tilde{\boldsymbol{\eta}}(t), \tag{2.17}$$

where $\tilde{\mathbf{F}}(\mathbf{r})$ is a spatially dependent force that can be designed to model the forces acting on a tagged particle in the liquid, $h$ is a parameter that tunes the coupling between the fluid and the tracer particle, and $\tilde{\boldsymbol{\eta}}(t)$ is a Gaussian $\delta$-correlated white noise. In order to ensure no net drift, we constrain $\int \tilde{\mathbf{F}}(\mathbf{r})d\mathbf{r} = 0$. We imagine sampling over many realizations of the force $\tilde{\mathbf{F}}(\mathbf{r})$ from a distribution in order to model the forces exerted by the fluid on a tracer particle. In the liquid considered in the previous sections, the statistics of force fluctuations on a tagged particle satisfy $\langle \mathbf{F}^2 \rangle = \langle \mathbf{F}^2 \rangle_0$ when the system is in equilibrium. Equilibrium dynamics in Eq. 2.17 are achieved whenever $\tilde{\mathbf{F}}(\mathbf{r})$ can be expressed as a gradient of a potential, $\tilde{\mathbf{F}}(\mathbf{r}) = -\nabla\tilde{U}(\mathbf{r})$. In such cases, it can be demonstrated that the equivalent relation $\langle \tilde{\mathbf{F}}^2 \rangle = \langle \tilde{\mathbf{F}}^2 \rangle_0$ holds, where the averages $\langle \dots \rangle$ are taken both over the statistics of the random noise $\tilde{\boldsymbol{\eta}}(t)$ and over many realizations of the force. The system can be driven out of equilibrium by ensuring that the force in Eq. 2.17 has a non-conservative component, $\tilde{\mathbf{F}}(\mathbf{r}) = \tilde{\mathbf{F}}_{\mathrm{c}}(\mathbf{r}) + Pe\tilde{\mathbf{F}}_{\mathrm{d}}(\mathbf{r})$, where $\tilde{\mathbf{F}}_{\mathrm{c}}(\mathbf{r}) = -\nabla\tilde{U}(\mathbf{r})$, $\tilde{\mathbf{F}}_{\mathrm{d}}(\mathbf{r}) = \nabla \times \tilde{\mathbf{A}}(\mathbf{r})$. Like in the previous sections, $Pe$ tunes the magnitude of the non-conservative force.

A rough qualitative mapping between the forces in the minimal model and the forces

in the atomistic liquid can be obtained using the following reasoning. The non-equilibrium driving in the atomistic liquid alters the pair correlation function between a driven and an undriven particle in two ways. First, the scaling of $\langle w \rangle$ with $Pe^2$ reveals that due to the driving forces the pair correlation function develops an anisotropic, time-dependent component that is also proportional to $Pe$. The non-conservative force introduced in the minimal model, $Pe\tilde{\mathbf{F}}_{\mathrm{d}}$, is meant to qualitatively simulate the effect of the anisotropic component of the atomistic pair correlation function. Second, the isotropic component of the pair correlation function can also be affected by the driving forces (Fig. A1.1). The potential $\tilde{U}$ in the minimal model is meant to simulate the effects of the isotropic component of the steady state pair correlation function.

In the presence of the non-conservative forces in the minimal model, as in the many-particle driven liquid, a breakdown of the fluctuation dissipation relation is predicted by the total amount of entropy dissipated by the system [80]. Specifically, we use a perturbation theory [39, 55] (detailed in Appendices A1.3- A1.5) to show that to $O(h^2)$,

$$\langle \dot{\sigma} \rangle = \langle \tilde{\mathbf{F}}^2 \rangle - \langle \tilde{\mathbf{F}}^2 \rangle_0 = \frac{Pe^2 h^2}{V} \int \langle \tilde{\mathbf{F}}_{\mathrm{d}}(\mathbf{q}) \cdot \tilde{\mathbf{F}}_{\mathrm{d}}(-\mathbf{q}) \rangle d\mathbf{q} \tag{2.18}$$

where $\langle \dot{\sigma} \rangle$ denotes the average rate of entropy dissipation, analogous to $\langle \dot{w} \rangle$ in the many-body liquid.

For this minimal model, we obtained expressions for the diffusion constant in Eq. 2.17 to quadratic order in the parameters $h$ and $hPe$:

$$D - D_0 = \alpha \frac{h^2}{V} \left( Pe^2 \int \frac{\langle \tilde{\mathbf{F}}_d(\mathbf{q}) \cdot \tilde{\mathbf{F}}_d(-\mathbf{q}) \rangle}{|\mathbf{q}|^2} d\mathbf{q} - \int \langle \tilde{U}^2(\mathbf{r}) \rangle d\mathbf{r} \right) \tag{2.19}$$

where $\alpha = D_0/(d(k_B T)^2)$, $d$ is the dimension, and we set $\langle \tilde{U} \rangle = 0$ without loss of generality so that $\langle \tilde{U}^2 \rangle$ is simply the variance of energy fluctuations. In instances where the spectrum of force fluctuations is strongly peaked at a particular wave vector $\mathbf{q}^*$, the diffusion constant

can be simply related to the average entropy dissipation rate:

$$D - D_0 = \alpha \frac{\langle \dot{\sigma} \rangle}{|\mathbf{q}^*|^2} - \alpha h^2 \langle \tilde{U}^2 \rangle. \tag{2.20}$$

The dynamics of our driven liquid, specified by Eq. 2.1, are substantially more complicated than the minimal model considered. Nonetheless, the expressions in Eq. 2.19 and Eq. 2.20 provide useful insight. By using the variance $\langle \tilde{U}^2 \rangle$ as a measure of the microscopic environment around a tagged particle in the many-body driven liquid, and by associating the entropy production rate $\langle \dot{\sigma} \rangle$ with the rate of work performed $\langle \dot{w} \rangle$ in the many-body liquid, Eq. 2.20 demonstrates how the non-conservative forces can modify the diffusion properties of a particle in the fluid. In particular, the minimal model predicts that the diffusion constant can increase as $Pe^2$ due to the energy dissipation from non-conservative forces. To probe how the microscopic environment around a tagged particle changes due to the driving forces, we measured the average energy of a driven tracer particle in a fluid, $\langle u \rangle = \int g(\mathbf{r}) u(\mathbf{r}) d\mathbf{r}$, where $g(r)$ is the steady-state two-body pair correlation function of a tagged active particle in the driven liquid with 10% active particles (Appendix Fig. A1.1). The average energy increases with $Pe$. Effectively, the active driven particles are sampling configurations characteristic of an equilibrium state with higher temperature. In other words, the variance $\langle \tilde{U}^2 \rangle$, used in our minimal model should increase with $Pe$. Taken together, these qualitative arguments suggest that the increase in the diffusion constant in our driven liquid is caused mainly by the work performed by the non-conservative forces.

Finally, we investigate how the increase of the diffusion constant due to the non-conservative forces renders the diffusion constant composition dependent. Specifically, unmixed regions with particles either being all driven or all un-driven effectively have equilibrium dynamics and diffusion properties since the nonconservative forces do not induce any collisions in such regions. The diffusion constant in regions with mixed compositions can be enhanced (as described above) due to collisions induced by the non-conservative forces. As we demonstrate

in Appendix A1.6 such a composition dependent diffusion constant can drive a transition from the low drive mixed phase (a) to the phase separated region (b) in the phase diagram in Fig. 2.1 when

$$\lambda = \frac{b(\tau)Pe^2}{D_{eq} + b(\tau)Pe^2} \tag{2.21}$$

where $b(\tau)$ is defined in Eq. 2.16 and $\lambda$ is a system-specific constant defined in Appendix A1.6, which in our atomistic simulations ends up being approximately 0.5. Eq. 2.21 provides us with a lower bound on the value of $Pe$ where phase separation first occurs.

Indeed, we find that the shape of the phase transition curve in Fig. 2.1 (solid green line) is well described by $b(\tau)Pe^2 = c \approx D_{\text{eq}}$, where $c$ is a constant, in accordance with Eq. 2.21. Further, our numerical simulations show that $b(\tau)Pe^2 \propto \langle w \rangle$ to a very good approximation (Fig. 2.9), and we also find that the shape of the phase transition curve is well-described by $\langle w \rangle \propto c$ (purple line in Fig. 2.1). The results in Eqs. 2.19, 2.20 and 2.21 show how the increase in diffusion due to energy dissipation can control the phase properties of our non-equilibrium liquid. Numerically, we find that the phase transition properties of our driven liquid are effectively controlled by $\langle w \rangle$, the energy injected into it per cycle by the non-equilibrium driving forces.

## 2.8 Conclusions

From the rich physics of non-equilibrium materials, and in particular of active particles with rotating dynamics, it is clear that dissipation plays an important role in modifying the structural and dynamical properties of the steady states of these systems. Here, in the context of a class of systems with rotating dynamics, we have identified how the rate of work done by the external forces renormalizes force fluctuations. Using simplified descriptions of density fluctuations, we were also able to model how the work performed in this many body system depends on the non-equilibrium forces. Finally, using a minimal model, we explained

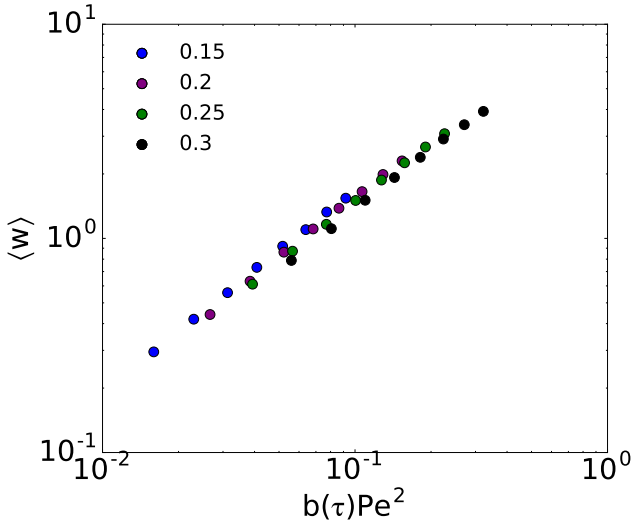



Figure 2.9: Numerical connection between diffusion and work: The last term in $D = D_{eq} - a(\tau)Pe + b(\tau)Pe^2$ is approximately proportional to the average amount of work done per cycle on the system. Data points are for the same values of $Pe$ and $\tau$ included in Fig. 2.8.

the observed enhancement of the diffusion due to the non-equilibrium driving forces and proposed a relation between diffusion and dissipation. The analysis based on the minimal model also helped explain the observed dependence of the phase behavior on the magnitude $Pe$ of the driving force. These results demonstrate how the material properties of driven liquids, which are experimentally realizable as colloids, can be tuned simply by violating time reversal symmetry and controlling the amount of energy put in to the system.

# CHAPTER 3

# ENERGY DISSIPATION, STRUCTURE, TRANSPORT AND BIASED ENSEMBLES IN ACTIVE MATTER

The material in this chapter is adapted with permission from the following publication: [180] Laura Tociu et al. "How Dissipation Constrains Fluctuations in Nonequilibrium Liquids: Diffusion, Structure, and Biased Interactions". *Phys. Rev. X*, 9 (2019).

## 3.1 Introduction

The work described in this chapter is concerned with a theoretical framework to precisely relate energy dissipation due to nonequilibrium forces to density fluctuations and transport in generic active matter. To elucidate the role of nonequilibrium forces in materials, it is crucial to examine how dissipation affects the emerging dynamics and structure. While equilibrium features are well established, progress in controlling systems with sustained dissipation has been hampered by a lack of general principles [22, 160, 132, 61, 122, 126, 131]. In this context, minimal models of active and driven systems provide analytically and numerically tractable test beds to investigate the interplay between dissipation and material properties far from equilibrium [112, 76, 8, 93, 59]. They have illustrated, for instance, how nonequilibrium driving can induce phase transitions and excite novel collective responses in soft media [189, 173, 76, 132, 200]. Recent theoretical work has proposed extending equilibrium concepts to active media, such as the definition of pressure [175, 160], to rationalize their phenomenology [161, 162]. Others have striven to obtain stationary properties of active matter through perturbation close to equilibrium [51, 61, 195], inspired by other approaches on driven systems [118, 95, 108, 107].

To investigate how dissipation controls emerging behavior, yet another approach has focused on introducing a bias in dynamical ensembles. Using large deviation techniques,

trajectories are conditioned to promote atypical realizations of the dynamics [182, 90]. Such techniques have been used, for instance, to investigate the role of dynamical heterogeneities in glassy systems [64, 82, 138, 165, 14, 106, 127] and soliton solutions in high-dimensional chaotic chains [174, 99]. More recently, it has been shown that changing dissipation, through a dynamical bias, strongly affects the internal transport and the density fluctuations of nonequilibrium liquids [19, 128], thus confirming that controlling dissipation is indeed a fruitful route to tailoring material properties. In spite of these advances, anticipating the emergent dynamics and structure of biased nonequilibrium systems is still challenging in the presence of many-body interactions [27, 90], so that precise control has remained elusive so far in this context. Consequently, any generic principle rationalizing spatial organization in terms of dissipation is still lacking.

In this chapter, we explore how dissipation affects the dynamics and structure of many-body diffusive systems. First, we consider in Sec. 3.2 two types of assemblies of Brownian particles: one in which only a subset is driven by an external force, and one in which a subset of the particles experience an internal active force. We first focus on instances where the fraction of driven particles is less than the fraction of undriven particles, so that driven and undriven particles are respectively referred to as *tracer* and *bath* particles. Using the diffusion coefficient of a tagged tracer particle and the density correlations between tracer and bath particles, we connect dissipation to liquid properties. In contrast with the work in the previous chapter, our prediction for diffusion follows from a systematic coarse-graining with explicit dependence in terms of microscopic details [36, 40, 39].

Next, importantly, we put forward a generic relation between density correlations and dissipation valid for an arbitrary driving force: this is our first main result. We demonstrate that this result holds both for fluids in which a fraction of the particles are driven by a fixed external drive and for fluids in which either a fraction of the liquid or the entire liquid is driven by an internal noise, analogous to the driving used in model active matter systems.

This result opens the door to estimating dissipation directly from the liquid structure, in contrast with previous approaches based either on perturbing the system [81, 119, 57, 186, 1] or on analyzing trajectories and currents in phase space [7, 67, 148, 116, 103]. We illustrate this with numerical simulations for which dissipation is quantified by the deviation from equilibrium tracer-bath correlations. Using these results as a basis, we also show how various aspects of the pair correlation function of a nonequilibrium liquid are effectively constrained by the energy dissipation. Altogether, this set of results clarify how nonequilibrium forces affect the transport and structure of the liquid, thus showing how liquid properties can be modified at the cost of energy dissipation.

Motivated by these results, and to provide concrete intuition for how particular configurations can be stabilized by nonequilibrium forces, we next investigate in Sec. 3.6 the emerging structure of Brownian particles subject to a dynamical bias. The explicit form of the bias is inspired by the results of Sec. 3.5 connecting dissipation to many-body interactions. Using analytical calculations and numerical simulations based on the cloning algorithm [66, 174, 86, 129, 143, 94, 15], we show that biased sampling trajectories can be used to renormalize any specific interparticle interaction in a multi-component liquid. The rare noise fluctuations sampled with dynamical bias effectively drive the system away from typical behavior [64, 82, 90, 138, 165, 14, 27, 106, 127]. Such noise realizations can then serve as proxies of how to control the dynamics by applying an external force with complex protocols. We also illustrate the generality of our ideas by considering an assembly of aligning self-propelled particles [52]. Specifically, we show how biased energy flows can renormalize interactions between particles and stabilize a flocking transition. Overall, our results lay the groundwork for precise control of the emerging structure and collective dynamics in many-body diffusive nonequilibrium systems.

## 3.2 Model and Methods

We consider interacting Brownian particles where a specific set of particles $\Omega$ is driven by a non-conservative force $\mathbf{F}_{\mathrm{d},i}$:

$$\gamma\dot{\mathbf{r}}_i = \delta_{i\in\Omega}\mathbf{F}_{\mathrm{d},i} - \nabla_i \sum_j v(\mathbf{r}_i - \mathbf{r}_j) + \boldsymbol{\xi}_i, \tag{3.1}$$

where $\delta_{i\in\Omega} = 1$ if $i \in \Omega$ and $\delta_{i\in\Omega} = 0$ otherwise. The driven particles belonging to the set $\Omega$ are referred to as *tracers*, and others as *bath* particles. The potential $v$ yields a conservative force through which the particles interact with each other. The fluctuating term $\boldsymbol{\xi}_i$ is a zero-mean Gaussian white noise with correlations $\langle \xi_{i\alpha}(t)\xi_{j\beta}(0)\rangle = 2\gamma T\delta_{ij}\delta_{\alpha\beta}\delta(t)$, where $\gamma$ and $T$ respectively denote the damping coefficient and the bath temperature, with the Boltzmann constant set to unity ($k_{\mathrm{B}} = 1$).

In what follows, we consider two types of drive: (i) an external force following the same deterministic protocol for all driven particles, and (ii) an internal force given by a noise term independent for each driven particle. Building on the driving forces in the previous chapter, we take for drive (i) a time-periodic protocol given in two dimensions by

$$\mathbf{F}_{\mathrm{d}}(t) = f\big[\sin(\omega t)\hat{\mathbf{e}}_x + \cos(\omega t)\hat{\mathbf{e}}_y\big], \tag{3.2}$$

where $f$ and $\omega$ are respectively the amplitude and the frequency of the drive, so that the drive persistence reads $\tau = 2\pi/\omega$. The relative strength of the drive is given by the Péclet number $\mathrm{Pe} = \sigma f/T$, where $\sigma$ is the typical particle size [76, 93]. In the absence of interactions ($v = 0$), the average position of driven tracers follows a periodic orbit, describing a circle in two dimensions. In contrast, drive (ii) corresponds to a random self-propulsion as is often considered in active liquids [54, 144, 109]. Specifically, we use a set of zero-mean Gaussian

noises with correlations

$$\langle F_{\mathrm{d},i\alpha}(t)F_{\mathrm{d},j\beta}(0)\rangle = \delta_{ij}\delta_{\alpha\beta}\frac{f^2}{d}\mathrm{e}^{-|t|/\tau}, \tag{3.3}$$

where $d$ is the spatial dimension. The parameters $f$ and $\tau$ respectively control the amplitude and the persistence of fluctuations.

Interestingly, the active force with correlations (Eq. 3.3) can be obtained from a generalized version of the deterministic force in Eq. 3.2, where each particle $i$ is now subjected to an independent drive. The period of the orbit is determined by a series of $n$ oscillators with identical frequencies for all particles, yet independent amplitudes:

$$\mathbf{F}_{\mathrm{d},i}(t) = \frac{f}{\sqrt{nd}}\sum_{a=1}^{n}\left[\mathbf{A}_{ai}\cos(\omega_a t) + \mathbf{B}_{ai}\sin(\omega_a t)\right]. \tag{3.4}$$

The essential ingredient of the mapping into active force is to implement disorder in the drive. This is done by taking the oscillator amplitudes as uncorrelated zero-mean Gaussian variables with unit variance:

$$\langle A_{ai\alpha}A_{bj\beta}\rangle_{\mathrm{d}} = \delta_{ab}\delta_{ij}\delta_{\alpha\beta} = \langle B_{ai\alpha}B_{bj\beta}\rangle_{\mathrm{d}}, \tag{3.5}$$

where $\langle\cdot\rangle_{\mathrm{d}}$ denotes an average over the disorder. It follows that $\mathbf{F}_{\mathrm{d},i}$ is also a Gaussian process with zero mean and correlations given by:

$$\langle F_{\mathrm{d},i\alpha}(t)F_{\mathrm{d},j\beta}(0)\rangle_{\mathrm{d}} = \delta_{ij}\delta_{\alpha\beta}\frac{f^2}{nd}\sum_{a=1}^{n}\cos(\omega_a t). \tag{3.6}$$

In the limit of a large number of oscillators ($n \gg 1$), we express these correlations in terms of the density of driving frequencies $\phi$ as

$$\langle F_{\mathrm{d},i\alpha}(t)F_{\mathrm{d},j\beta}(0)\rangle_{\mathrm{d}} \underset{n\gg 1}{=} \delta_{ij}\delta_{\alpha\beta}\frac{f^2}{d}\int \phi(\omega')\mathrm{e}^{\omega'|t|}\frac{\mathrm{d}\omega'}{2\pi}. \tag{3.7}$$

This establishes that, in the limit of many oscillators, the deterministic drive in Eq. 3.4 with disordered amplitude is equivalent to a noise term with spectrum $\phi$. In particular, by choosing $\phi(\omega') = 2\tau/\big[1+(\omega'\tau)^2\big]$, the drive correlations (Eq. 3.7) reproduce exactly the ones of the random force in Eq. 3.3.

In Sec. 3.3 where we discuss the relevance of this mapping, we employ a custom code of molecular dynamics, based on finite time difference, to perform the simulations in a two-dimensional box $10^2\sigma \times 10^2\sigma$ with periodic boundary conditions. The time step is $\delta t = 10^{-4}$ and the initial condition is homogeneous. We use $\rho_0 = 0.7$, $T = 0$, $\gamma = 1$, $f = 3 \times 10^{-2}$, $\tau = 10^3$, $v_0 = 5$, $n = 10^2$ and $\sigma = 1$. We use the potential $v(\mathbf{r}) = v_0(1 - |\mathbf{r}|/\sigma)^2\Theta(\sigma - |\mathbf{r}|)$, where $\Theta$ denotes the Heaviside step function, which sets purely repulsive interactions.

In Secs. 3.4 and 3.5 where we relate energy dissipation to diffusion and structure, respectively, we perform numerical simulations of the dynamics in Eq. 3.1 using a custom build of the LAMMPS simulation package in a two-dimensional box $10^2\sigma \times 10^2\sigma$, where $\sigma$ is the particle diameter, with periodic boundary conditions. Our custom build implements overdamped Langevin equations of motion with finite time difference, and simply utilizes the efficient force computation routines that are built as a part of the molecular dynamics package. The system is first relaxed for $10^3$ conjugate gradient descent steps, and later equilibrated during $50\tau$. We evaluated average values over 10 independent trajectories with duration $150\tau$. The density pair correlations were constructed using 10 independent trajectories each with duration $50\tau$. We performed error analysis from the independent simulations and obtained negligible errors for all the data in Figs. 3.2-3.4. The time step is $5 \times 10^{-4}$ and the bin size for computing the pair correlations is $0.01\sigma$. We performed simulations at other values of the time step $\{10^{-4}, 10^{-5}\}$ and of the bin size $\{5 \times 10^{-3}\sigma, 2 \times 10^{-2}\sigma\}$ to confirm that our calculations of the $\alpha$ coefficients are well converged. Parameter values: $T = 1$, $\gamma = 100$, $v_0 = 1$, $\sigma = 1$, $\rho = 0.45$. Interactions are set by the WCA potential $v(\mathbf{r}) = 4v_0\big[(\sigma/|\mathbf{r}|)^{12} - (\sigma/|\mathbf{r}|)^6\big]\Theta(2^{1/6}\sigma - |\mathbf{r}|)$ [193].

In Sec. 3.6 where we explore biased ensembles and auxiliary dynamics, a custom code of molecular dynamics, based on finite time difference, is used to perform the simulations in a two-dimensional box $10\sigma \times 10\sigma$ with periodic boundary conditions. We bias the pair potential between 8 blue particles and 16 red particles. To sample the biased ensemble, we use the cloning algorithm described in Appendix A of [129]. The time interval for cloning is $\Delta t = 10\delta t$ and the number of clones is 1600. The time step is $\delta t = 10^{-4}$, the initial relaxation time is $10^4 \Delta t$, and the total simulation time is $10^6 \Delta t$. Parameter values that we used are $T = 1$, $\gamma = 1$, $v_0 = 4$ and $\sigma = 1$.

When simulating Vicsek particles using the equations of motion in Eq. 3.39, we employ a custom code of molecular dynamics based on finite time difference to perform the simulations. $N = 128$ particles are simulated in a two dimensional box of size $4\sigma \times 4\sigma$ with periodic boundary conditions. To sample the biased ensemble, we use the cloning algorithm described in Appendix A of [129]. The time interval for cloning is $\Delta t = 10\delta t$ and the number of clones is 200. The time step is $\delta t = 10^{-3}$, and the total simulation time is $500\Delta t$. We used the following parameters: $V_0 = 2$, $\mu_{\rm r} = 2$, $\rho_0 = 8$, $\sigma = 1$.

## 3.3 Deterministic versus Active Driving

This Section describes a framework to connect active and deterministic driving devised by my collaborator Étienne Fodor. To illustrate the relevance of this active to deterministic mapping that we began to describe in the previous section, we simulate numerically the many-body dynamics in Eq. 3.1 where every particle is subjected to a disordered drive of the form in Eq. 3.4. To implement numerically the disorder in driving, it is sufficient to sample the amplitudes $\{\mathbf{A}_{ai}, \mathbf{B}_{ai}\}$ and frequencies $\{\omega_a\}$ at initial time, according to the statistics indicated in the Methods section. In the regime of high persistence $\tau$ and large average density $\rho_0$, we observe the spontaneous formation of clusters up to a complete phase separation at large time, see Fig. 3.1. This is analogous to the motility-induced phase

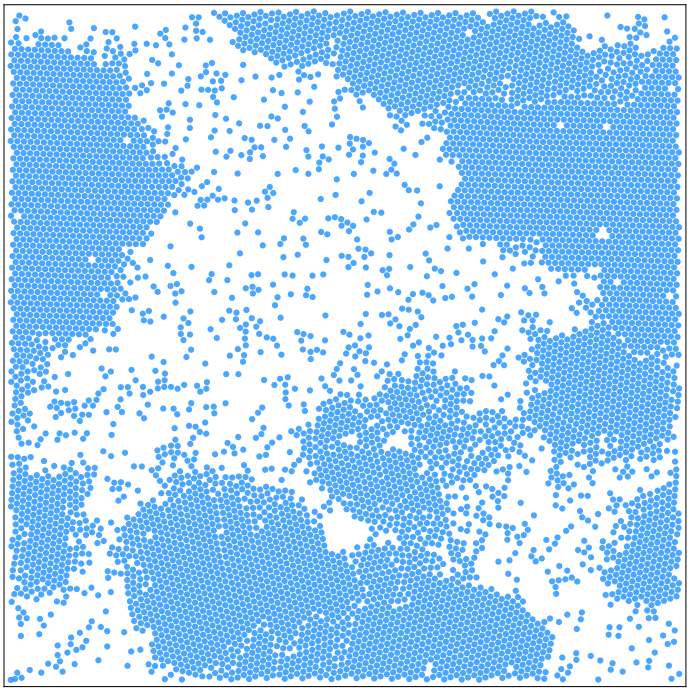


Figure 3.1: Snapshot of particles subjected to a disordered drive, from simulations described in the Methods section. A phase separation emerges which is analogous to the motility-induced phase separation of active particles [173, 22].

separation commonly reported in standard models of active particles [173, 22]. Interestingly, it appears in our case even in the absence of fluctuations ($T = 0$), namely for a purely deterministic set of equations.

In short, we thus demonstrate that the disordered drive alone reproduces the emerging physics of active systems. This important result bridges the gap between two main classes of nonequilibrium liquids, where the driving force stems from either a deterministic protocol or a random noise. In what follows, we obtain analytic and numerical results for both drives to illustrate the broad applicability of our framework, ranging from systems driven by deterministic fields to active matter systems.

## 3.4 Dissipation and Diffusion

To connect tracer diffusion with dissipation, we first describe the dynamics of undriven particles in terms of a coarse-grained variable. Using standard techniques, the dynamics of the density field $\rho(\mathbf{r},t) = \sum_{i\notin\Omega}\delta[\mathbf{r}-\mathbf{r}_i(t)]$ can be written as a non-linear Langevin equation [36]. In the regime of weak interactions, the density fluctuations $\delta\rho(\mathbf{r},t) = \rho(\mathbf{r},t) - \rho_0$ around the average density $\rho_0$ are Gaussian and captured by the following Hamiltonian [24, 39, 96]:

$$\begin{aligned}\mathcal{H} = &\frac{T}{2}\int \delta\rho(\mathbf{r})K(\mathbf{r}-\mathbf{r}')\delta\rho(\mathbf{r}')\mathrm{d}\mathbf{r}\mathrm{d}\mathbf{r}' \\ &+ \int\sum_{i\in\Omega} v(\mathbf{r}-\mathbf{r}_i)\rho(\mathbf{r})\mathrm{d}\mathbf{r},\end{aligned} \tag{3.8}$$

where $K(\mathbf{r}) = \delta(\mathbf{r})/\rho_0 + v(\mathbf{r})/T$. Note that density fluctuations remain generally Gaussian even for a homogeneous active liquid [54]. The conserved density dynamics reads

$$\begin{aligned}\frac{\partial\delta\rho(\mathbf{r},t)}{\partial t} = &D_\mathrm{G}\nabla^2\int K(\mathbf{r}-\mathbf{r}')\delta\rho(\mathbf{r}',t)\mathrm{d}\mathbf{r}' \\ &+ \frac{1}{\gamma_\mathrm{G}}\nabla^2\sum_{i\in\Omega} v(\mathbf{r}-\mathbf{r}_i(t)) + \nabla\cdot\mathbf{\Lambda}(\mathbf{r},t),\end{aligned} \tag{3.9}$$

where $D_\mathrm{G} = \rho_0 T/\gamma$ and $\gamma_\mathrm{G} = \gamma/\rho_0$ are respectively the field diffusion coefficient and the field damping coefficient. The term $\mathbf{\Lambda}$ is a zero-mean Gaussian white noise with correlations $\langle\Lambda_\alpha(\mathbf{r},t)\Lambda_\beta(\mathbf{r}',t')\rangle = 2D_\mathrm{G}\delta_{\alpha\beta}\delta(\mathbf{r}-\mathbf{r}')\delta(t-t')$.

Owing to the linearity of the density dynamics in Eq. 3.9, it can be readily written in Fourier space $\delta\rho(\mathbf{q},t) = \int\rho(\mathbf{r},t)\mathrm{e}^{-\mathrm{i}\mathbf{q}\cdot\mathbf{r}}\mathrm{d}\mathbf{r}$ as

$$\begin{aligned}\frac{\partial\delta\rho(\mathbf{q},t)}{\partial t} = &-|\mathbf{q}|^2 D_\mathrm{G}K(\mathbf{q})\delta\rho(\mathbf{q},t) \\ &- |\mathbf{q}|^2\frac{v(\mathbf{q})}{\gamma_\mathrm{G}}\sum_{j\in\Omega}\mathrm{e}^{-\mathbf{q}\cdot\mathbf{r}_j(t)} + \mathbf{q}\cdot\mathbf{\Lambda}(\mathbf{q},t),\end{aligned} \tag{3.10}$$

so that the field dynamics can be directly solved as

$$\begin{aligned}\delta\rho(\mathbf{q},t) = &\int_{-\infty}^{t} \mathrm{d}s \mathrm{e}^{-D_{\mathrm{G}}|\mathbf{q}|^2 K(\mathbf{q})(t-s)} \\ &\times \left[\mathbf{q}\cdot\mathbf{\Lambda}(\mathbf{q},s) - |\mathbf{q}|^2 \frac{v(\mathbf{q})}{\gamma_{\mathrm{G}}} \sum_{j\in\Omega} \mathrm{e}^{-\mathbf{q}\cdot\mathbf{r}_j(s)}\right].\end{aligned} \tag{3.11}$$

Considering the limit of dilute driven tracers, where interactions among them are negligible, their dynamics reads

$$\gamma\dot{\mathbf{r}}_j = \mathbf{F}_{\mathrm{d}} + \boldsymbol{\xi}_j - \int_{\mathbf{q}} \mathbf{q} v(-\mathbf{q}) \mathrm{e}^{\mathbf{q}\cdot\mathbf{r}_j(t)} \delta\rho(\mathbf{q},t), \tag{3.12}$$

with $\int_{\mathbf{q}} = \int \mathrm{d}\mathbf{q}/(2\pi)^d$. As a result, Eq. 3.11 and Eq. 3.12 provide closed time-evolution equations for tracers only. It should only be valid for weak interactions *a priori*, yet previous works have shown that it remains qualitatively relevant even beyond this regime in practice [38, 114, 41]. Indeed, Gaussian field theories for density fluctuations provide a very good description of simple liquids [24].

To characterize the transport properties of the liquid in the presence of driving forces, our first goal is to obtain an explicit expression, in terms of microscopic details, for the tracer diffusion coefficient:

$$D = \lim_{t\to\infty} \frac{1}{2dt} \left\langle \left[\langle \mathbf{r}_i(t)\rangle - \mathbf{r}_i(t)\right]^2 \right\rangle. \tag{3.13}$$

We aim to explore connections between $D$ and dissipation, which is defined from stochastic thermodynamics as the power of the forces exerted by all tracers on solvent: $\mathcal{J} = \sum_i \langle \dot{\mathbf{r}}_i \cdot (\gamma\dot{\mathbf{r}}_i - \boldsymbol{\xi}_i)\rangle$, where $\cdot$ denotes a Stratonovich product [157, 155]. Dissipation is directly related to entropy production, as a measure of irreversibility, both when the drive is deterministic [157, 155] and when it is a correlated noise [111, 136, 158, 32]. Substituting Eq. 3.1, the dissipation coincides with the power of driving forces: $\mathcal{J} = \sum_{i\in\Omega} \langle \dot{\mathbf{r}}_i \cdot \mathbf{F}_{\mathrm{d},i}\rangle$. Besides, replacing $\dot{\mathbf{r}}_i$ by its expression (Eq. 3.1), and using the fact that $\boldsymbol{\xi}_i$ and $\mathbf{F}_{\mathrm{d},i}$ are uncorrelated,

we deduce that the dissipation can be further separated into free-motion and interaction contributions as $\mathcal{J} = Nf^2/\gamma - \dot{w}$, where the *rate of work* reads

$$\dot{w} = \frac{1}{\gamma} \sum_{i\in\Omega,j} \big\langle \mathbf{F}_{\mathrm{d},i} \cdot \nabla_i v(\mathbf{r}_i - \mathbf{r}_j) \big\rangle. \tag{3.14}$$

Given that $\dot{w}$ is the only non-trivial contribution to dissipation, connecting diffusion and dissipation simply amounts to expressing $D$ in terms of $\dot{w}$.

Deriving transport coefficients in nonequilibrium many-body systems, whose collective effects result from the complex interplay between interaction and driving forces, is a notoriously difficult task [9, 17, 168, 10, 87]. We set up a proper perturbation scheme by scaling the pair potential $v$ with a small dimensionless parameter $h \ll 1$ which controls the coupling between tracer and bath. In Appendix A2.1, we obtain some explicit expressions for $D$ and $\dot{w}$ to quadratic order in $h$ and in the scaled driving amplitude Pe.

First, we discuss the case of the deterministic drive in Eq. 3.2, and we focus on the limits of small and large driving frequency, respectively $\omega\tau_\mathrm{r} \ll 1$ and $\omega\tau_\mathrm{r} \gg 1$, where the relaxation time scale $\tau_\mathrm{r} = (D_\mathrm{G}/\sigma^2)K(|\mathbf{q}| = 1/\sigma)$ is set by density diffusion over the tracer size $\sigma$. First, at high frequencies $\omega\tau_\mathrm{r} \gg 1$, the rate of work per particle $\dot{w}/N$ and the deviation from equilibrium diffusion $D - D_\mathrm{eq}$, where $D_\mathrm{eq}$ is the diffusion coefficient for $\mathrm{Pe} = 0$, are given by

$$\begin{aligned} \frac{\dot{w}}{N} &= \Big(\frac{h\mathrm{Pe}}{\omega}\Big)^2 \cdot \frac{(T/\sigma)^2}{d\gamma^3} \int_\mathbf{q} |\mathbf{q}|^4 |v(\mathbf{q})|^2 \frac{1+\rho_0 K(\mathbf{q})}{K(\mathbf{q})}, \\ D - D_\mathrm{eq} &= \Big(\frac{h\mathrm{Pe}}{\omega}\Big)^2 \cdot \frac{T/\sigma^2}{d\gamma^3} \int_\mathbf{q} \frac{|\mathbf{q}|^2 |v(\mathbf{q})|^2}{K(\mathbf{q})\big[1+\rho_0 K(\mathbf{q})\big]}. \end{aligned} \tag{3.15}$$

In the opposite limit of low frequencies $\omega\tau_\mathrm{r} \ll 1$, we get

$$\begin{aligned} \frac{\dot{w}}{N} &= \frac{(h\mathrm{Pe})^2}{d\gamma\sigma^2} \int_\mathbf{q} \frac{|v(\mathbf{q})|^2}{K(\mathbf{q})\big[1+\rho_0 K(\mathbf{q})\big]}, \\ D - D_\mathrm{eq} &= \frac{5(h\mathrm{Pe})^2}{d\gamma T\sigma^2} \int_\mathbf{q} \frac{|v(\mathbf{q})|^2}{|\mathbf{q}|^2 K(\mathbf{q})\big[1+\rho_0 K(\mathbf{q})\big]^3}. \end{aligned} \tag{3.16}$$

Both $\dot{w}/N$ and $D - D_{\rm eq}$ are now independent of the driving frequency $\omega$. As a result, our perturbation theory shows that the scalings of $\dot{w}$ and $D - D_{\rm eq}$ are identical, both in terms of the drive amplitude Pe and of its frequency $\omega$, in asymptotic frequency regimes. Note that the scaled rate of work $\gamma\dot{w}/(Nf^2)$ coincides with the reduced equilibrium diffusion $\gamma D_{\rm eq}/T - 1$ to this order [40, 39], as expected from linear response.

The case of the active drive with correlations (Eq. 3.3) follows by using the mapping between disordered drive and active forcing in Sec. 3.3. In practice, we first derive the diffusion coefficient $D$ and the rate of work $\dot{w}$ for the driving force Eq. 3.4 at fixed disorder, as a straightforward generalization of the deterministic driving case, and we then average over the disorder. At small persistence $\tau \ll \tau_{\rm r}$, we get

$$\begin{aligned} \frac{\dot{w}}{N} &= \frac{\tau T (h\mathrm{Pe})^2}{d(\sigma\gamma)^2} \int_{\mathbf{q}} \frac{|\mathbf{q}|^2 |v(\mathbf{q})|^2}{K(\mathbf{q})}, \\ D - D_{\rm eq} &= \frac{3\tau (h\mathrm{Pe})^2}{d(\sigma\gamma)^2} \int_{\mathbf{q}} \frac{|v(\mathbf{q})|^2}{K(\mathbf{q})\big[1 + \rho_0 K(\mathbf{q})\big]^2}. \end{aligned} \tag{3.17}$$

In contrast, the large persistence limit $\tau \gg \tau_{\rm r}$ yields the same results as for the low frequencies regime of deterministic drive, namely the expressions Eq. 3.16. Indeed, the force $\mathbf{F}_{{\rm d},i}$ has a constant direction in such a limit, and the difference between deterministic and active drives, which respectively correspond to independent or similar directions for each tracer, is irrelevant in the limit of dilute tracers.

When the size of the bath particles, let's call it $\sigma_b$, is significantly smaller than the tracer size $\sigma \gg \sigma_b$, which amounts to setting different pair potential $v$ for bath-bath and for bath-tracer interactions, one can safely neglect the variation of $K(\mathbf{q})$ in Eqs. (3.15-3.17), so that $K(\mathbf{q}) \simeq K(|\mathbf{q}| = 1/\sigma_b)$. Then, in both regimes $\omega\tau_{\rm r} \gg 1$ ($\tau \ll \tau_{\rm r}$) and $\omega\tau_{\rm r} \ll 1$ ($\tau \gg \tau_{\rm r}$), the renormalization of the diffusion coefficient $D - D_{\rm eq}$ can be simply written in terms of the

rate of work per particle $\dot{w}/N$ for $\mathrm{Pe} \ll 1$ as

$$\frac{D - D_{\mathrm{eq}}}{\sigma^2} \sim \frac{\dot{w}}{NT}. \tag{3.18}$$

Thus, the excess rate at which tracers move over their own size compared to equilibrium, set by the lhs of Eq. 3.18, is controlled by the rate at which work is applied on tracers by nonequilibrium forces, set by the rhs of Eq. 3.18. The proportionality factor depends on the details of interactions and of density fluctuations. Interestingly, this result is valid both for deterministic and active drives. It corroborates numerical observations obtained in the model active liquid considered in the previous chapter.

## 3.5 Dissipation and Density Correlations

We now explore how dissipation relates to static density correlations of the liquid. To this end, we treat undriven bath particles without any approximation in what follows, instead of relying on the Gaussian density field theory for $\delta\rho$ as in Sec. 3.4, and we consider an arbitrary set of driving forces $\mathbf{F}_{\mathrm{d},i}$. In equilibrium, the liquid structure can be derived from a hierarchy of equations for density correlations, whose explicit form reflects the steady-state condition on the many-body distribution function [79]. In our settings, steady-state conditions should now provide modified equations for density correlations, which can potentially make apparent the connection with dissipation.

This motivates us to consider the average rate at which the potential $U = \sum_{i\in\Omega,j} v(\mathbf{r}_i - \mathbf{r}_j)$ changes, which can be written using Itô calculus as:

$$\gamma\langle\dot{U}\rangle = \sum_{i\in\Omega,j} \left\langle \left[\gamma(\dot{\mathbf{r}}_i - \dot{\mathbf{r}}_j) + 2T\nabla_i\right] \cdot \nabla_i v(\mathbf{r}_i - \mathbf{r}_j)\right\rangle. \tag{3.19}$$

Substituting the dynamics (Eq. 3.1) and using $\langle \boldsymbol{\xi}_i \cdot \nabla_i v \rangle = 0$ within Itô convention, we get:

$$\begin{aligned}\gamma\langle \dot{U} \rangle = &\sum_{i\in\Omega,j} (1+\delta_{j\in\Omega}) \big\langle \mathbf{F}_{\mathrm{d},i} \cdot \nabla_i v(\mathbf{r}_i - \mathbf{r}_j) \big\rangle \\ &+ \sum_{i\in\Omega,j,k} \big\langle \big[\nabla_i v(\mathbf{r}_i - \mathbf{r}_j)\big] \cdot \nabla_k \big[v(\mathbf{r}_i - \mathbf{r}_k) - v(\mathbf{r}_j - \mathbf{r}_k)\big] \big\rangle \\ &+ \sum_{i\in\Omega,j} 2T \big\langle \nabla_i^2 v(\mathbf{r}_i - \mathbf{r}_j) \big\rangle .\end{aligned} \tag{3.20}$$

In the first line of Eq. 3.20, we recognize the rate of work $\dot{w}$ as defined in Eq. 3.14, and the term $\gamma \dot{w}_{\mathrm{act}} = \sum_{\{i,j\}\in\Omega} \big\langle \mathbf{F}_{\mathrm{d},i} \cdot \nabla_i v(\mathbf{r}_i - \mathbf{r}_j) \big\rangle$ which quantifies the contribution of interactions among driven particles to dissipation. The latter vanishes exactly when the drive is identical for all particles, since $\sum_{\{i,j\}\in\Omega} \nabla_i v(\mathbf{r}_i - \mathbf{r}_j) = 0$ by symmetry, and it can be neglected for active drive when the fraction of driven particles is small. Then, using the steady-state condition $\langle \dot{U} \rangle = 0$, we deduce:

$$\begin{aligned}\dot{w} + \dot{w}_{\mathrm{act}} = &\frac{2\rho_0}{\gamma} \int g(\mathbf{r}) \big[ (\nabla v(\mathbf{r}))^2 - T\nabla^2 v(\mathbf{r}) \big] \mathrm{d}\mathbf{r} \\ &+ \frac{\rho_0^2}{\gamma} \iint \big[ g_{3a}(\mathbf{r},\mathbf{r}') + g_{3b}(\mathbf{r},\mathbf{r}') \big] \big[\nabla v(\mathbf{r})\big] \cdot \big[\nabla v(\mathbf{r}')\big] \mathrm{d}\mathbf{r}\mathrm{d}\mathbf{r}',\end{aligned} \tag{3.21}$$

where

$$\begin{aligned} g(\mathbf{r}) &= \frac{1}{N} \sum_{i\in\Omega,j}{}' \big\langle \delta(\mathbf{r} - \mathbf{r}_i + \mathbf{r}_j) \big\rangle, \\ g_{3a}(\mathbf{r},\mathbf{r}') &= \frac{1}{N^2} \sum_{i\in\Omega,j,k}{}' \big\langle \delta(\mathbf{r} - \mathbf{r}_i + \mathbf{r}_j) \delta(\mathbf{r}' - \mathbf{r}_i + \mathbf{r}_k) \big\rangle, \\ g_{3b}(\mathbf{r},\mathbf{r}') &= \frac{1}{N^2} \sum_{i\in\Omega,j,k}{}' \big\langle \delta(\mathbf{r} - \mathbf{r}_i + \mathbf{r}_j) \delta(\mathbf{r}' - \mathbf{r}_j + \mathbf{r}_k) \big\rangle, \end{aligned} \tag{3.22}$$

and $\sum'$ denotes a sum without the overlap of indices: $i \neq j$, $k \neq i$ and $k \neq j$. The power balance (Eq. 3.21), valid for an arbitrary driving, either deterministic or active, is our first main result. Importantly, it holds for generic interactions and for any number of driven particles, namely not only in the limit of dilute tracers, in contrast with the results

in Sec. 3.4.

In practice, it reflects how density correlations adapt to the presence of nonequilibrium forces. For a vanishing rate of work ($\dot{w} = 0 = \dot{w}_{\mathrm{act}}$), one recovers the first order of the equilibrium Yvon-Born-Green (YBG) hierarchy, in its integral form, for two-component fluids [79]. At finite rate of work ($\dot{w} \neq 0$), the relation between the two-body correlation $g$ and the three-body terms $\{g_{3a}, g_{3b}\}$ is now implicitly constrained by dissipation. A direct implication is that the rate of work can now be inferred simply by measuring static density correlations, provided that the pair-wise interaction potential is known, for a given driven liquid. Importantly, such an approach does not require any invasive methods based on comparing fluctuations and response [81, 119, 57, 186, 1], and it does not rely on a detailed analysis of particle trajectories [148, 116] or currents in phase space [67, 103], whose experimental implementation can require elaborate techniques [7, 70].

However, the power balance (Eq. 3.21) is not straightforward to test, either numerically or experimentally, due to the three-body correlations. In equilibrium, where tracer and bath particles are indistinguishable, we get $g_{3a} = g_{3b}$. Assuming that this remains approximately valid in the driven case for a small fraction of tracers, the rate of work can simply be written in terms of the force exerted on a tracer $\mathbf{F}_i = -\sum_j \nabla_i v(\mathbf{r}_i - \mathbf{r}_j)$ as

$$\dot{w} \simeq \frac{2}{\gamma} \sum_{i \in \Omega} \left[ \left\langle \mathbf{F}_i^2 \right\rangle + T \left\langle \nabla_i \cdot \mathbf{F}_i \right\rangle \right]. \tag{3.23}$$

To probe the validity of this result, we simulate the dynamics in Eq. 3.1 where 10% of particles are subject to the driving force, considering either the deterministic periodic drive (Eq. 3.2) or the active noise drive (Eq. 3.3). Our measurements in Fig. 3.2 show that Eq. 3.23 is indeed a good approximation at small Pe and small $\tau$, namely when the drive only weakly perturbs the liquid. The discrepancy is higher for the active case compared with the deterministic one, since $\dot{w}_{\mathrm{act}} = 0$ in the latter without any approximation. In contrast with previous approaches [81, 101, 7], which rely on prospecting the whole system, our results demonstrate

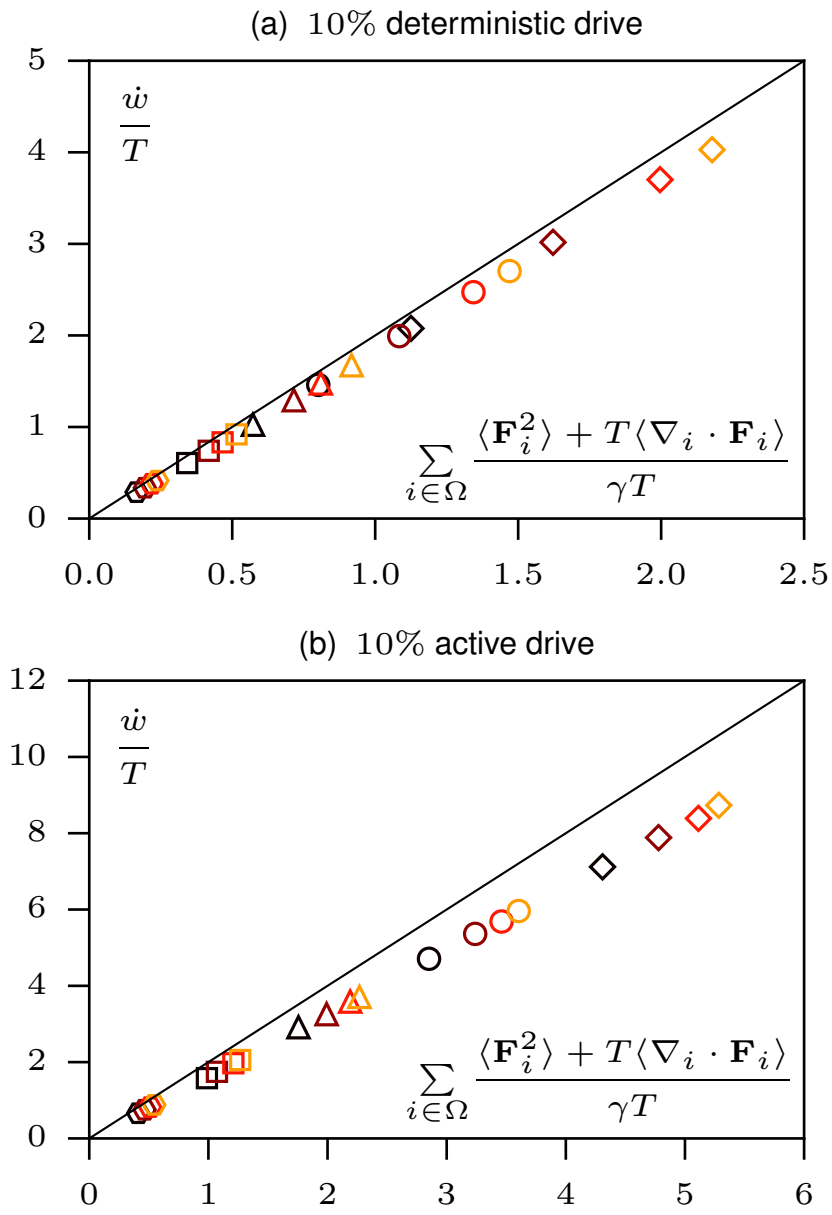



Figure 3.2: Parametric plot of the rate of work $\dot{w}/T$ and of the statistics of bath-tracer forces $\sum_{i\in\Omega}\left[\langle \mathbf{F}_i^2\rangle + T\langle \nabla_i \cdot \mathbf{F}_i\rangle\right]/(\gamma T)$ when 10% of particles are driven by either (a) a deterministic force, or (b) an active force. The solid line with slope 2 refers to the approximate relation Eq. 3.23. The satisfying agreement with numerical data indicates that the rate of work can be estimated by only measuring bath-tracer forces. The simulations were performed using the procedure described in the Methods section. Parameters: Pe = 12 (hexagons), 18 (squares), 24 (triangles), 30 (circles), 36 (diamonds); (a) $\tau T/(\gamma\sigma^2) = 2\times10^{-1}$ (black), $3\times10^{-1}$ (brown), $4\times10^{-1}$ (red), $5\times10^{-1}$ (orange); (b) $\tau T/(\gamma\sigma^2) = 2\times10^{-2}$ (black), $3\times10^{-2}$ (brown), $4\times10^{-2}$ (red), $5\times10^{-2}$ (orange).

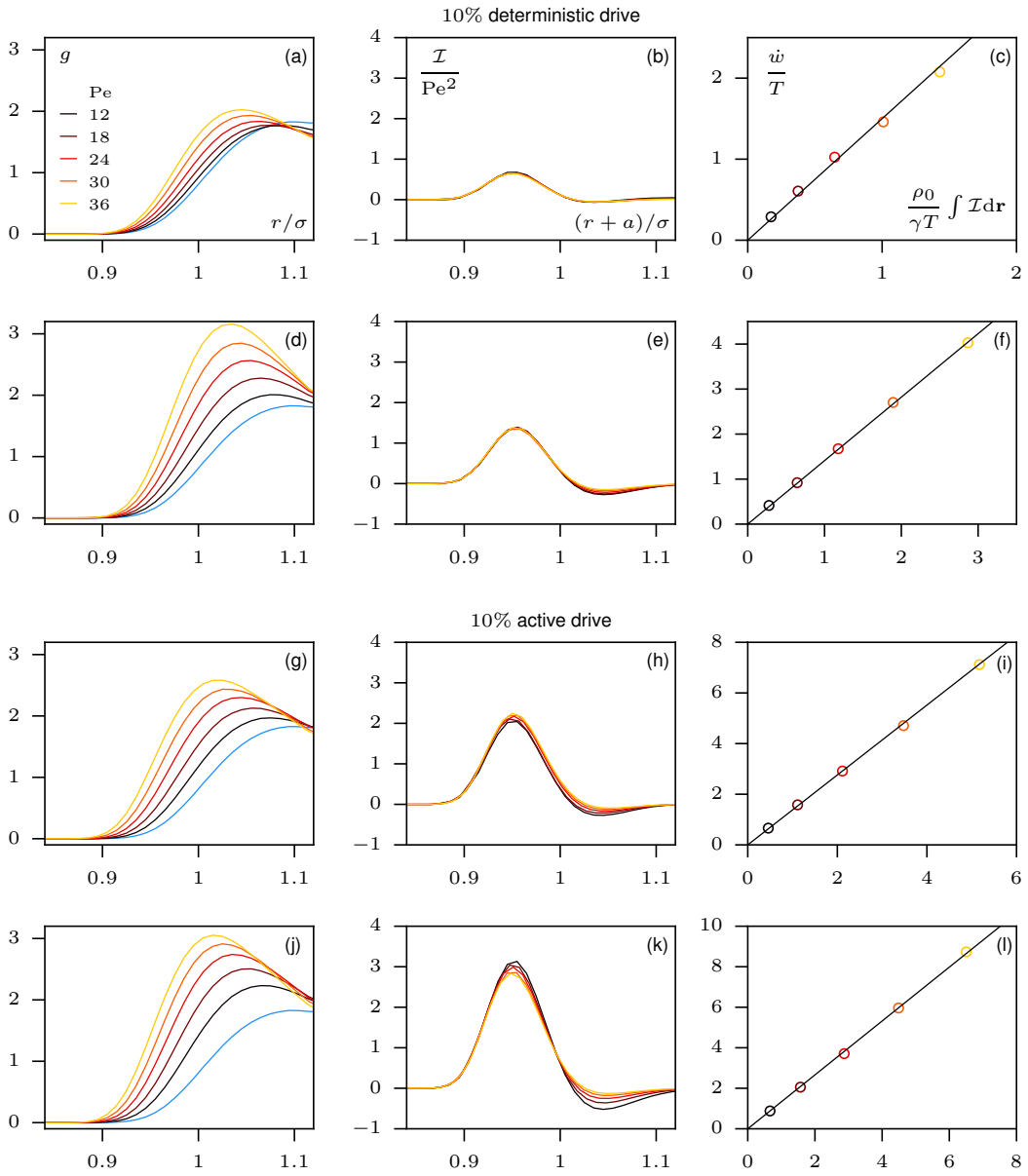



Figure 3.3: Connecting dissipation and structure for a liquid where 10% of particles are driven by either deterministic or active forces. (Left) Bath-tracer density correlation $g$ as a function of interparticle distance $r/\sigma$. The blue solid line corresponds to the equilibrium correlation function $g_{\mathrm{eq}}$ for $\mathrm{Pe} = 0$. (Middle) Deviation from equilibrium correlations $\mathcal{I} = \left[(\nabla v)^2 - T\nabla^2 v\right](g - g_{\mathrm{eq}})$ scaled by $\mathrm{Pe}^2$ as a function of $(r+a)/\sigma$ where $a(\mathrm{Pe})$ is a fitting parameter. The data almost collapse into a master curve for each row, namely at a given $\tau$. (Right) Parametric plot of the rate of work $\dot{w}/T$ and the integrated deviation from equilibrium correlations $\rho_0 \int \mathcal{I}(\mathbf{r})\mathrm{d}\mathbf{r}/(\gamma T)$ showing a linear relation. The black solid line with slope $\alpha$ is the best linear fit, and the marker colors refer to the same Péclet values as in the left and right columns. Parameters: $\{\tau T/(\gamma\sigma^2), \alpha\} = \{2\times10^{-1}, 1.50\}$ (a-c), $\{5\times10^{-1}, 1.41\}$ (d-f), $\{2\times10^{-2}, 1.38\}$ (g-i), $\{5\times10^{-2}, 1.33\}$ (j-l).

that the rate of work can actually be evaluated with only a small error by considering solely forces acting on tracer: the contribution of forces on other particles is negligible for a small fraction of driven tracers.

To further evaluate the change in liquid structure induced by dissipation, we measure the deviation from equilibrium pair correlations $g - g_{\mathrm{eq}}$ due to the driving forces (left column in Fig. 3.3). In particular, inspired by the two-body contribution in the power balance (Eq. 3.21), we focus on the observable $\mathcal{I} = \big[(\nabla v)^2 - T\nabla^2 v\big](g - g_{\mathrm{eq}})$. At a given $\tau$, scaling $\mathcal{I}$ by $\mathrm{Pe}^2$ reveals that all curves almost collapse into a master curve for our numerical range $\mathrm{Pe} \in [12, 36]$, as reported in the middle column of Fig. 3.3. In practice, particles overlap more for a stronger drive, so that $g$ departs from zero at smaller interparticle distance. To correct for this, we introduce a shift of the curves $\mathcal{I}(|\mathbf{r}|)$ as $|\mathbf{r}| \to |\mathbf{r}| + a$, where $a(\mathrm{Pe})$ is a fitting parameter. Given that the rate of work also scales like $\mathrm{Pe}^2$, it suggests the existence of an underlying relation between $\int \mathcal{I}(\mathbf{r})\mathrm{d}\mathbf{r}$ and $\dot{w}$. In practice, a linear fitting provides a satisfactory agreement between them, as shown in the right column of Fig. 3.3:

$$\dot{w} = \frac{\alpha\rho_0}{\gamma} \int \big[g(\mathbf{r}) - g_{\mathrm{eq}}(\mathbf{r})\big]\big[(\nabla v(\mathbf{r}))^2 - T\nabla^2 v(\mathbf{r})\big]\mathrm{d}\mathbf{r}, \tag{3.24}$$

where $\alpha$ is a fitting parameter independent of the Péclet number. This empirical relation demonstrates that, in the limit of dilute tracers, the rate of work can actually be directly estimated by comparing driven and equilibrium pair correlations for both deterministic and active drives. Comparing Eq. 3.21 and Eq. 3.24, we deduce the following integral relation between density correlations

$$\begin{aligned} &\int \big[(2-\alpha)g(\mathbf{r}) + \alpha\, g_{\mathrm{eq}}(\mathbf{r})\big]\big[T\nabla^2 v(\mathbf{r}) - (\nabla v(\mathbf{r}))^2\big]\mathrm{d}\mathbf{r} \\ &= \rho_0 \iint \big[g_{3a}(\mathbf{r}, \mathbf{r}') + g_{3b}(\mathbf{r}, \mathbf{r}')\big]\big[\nabla v(\mathbf{r})\big] \cdot \big[\nabla v(\mathbf{r}')\big]\mathrm{d}\mathbf{r}\mathrm{d}\mathbf{r}'. \end{aligned} \tag{3.25}$$

Interestingly, it is reminiscent again of the connection between density correlations provided

by the YBG hierarchy at equilibrium [79]. Similarly, the relation in Eq. 3.25 amounts to a constraint on density correlations, now valid for nonequilibrium liquids, which could guide the search for explicit predictions on the emerging structure. Importantly, it does not rely on any equilibrium mapping, in contrast with previous works [113, 146, 194], since it remains valid for non-negligible dissipation.

The power balance in Eq. 3.21 can actually be extended to the case where all particles in the liquid are driven as

$$\begin{aligned} \dot{w} = \frac{\rho_0}{\gamma} \int g(\mathbf{r}) \big[ (\nabla v(\mathbf{r}))^2 - T \nabla^2 v(\mathbf{r}) \big] \mathrm{d}\mathbf{r}, \\ + \frac{\rho_0^2}{\gamma} \iint g_3(\mathbf{r}, \mathbf{r}') \big[ \nabla v(\mathbf{r}) \big] \cdot \big[ \nabla v(\mathbf{r}') \big] \mathrm{d}\mathbf{r} \mathrm{d}\mathbf{r}', \end{aligned} \tag{3.26}$$

where $g$ and $g_3$ now refer respectively to the two-body and three-body density correlations among all particles. This leads to an exact relation between the rate of work and the forces $\mathbf{F}_i$ applied to particles as:

$$\dot{w} = \frac{1}{\gamma} \sum_i \big[ \langle \mathbf{F}_i^2 \rangle + T \langle \nabla_i \cdot \mathbf{F}_i \rangle \big], \tag{3.27}$$

which differs from the relation in Eq. 3.23 for driven tracers by an overall factor of 2. A result analogous to Eq. 3.27 was found previously for deterministic drive [107, 6]. The main difference is that Eq. 3.27 only features interaction forces $\mathbf{F}_i$ in the rhs, thus allowing one to evaluate the rate of work without any prior knowledge on the driving force. Besides, it is valid for both deterministic and active drives. Moreover, conducting the same analysis of density correlations as for driven tracers, $\mathcal{I}$ exhibits again a scaling with $\mathrm{Pe}^2$, as reported in Fig. 3.4. We show that $\dot{w}$ and $\int \mathcal{I}(\mathbf{r}) \mathrm{d}\mathbf{r}$ are also linearly related. Introducing the linear coefficient as $\alpha \rho_0 / (2\gamma)$ is consistent with substituting Eq. 3.25 into Eq. 3.26, where $g_{3,a} + g_{3,b}$ is now replaced by $2g_3$. Hence, it demonstrates that the rate of work is also accessible from the nonequilibrium deviation of pair correlations in fully driven liquids.

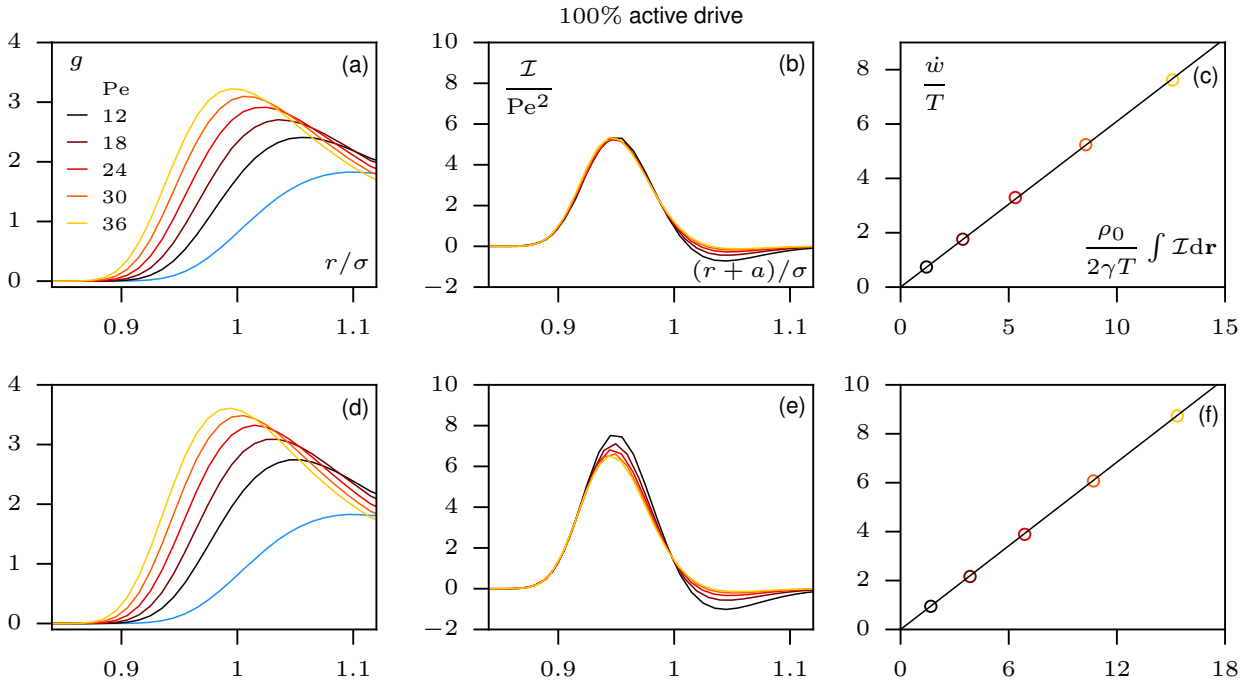



Figure 3.4: Connecting dissipation and structure for a liquid where 100% of particles are driven by an active force. (Left) Density correlation $g$ as a function of interparticle distance $r/\sigma$. The blue solid line corresponds to the equilibrium correlation function $g_{\rm eq}$ for $\mathrm{Pe} = 0$. (Middle) Deviation from equilibrium correlations $\mathcal{I} = \left[(\nabla v)^2 - T\nabla^2 v\right](g - g_{\rm eq})$ scaled by $\mathrm{Pe}^2$ as a function of $(r + a)/\sigma$ where $a(\mathrm{Pe})$ is a fitting parameter. The data almost collapse into a master curve for each row, namely at a given $\tau$. (Right) Parametric plot of the rate of work $\dot{w}/T$ and the integrated deviation from equilibrium correlations $\rho_0 \int \mathcal{I}(\mathbf{r})\mathrm{d}\mathbf{r}/(2\gamma T)$ showing a linear relation. The black solid line with slope $\alpha$ is the best linear fit, and the marker colors refer to the same Péclet values as in the left and right columns. Simulation details can be found in the Methods section. Parameters: $\{\tau T/(\gamma\sigma^2), \alpha\} = \{2 \times 10^{-2}, 1.22\}$ (a-c), $\{5 \times 10^{-2}, 1.14\}$ (d-f).

Overall, the results of this Section illustrate how dissipation affects the transport and structural properties of driven liquids, measured in terms of diffusion coefficient and density fluctuations. These findings motivate the following question: can nonequilibrium forces be tuned to reliably stabilize target configurations? To explore this, we rely in what follows on the framework of large deviation theory. In practice, our strategy amounts to biasing trajectories in terms of dissipation, related to many-body interactions by Eq. 3.21, to mimic the effect of an external drive. Following this route, our analytical and numerical results provide some concrete intuition for how interactions in a multicomponent system can be controllably

renormalized by nonequilibrium forces. Hence, we demonstrate the ability to nucleate structures different from those characteristic of the equilibrium Boltzmann distribution to help guide self-assembly [12] and collective motion [189] far from equilibrium. These results further illustrate the interplay between energy dissipation and organization in nonequilibrium many-body settings.

## 3.6 Nonequilibrium Biased Trajectories and Renormalized Interactions

This Section describes how target structures and dynamics can be promoted by means of a dynamical bias, an extensive project that was done primarily by my collaborator Takahiro Nemoto and my adviser Suri Vaikuntanathan. We begin by considering a system of interacting Brownian particles without any driving force

$$\gamma \dot{\mathbf{r}}_i = -\nabla_i \sum_j v(\mathbf{r}_i - \mathbf{r}_j) + \boldsymbol{\xi}_i, \tag{3.28}$$

where the statistics of the noise term $\boldsymbol{\xi}_i$ is the same as the one in Eq. 3.1. The rate of work $\dot{w}$ defined in Eq. 3.14 is zero because of the absence of driving. In Sec. 3.2, to obtain a non-zero rate of energy flow through the system, we consider an explicit driving force $\mathbf{F}_{\mathrm{d},i}$ and we explore its effects on the transport and structural properties of the liquid. In practice, different types of driving can lead to the same dissipation. In this section, using the framework of large deviation theory, we take an alternative approach where the dynamics is now conditioned by enforcing a required energy flow without any explicit driving. Thus, exploring how the system adapts to this requirement provides a new insight into the relation between dissipation and organization in driven systems, which is distinct from yet complements the approach in Sec. 3.2.

To this end, we focus on the subset of noise realizations that are conditioned on a non-

zero rate of work. In particular, these realizations no longer have zero average, so that one can re-define the noise term in Eq. 3.28 as $\boldsymbol{\xi}_i \to \boldsymbol{\xi}_i + \mathbf{F}_{\mathrm{aux},i}$ by introducing an *auxiliary force* $\mathbf{F}_{\mathrm{aux},i}$ [90, 27]. Hence, the stochastic dynamics given by Eq. 3.28 with added force $\mathbf{F}_{\mathrm{aux},i}$ provides an explicit case which ensures a non-zero energy flow rate. In practice, this dynamics can be drastically different from the original one, thus opening the door to stabilizing unexpected structure and to promoting novel collective effects. Interestingly, such a dynamics can actually be regarded as the optimal strategy to effectively enforce a target condition on rate of work [49].

Formally, to study the dynamics conditioned by dissipation, we bias the probability of trajectories. This is done by introducing an exponential weighting factor $\exp\left[\kappa \int_0^t \mathcal{E}(s)\mathrm{d}s\right]$ where $\mathcal{E}$ is the observable which conditions the dynamics, *e.g.* energy flow rate, and $\kappa$ is a conjugate field. In practice, the relative importance of biasing in the dynamics is controlled by $\kappa$, which in turn controls the average value of $\mathcal{E}$ [182]. Before deriving the central results of this section, namely relations between biased energy flow rates and organization, we first introduce a simple example in which the connection between auxiliary forces and exponentially biased ensemble can be clearly seen.

### *3.6.1 Dynamical Bias and External Forces*

To introduce pedagogically our methods, we first show how biasing trajectories can lead to effectively introducing a driving force. Inspired by the role of dissipation in emerging liquid properties, as discussed in Sec. 3.2, we bias the equilibrium dynamics (Eq. 3.28) with the sum of the dissipation and the rate of work, scaled by $T$, that would be produced by applying a constant force $\mathbf{F}_{\mathrm{d}}$ to a subset $\Omega$ of particles:

$$\mathcal{E} = \frac{1}{\gamma T} \sum_{i \in \Omega} \mathbf{F}_{\mathrm{d}} \cdot \left[\gamma \dot{\mathbf{r}}_i + \nabla_i V\right], \tag{3.29}$$

where $V = (1/2)\sum_{i,j} v(\mathbf{r}_i - \mathbf{r}_j)$. The path probability $\mathcal{P} \sim \exp\big[-\sum_i \int_0^t \mathbb{A}_i(s)\mathrm{d}s\big]$ corresponding to this biased ensemble is obtained with standard methods [115, 34]:

$$\mathbb{A}_i = \frac{1}{4\gamma T}\big[\gamma\dot{\mathbf{r}}_i + \nabla_i V\big]^2 - \frac{1}{2\gamma}\nabla_i^2 V - \frac{\kappa}{\gamma T}\delta_{i\in\Omega}\mathbf{F}_\mathrm{d}\cdot\big[\gamma\dot{\mathbf{r}}_i + \nabla_i V\big], \tag{3.30}$$

where the two first terms correspond to the unbiased dynamics (Eq. 3.28), and the third one to the bias in Eq. 3.29. It can also be written as

$$\mathbb{A}_i = \frac{1}{4\gamma T}\big[\gamma\dot{\mathbf{r}}_i - 2\kappa\delta_{i\in\Omega}\mathbf{F}_\mathrm{d} + \nabla_i V\big]^2 - \frac{1}{2\gamma}\nabla_i^2 V - \frac{\kappa^2}{\gamma T}\delta_{i\in\Omega}\mathbf{F}_\mathrm{d}^2. \tag{3.31}$$

As a result, given that the last term in Eq. 3.31 can be absorbed in a normalization factor, we deduce that the trajectories biased by Eq. 3.29 can be generated, at leading order, in a physical dynamics where the external force $2\kappa\mathbf{F}_\mathrm{d}$ is applied to every particle in $\Omega$. In particular, it does not feature any long-range interactions which are usually found in auxiliary dynamics [88].

### *3.6.2 Dynamical Bias and Modified Interactions*

To go beyond the case of applying a constant force, we now seek a dynamical bias which regulates particle interactions in a controlled manner. In particular, we examine cases where the control parameters $\kappa_{ij}$ are specific to particle pairs $\{i,j\}$, so that the biasing factor in path probability now reads $\exp\big[\sum_{i,j}\kappa_{ij}\int_0^t \mathcal{E}_{ij}(s)\mathrm{d}s\big]$. Now, our choice for the biasing function $\mathcal{E}_{ij}$ is informed by the connection between the rate of work and many-body interactions in driven liquids, as detailed in Sec. 3.5. Specifically, we observe that the power balance (Eq. 3.21) for deterministic drive ($\dot{w}_\mathrm{act} = 0$) can be written as $\dot{w} = -\sum_{i\in\Omega,j}\langle\mathcal{L}v(\mathbf{r}_i - \mathbf{r}_j)\rangle$ in terms of the evolution operator of the equilibrium dynamics (Eq. 3.28) defined by $\gamma\mathcal{L} =$

$\sum_i \left[T\nabla_i - \nabla_i V\right] \cdot \nabla_i$. This motivates us to consider the following bias

$$\mathcal{E}_{ij} = \frac{1}{4T}\mathcal{L}v(\mathbf{r}_i - \mathbf{r}_j). \tag{3.32}$$

In the unbiased ensembles of Sec. 3.5, $\langle \mathcal{E}_{ij} \rangle$ provides a measure of the rate at which driving forces pump energy into or extract energy from the specific interaction between the $i^{\text{th}}$ and $j^{\text{th}}$ particles. Here, instead of driving the system with a specific driving force, trajectories are driven by atypical realizations of the noise generated by biased sampling.

To explore how this bias modifies interactions, we first employ a derivation different from the path integral approach in Sec. 3.6.1. Based on the procedure in [90, 27], the auxiliary physical dynamics, which has the same statistical properties as in the biased ensemble, can be constructed by solving the eigenvalue equation

$$\left[\mathcal{L} + \sum_{i,j} \kappa_{ij}\mathcal{E}_{ij}\right] \mathcal{G}(\{\mathbf{r}_k\}, \kappa) = \lambda(\kappa)\, \mathcal{G}(\{\mathbf{r}_k\}, \kappa), \tag{3.33}$$

where the eigenvalue $\lambda$, parametrized by $\kappa_{ij}$, is the scaled cumulant generating function appropriate to $\mathcal{E}_{ij}$. The auxiliary dynamics is then defined by replacing the interaction potential in Eq. 3.28 by the following auxiliary potential:

$$\tilde{V} = \frac{1}{2}\sum_{i,j} v(\mathbf{r}_i - \mathbf{r}_j) - 2T \ln \mathcal{G}. \tag{3.34}$$

In practice, computing $\mathcal{G}$ is a highly non-trivial procedure for many-body systems. The explicit solutions considered so far concern either exclusion processes [141, 140, 143] or particle-based diffusive systems restricted to small noise regimes [178, 142] and non-interacting cases in some specific potentials [110, 185, 133].

In our case, a simple expression can be obtained for the auxiliary potential $\tilde{V}_i$ by solving

Eq. 3.33 perturbatively at small bias parameter $\kappa$. Specifically, we expand

$$\begin{aligned}\lambda(\kappa) &= \sum_{ij} \kappa_{ij} \langle \mathcal{E}_{ij} \rangle + \mathcal{O}(\kappa^2), \\ \mathcal{G}(\{\mathbf{r}_k\}, \kappa) &= \mathcal{G}^{(0)} + \sum_{ij} \kappa_{ij}\, \mathcal{G}^{(1)}_{ij}(\{\mathbf{r}_k\}) + \mathcal{O}(\kappa^2),\end{aligned} \tag{3.35}$$

where $\mathcal{G}_0$ is the uniform eigenvector associated with the zero eigenvalue. Given that $\langle \mathcal{E}_{ij} \rangle = 0$ in steady state, which follows from the vanishing current condition in the unbiased dynamics ($\langle \dot{v} \rangle = 0$), the leading non-trivial order of Eq. 3.33 reads

$$\sum_{ij} \kappa_{ij} \Big[ \mathcal{L}\, \mathcal{G}^{(1)}_{ij} + \mathcal{G}^{(0)} \mathcal{E}_{ij} \Big] + \mathcal{O}(\kappa^2) = 0. \tag{3.36}$$

Substituting the explicit expressions for the biasing function in Eq. 3.32, we then deduce that $4T\mathcal{G}^{(1)}_{ij} = -\mathcal{G}_0\, v(\mathbf{r}_i - \mathbf{r}_j)$ is a solution of the eigenvalue problem to order $\kappa$. The auxiliary potential follows as

$$\tilde{V} = \frac{1}{2} \sum_{i,j} (1 + \kappa_{ij})\, v(\mathbf{r}_i - \mathbf{r}_j) + \mathcal{O}(\kappa^2). \tag{3.37}$$

Therefore, biasing with Eq. 3.32 amounts to changing the strength of particle interaction by a factor $\kappa_{ij}$ specifically for any pair $\{i, j\}$. This is the main result of this section.

While energy flows were sustained by explicit nonequilibrium forces in previous Sections, we now maintain a non-zero average for $\mathcal{E}_{ij}$ by a biased sampling of trajectories. The corresponding noise realizations can be thought of as an external protocol, which leads to modifying the energy landscape sampled by the biased system as given in Eq. 3.37. Note that tuning interaction strength between targeted pairs is qualitatively consistent with the effect of external driving. Indeed, phase separation in mixtures of driven and undriven particles, reported both experimentally and numerically, can be rationalized in terms of an effective decrease of specific interactions between these particles, as also shown in Chapter 2 of this dissertation.

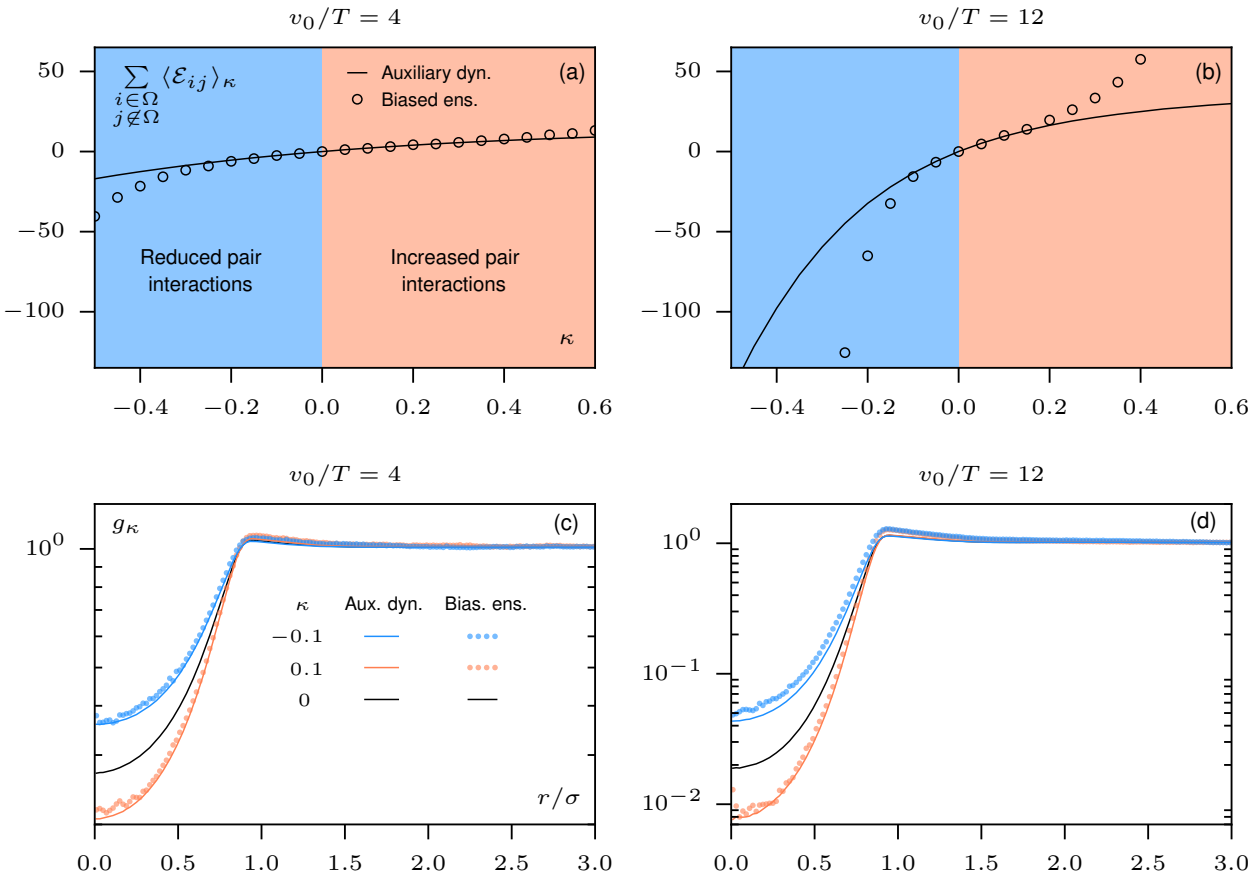



Figure 3.5: (a-b) Average biasing observable $\sum_{i\in\Omega,j\notin\Omega}\langle\mathcal{E}_{ij}\rangle_\kappa = \sum_{i\in\Omega,j\notin\Omega}\langle\mathcal{L}v(\mathbf{r}_i-\mathbf{r}_j)\rangle_\kappa/T$ as a function of bias parameter $\kappa$, where $\mathcal{L}$ and $v$ respectively denote the evolution operator and the pair potential of the equilibrium dynamics (Eq. 3.28). Results from the first-order auxiliary dynamics (solid lines) and from a direct sampling of the biased ensemble (circles) coincide for a finite range of $\kappa$. (c-d) Biased density correlation $g_\kappa$ as a function of interparticle distance $r/\sigma$ obtained from auxiliary dynamics (solid lines) and direct sampling (dotted lines). At leading order, our dynamical bias effectively renormalizes the potential $v$ by a factor $\kappa$ for specific pairs of particles $\{i \in \Omega, j \notin \Omega\}$, in satisfying agreement with direct sampling. This illustrates the control of liquid structure at small $\kappa$ and weak interactions.

Moreover, the techniques in this chapter allow one to anticipate the trajectories generated at higher-order when now biasing with Eq. 3.32. To this end, we consider the ensemble where the first-order dynamics, given by the potential in Eq. 3.37, is biased with $\exp\left[\int_0^t \varepsilon(s)\mathrm{d}s\right]$ defined in terms of

$$\varepsilon = \frac{1}{4\gamma T}\sum_k \Big[\sum_{i,j}\kappa_{ij}\nabla_k v(\mathbf{r}_i(s)-\mathbf{r}_j(s))\Big]^2. \tag{3.38}$$

As detailed in Appendix A2.2, this ensemble is equivalent to biasing the original dynamics (Eq. 3.28) with Eq. 3.32. Thus, the effect of higher-order bias on trajectories amounts to maximizing the squared forces in the integrand of Eq. 3.38, which effectively tends to cluster particles for both signs of $\kappa_{ij}$.

Finally, the decomposition between first-order auxiliary potential and higher-order symmetric bias can be extended to a generic class of biases of the form $T\mathcal{E}_{ij} = \mathcal{L}A(\mathbf{r}_i - \mathbf{r}_j)$ for an arbitrary observable $A$: the corresponding first-order auxiliary potential $V + 2\sum_{i,j}\kappa_{ij}A(\mathbf{r}_i - \mathbf{r}_j)$ is now complemented with the higher-order bias (Eq. 3.38) where $A$ replaces $v$. Such a bias is reminiscent of, yet qualitatively different from, the escape rate used to promote dynamical heterogeneity in glassy systems [138, 62]. In our case, clustering is favored for both positive and negative bias parameters $\kappa_{ij}$. In particular, this is in contrast with the emergence of a hyperuniform phase, where large scale fluctuations are suppressed, reported when biasing some hydrodynamic theories of diffusive systems [91].

To illustrate the potential of our bias to control liquid properties, we focus in what follows on the specific case $\kappa_{ij} = \kappa\delta_{i\in\Omega}\delta_{j\notin\Omega}$ where all pairs between a subset $\Omega$ and other particles are biased with the same strength $\kappa$. Here, the set $\Omega$ could for instance refer to some tracer particles immersed in the liquid, to connect with the settings in Sec. 3.2.

To confirm numerically the validity of our approach, we first probe the range of the first-order auxiliary dynamics where interactions are predicted to be simply renormalized. We compare measurements of $\sum_{i\in\Omega,j\notin\Omega}\langle\mathcal{E}_{ij}\rangle_\kappa$, where $\langle\cdot\rangle_\kappa$ denotes an average in the biased ensemble, obtained from simulations with the renormalized potentials in Eq. 3.37 and from a direct sampling of the biased ensemble. The latter is implemented with a cloning algorithm which regularly selects and multiplies rare realizations for efficient sampling [66, 174, 86, 129, 143, 94, 15]. For convenience, interactions are now given by the soft-core potential $v(\mathbf{r}) = v_0 \exp\big[-1/(1-(|\mathbf{r}|/\sigma)^2)\big]\Theta(\sigma - |\mathbf{r}|)$. For weak interactions ($v_0 = 4T$), we observe a very satisfying agreement between the two measurements for a finite range of $\kappa$, as reported

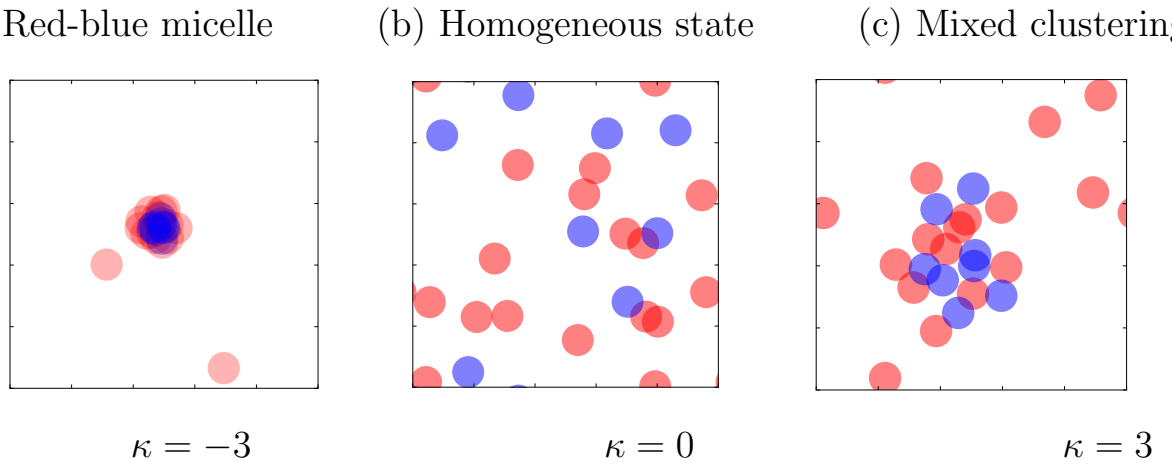



Figure 3.6: Configurations obtained from a direct sampling of the biased ensemble where the pair interactions between red and blue particles are selectively modified. In the unbiased dynamics ($\kappa = 0$), interactions are purely repulsive with a soft core which has a similar strength for all particles, either red or blue, so that the system is homogeneous. The dynamical bias promotes clustering for both signs of $\kappa$, yet it changes interaction selectively for either sign. The repulsion is increased between red and blue particles for $\kappa = 3$, and their interactions become effectively attractive for $\kappa = -3$. As a result, the clusters which emerge spontaneously have different structures: either a random composition of mixed reds and blues ($\kappa = 3$) or a micelle-like structure with a blue core ($\kappa = -3$). This illustrates how biasing specific pairs leads to supervised spatial organization.

in Fig. 3.5(a), which supports the validity of our perturbation up to interaction change between $-20\%$ and $+40\%$. The range of validity decreases as $v_0/T$ increases, as shown in Fig. 3.5(b), and we expect a similar trend when also increasing the number of biased pairs.

To explore further the features of this biased ensemble, we now compare the density correlations of biased pairs $g_\kappa(\mathbf{r}) \sim \sum_{i\in\Omega, j\notin\Omega}\langle\delta(\mathbf{r} - \mathbf{r}_i + \mathbf{r}_j)\rangle_\kappa$ obtained from both direct sampling and first-order auxiliary dynamics. For $\kappa = \pm 0.1$, we observe that the structural modification induced by the bias becomes more dramatic as $v_0/T$ increases. The agreement between the cloning and auxiliary dynamics is good for the whole curve when $v_0/T = 4$, whereas a deviation appears beyond $r \simeq \sigma$ when $v_0/T = 12$, as shown in Figs. 3.5(c-d). In both cases, the region of particle overlap $r < \sigma$ is well reproduced. These results corroborate the ability of the first-order auxiliary dynamics to capture interaction changes as a simple

renormalization of potential strength. In contrast, the tendency for particles to cluster, manifested numerically in the increased peak value at $r \simeq \sigma$, is a higher-order effect missed by this auxiliary dynamics when $v_0/T = 12$. Yet, note that the peak value is comparable for $\kappa = \pm 0.1$, in agreement with Eq. 3.38 being symmetric in $\kappa$. Altogether, these results demonstrate that our bias modulates the liquid structure in a controlled manner for small bias and weak interactions as predicted by Eq. 3.37.

Finally, we probe numerically the effect of large bias ($|\kappa| > 1$) using direct sampling, to explore configurations significantly distinct from the one of the equilibrium dynamics (Eq. 3.28). The particles spontaneously tend to cluster for both positive and negative $\kappa$, as shown in Fig. 3.6. This confirms the propensity of trajectories to maximize interaction forces at high bias, as captured by Eq. 3.38. Importantly, the shape of clusters differs depending on the sign of $\kappa$: a micelle-like structure featuring the particles in $\Omega$ at the core (blue) surrounded by others (red) appears for $\kappa = -3$, whereas clusters have a random composition for $\kappa = 3$. Again, this agrees with the renormalized interactions being either increased ($\kappa > 0$) or decreased ($\kappa < 0$). In practice, the interaction strength changes sign when $\kappa < -1$ according to Eq. 3.37, so that the red-blue pairs are effectively attractive for $\kappa = -3$. To optimize the overall energy, the most favorable configuration then consists in maximizing (minimizing) overlap of particles in $\Omega$ (not in $\Omega$), which in turn stabilizes a cluster of blues surrounded by reds. In general, two types of configurations should generically be stabilized for a given interaction potential $v$, depending on the sign of the bias. Overall, this establishes a reliable proof of principle for the design of tailored self-assembled structures with our specific choice of biased ensembles.

As a final illustration of how collective effects can be controlled by dynamical bias, we consider a model of self-propelled particles where interactions are now only mediated via the

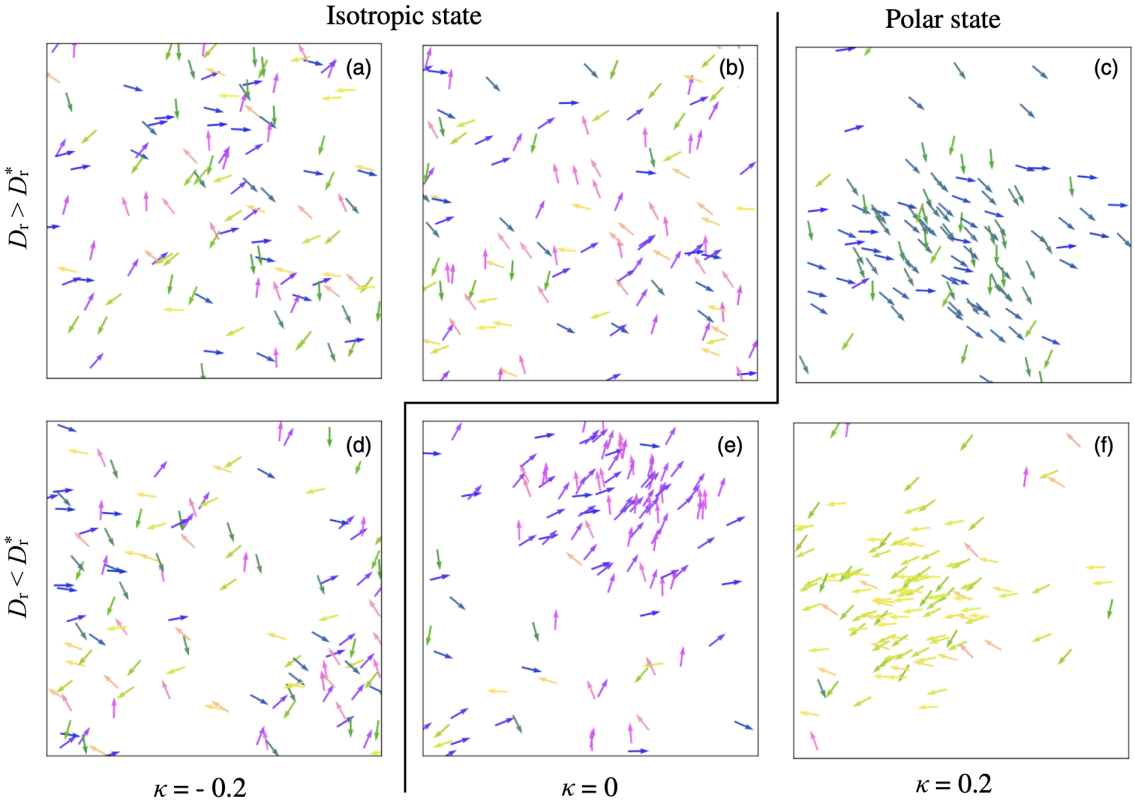



Figure 3.7: Configurations obtained from a direct sampling of the biased ensemble for aligning self-propelled particles. The color code refers to the orientation of particles. In the unbiased dynamics ($\kappa = 0$), we observe isotropic and polar states respectively at large noise ($D_r > D_r^*$) and small noise ($D_r < D_r^*$). Here, the critical noise is $D_r^* = 8$ and we take the noise values $D_r = \{7, 9\}$ for the polar and isotropic regimes, respectively. The dynamical bias leads to renormalized interactions in a controlled manner, which effectively changes the transition threshold as $D_r^* \to D_r^*(1+\kappa)$ at leading order. As a result, one can stabilize either isotropic or polar states, respectively for $\kappa < 0$ and $\kappa > 0$, illustrating the ability to trigger or inhibit collective effects in nonequilibrium systems.

angular dynamics [52]:

$$\dot{\mathbf{r}}_i = V_0 \mathbf{u}(\theta_i), \quad \dot{\theta}_i = \mu_r \sum_j \mathcal{T}(\theta_j - \theta_i, \mathbf{r}_i - \mathbf{r}_j) + \eta_i(t), \tag{3.39}$$

where $V_0$ denotes the self-propulsion velocity, $\mathbf{u}(\theta) = (\cos\theta, \sin\theta)$ is the unit vector, and $\mu_r$ is the rotational mobility. The term $\eta_i$ is a zero-mean Gaussian white noise with correlations

$\langle \eta_i(t)\eta_j(0)\rangle = 2D_{\rm r}\delta_{ij}\delta(t)$ given in terms of the rotational diffusion coefficient $D_{\rm r}$. To promote alignment between neighboring particles, we choose the pair-wise torque as $\mathcal{T}(\theta, \mathbf{r}) = \Theta(\sigma - |\mathbf{r}|)\sin\theta/(\pi\sigma^2)$. This dynamics was originally introduced as a generalization of the Vicsek model to continuous time [189]. Thus, it exhibits a transition between a isotropic state for small density $\rho_0$ and large noise $D_{\rm r}$, and a polar state for large density $\rho_0$ and small noise $D_{\rm r}$. In practice, the linear stability analysis of the corresponding hydrodynamic equations predicts the transition to occur when $2D_{\rm r} = \mu_{\rm r}\rho_0$ [52]. In what follows, our aim is to show that such a transition can also be mediated by tuning interactions with a dynamical bias.

To this end, we take the biasing factor in path probability as $\exp\left[\kappa \int_0^t \mathcal{E}_\theta(s)\mathrm{d}s\right]$, where the biasing observable $\mathcal{E}_\theta$ reads

$$\begin{aligned}\mathcal{E}_\theta = -\frac{1}{2}\sum_{i,j}\left[\frac{\partial}{\partial\theta_i} + \frac{\mu_{\rm r}}{D_{\rm r}}\sum_k \mathcal{T}(\theta_i - \theta_k, \mathbf{r}_i - \mathbf{r}_k)\right] \\ \times \mathcal{T}(\theta_i - \theta_j, \mathbf{r}_i - \mathbf{r}_j).\end{aligned} \tag{3.40}$$

The average value $\langle\mathcal{E}_\theta\rangle$ is proportional to $\frac{\mathrm{d}}{\mathrm{d}t}\sum_{i,j}\langle\cos(\theta_i - \theta_j)\Theta(\sigma - |\mathbf{r}_i - \mathbf{r}_j|)\rangle$, which vanishes in steady state. Then, following the procedure detailed in Sec. 3.6.2, we deduce that biasing the dynamics in Eq. 3.39 with $\mathcal{E}_\theta$ amounts to considering renormalized interactions of the form

$$\tilde{\mathcal{T}} = (1 + \kappa)\mathcal{T} + \mathcal{O}(\kappa^2). \tag{3.41}$$

Thus, by promoting a non-zero average for $\mathcal{E}_\theta$, aligning interactions can be tuned in a controlled manner at first order in $\kappa$. Higher order bias leads to maximizing the squared torque, as presented in Appendix A2.2.

We test this prediction numerically using a direct sampling of biased trajectories. We consider values of $\{D_{\rm r}, \mu_{\rm r}, \rho_0\}$ above and below the threshold $D_{\rm r}^* = \mu_{\rm r}\rho_0/2$, where the system exhibits either isotropic or polar states in the unbiased dynamics ($\kappa = 0$), as shown in Fig. 3.7. Specifically, when the original system is isotropic ($D_{\rm r} > D_{\rm r}^*$), we observe a

transition to polar for $\kappa > 0$, and, conversely, when it is polar ($D_{\mathrm{r}} < D_{\mathrm{r}}^{*}$) a transition to isotropic for $\kappa < 0$. This confirms our result Eq. 3.41 where the bias amounts to changing the angular mobility as $\tilde{\mu}_{\mathrm{r}} = (1+\kappa)\mu_{\mathrm{r}} + \mathcal{O}(\kappa^2)$ for weak $\kappa$, so that the linear instability is either triggered or suppressed by solely tuning $\kappa$, all other parameters being held the same. Thus, these results demonstrate how biasing the dynamics with an appropriate observable leads to control the emergence of spontaneous organization, with potential interest for other nonequilibrium dynamics.

## 3.7 Conclusions

Developing techniques to characterize and control the behavior of systems operating far from equilibrium remains a central and outstanding problem. Despite the apparently complex interplay between internal dissipation and emerging properties, we have demonstrated that tracer diffusion and density correlations can simply be connected to dissipation in driven liquids. We have also constructed a mapping between deterministic and active drives for a specific active matter model, thus showing how our approach can potentially be extended to a broad class of systems.

In practice, monitoring dissipation with a well-defined parameter remains an open challenge for many-body systems. To this end, biased ensembles enable one to specify the statistics of dissipation by introducing an additional control parameter, analogously to the change from micro-canonical to canonical ensemble in equilibrium thermodynamics [27, 90]. This is done by selecting rare noise realizations which drive the system away from typical behaviour, without introducing any driving force. Pioneering works were focused on favoring dynamical heterogeneities, without affecting the structure, of kinetically constrained models [64, 82, 138, 165, 14]. Yet, more recent studies have shown the potential to also modify density correlations in diffusive systems [89, 19, 128].

Using these large-deviation techniques, we have put forward a particular set of biased

ensembles which allows one to regulate the liquid structure in a controlled manner. The explicit form of the bias is motivated by the relations between dissipation and structure that we have derived for driven liquids. At leading order, any bias in this class simply leads to introducing additional interactions in the dynamics. Furthermore, higher-order bias systematically constrains the trajectories to favor the formation of clusters. Based on minimal case studies, we have sampled the biased configurations, using state-of-the-art numerics [66, 174, 86, 129, 143, 94, 15], to illustrate the ability to stabilize specific structures and collective effects in a controlled manner.

Overall, these results illustrate how specifying the amount of energy dissipated by nonequilibrium forces allows one to constrain the dynamics and structure of driven liquids. This paves the way towards controlling the emerging properties of such systems by tuning dissipation accurately. It remains to investigate whether similar results can be obtained in more complex systems which could, for instance, potentially include anisotropic building blocks, such as driven chiral objects or active liquid crystals [92, 200, 130].

# CHAPTER 4

# ENERGY DISSIPATION, STRUCTURE AND MACHINE LEARNING IN STRONGLY INTERACTING ACTIVE MATTER

The material in this chapter is adapted with permission from the following publication:

[181] Laura Tociu et al. "Inferring Dissipation from Static Structure in Active Matter". *ArXiv e-prints* (2020)

## 4.1 Introduction

This chapter describes our mean-field method to relate dissipated energy to structure in active matter systems with strong interactions. Active matter is a class of nonequilibrium systems in which every component consumes energy to produce an autonomous motion [112, 8, 59]. The energy fluxes stemming from individual self-propulsion lead to complex collective behaviors without any equilibrium equivalent, such as a collective directed motion [97] and phase separation despite purely repulsive interactions [134]. The possibility of exploiting such behaviors to design materials with innovative functions has motivated much research [56], with the goal of reliably predicting and controlling the features of active systems.

Minimal models have been proposed to capture active dynamics, for instance in particles with aligning interactions or self-propelled isotropic particles, yielding, respectively, collective motion [25] and motility-induced phase separation [22]. Based on these models, the challenge is to establish a nonequilibrium framework, by analogy with equilibrium statistical thermodynamics, which connects microscopic details and emergent physics. Progress has been made in this direction by characterizing protocol-based observables, such as pressure [176, 163], surface tension [166, 198], and chemical potential [75].

The dissipation induced by microscopic energy fluxes has recently attracted much at-

tention, since it measures the cost to drive the dynamics into nonequilibrium states [93, 153] and to extract work with original protocols [137, 104, 46]. In particular, it has been shown that dissipation constrains the transport of active particles [180, 60], and that changing dissipation with a dynamical bias changes material properties [180], ultimately inducing phase transitions [128, 60, 73]. Moreover, the dissipation can be connected to some mechanical properties, such as the so-called active pressure, of certain isotropic active matter models [163].

Despite recent advancements, understanding how to control quantitatively the dynamics and structure of many-body active systems, for instance by appropriately tuning external parameters, remains largely an open challenge [72]. A large part of the theoretical approaches used to predict the structure of active fluids generally rely on either equilibrium mappings [172, 146, 195, 170] or weak-interaction approximations [180, 60], thus limiting their applicability.

In this work, we use tools from liquid-state theories and machine learning to take on these challenges. We construct a novel mean-field theory whose applicability and ease of implementation surpasses existing approaches, and which quantitatively connects the static two-point density correlations to the rate of energy dissipation. Our results provide a gateway towards controlling the structure of a nonequilibrium many-body system by tuning the dissipation rate, and it demonstrates how artificial intelligence can potentially guide such a control.

The paper is organized as follows. First, we describe the model of active liquids that we consider throughout the manuscript, which is an assembly of self-propelled particles. Second, we derive a novel nonequilibrium mean-field theory that solves the structure of our active liquid with strong interactions, even close to phase transitions. Third, we use our theory to evaluate energy dissipation. Unlike other effective representations of active dynamics [172, 146, 194, 170], we do not rely on any equilibrium approximation, thus allowing

all nonequilibrium features to be retained, and in particular for the dissipation to be properly evaluated. Combining our results on structure and dissipation, we then elucidate the relationship between static density correlations and dissipation over various regimes of activity and interaction strength. Importantly, this relation highlights that dissipation can be deduced by analyzing solely the static structure. In contrast, existing approaches to quantify dissipation rely on evaluating either the violation of fluctuation-response relations [81, 183, 58, 126, 2], currents [5, 68, 69, 103], or irreversibility of trajectories [148, 116]. Lastly, as a proof of principle, we put forward a machine learning architecture, trained with static configurations, which accurately infers the dissipation rate without any information about the underlying dynamics.

## 4.2 Model and Methods

We primarily consider a popular model of active matter consisting of $N$ interacting self-propelled particles, often referred to as Active Ornstein-Uhlenbeck Particles [171, 109, 61], with two-dimensional overdamped dynamics:

$$\dot{\mathbf{r}}_i = -\nabla_i \sum_{j \neq i} U(\mathbf{r}_i - \mathbf{r}_j) + \mathbf{f}_i + \boldsymbol{\xi}_i, \tag{4.1}$$

where $U$ is the pair-wise potential, and the mobility is set to unity. The terms $\{\boldsymbol{\xi}_i, \mathbf{f}_i\}$ embody, respectively, the thermal noise and the self-propulsion velocity. They have Gaussian statistics with zero mean and uncorrelated variances, given by $\langle \xi_{i\alpha}(t)\xi_{j\beta}(0)\rangle = 2T\delta_{ij}\delta_{\alpha\beta}\delta(t)$ and $\langle f_{i\alpha}(t)f_{j\beta}(0)\rangle = (T_{\mathrm{A}}/\tau)\delta_{ij}\delta_{\alpha\beta}e^{-|t|/\tau}$, where $\tau$ is the persistence time. For a vanishing persistence ($\tau = 0$), the system reduces to a set of passive Brownian particles at temperature $T + T_{\mathrm{A}}$. At sufficiently high persistence, the system undergoes a phase separation even with purely repulsive interactions [61].

In order to test the general nature of our results, we also consider actively driven rotors

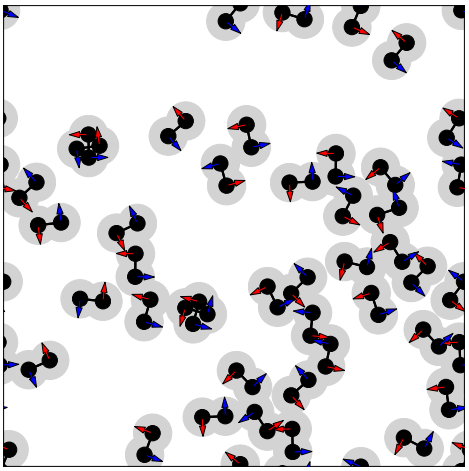


Figure 4.1: Representative snapshot of a system of active rotors interacting between their poles via the Yukawa potential. The inner black and outer light gray circles are centered on each rotor pole. The black circles have a radius of 1, roughly corresponding to the distance at which interaction energy is comparable to thermal energy, and the light gray circles have a radius of 2.5, which is the cutoff radius of the interaction potential. The red and blue arrows indicate the velocity of the rotor poles due to the rotational driving force. Parameters: $\rho_0 = 0.05$, $d = 1.50$, $A = 50$, $\kappa = 4.0$, $T = 1.0$, $v_0 = 25$

(Fig. 4.1) similar to those in [200]. Each rotor consists of 2 identical isotropic point particles kept at a fixed length from each other through an invisible bond of fixed length $d$. The rotor is subject to translational and rotational thermal noise and to conservative forces/torques due to the interactions of its poles with those of neighboring rotors. An active force with constant magnitude induces them to rotate in a counterclockwise direction around their joint center. The equations of motion, in the overdamped regime just as for AOUPs, are as follows:

$$\begin{aligned} \dot{\mathbf{r}}_i &= -\frac{1}{\gamma_t}\nabla_{i,k}\sum_{j\neq i}\sum_{k,l}U(\mathbf{r}_{i,k}-\mathbf{r}_{j,l}) + \boldsymbol{\xi}_{t,i} \\ \dot{\theta}_i &= \frac{v_0}{d/2} + \frac{1}{\gamma_r\,(d/2)^2}\left(\sum_{j\neq i}\sum_{k,l}\left(\mathbf{r}_{i,k}-\mathbf{r}_i\right) - \nabla_{i,k}U(\mathbf{r}_{i,k}-\mathbf{r}_{j,l})\right) + \xi_{r,i} \end{aligned} \tag{4.2}$$

Here, the sum over $k$ ($l$) is over the positions of the 2 poles, which we denote $h$ and $t$, of rotor $i$ ($j$). The center of rotor $i$ is denoted by $\mathbf{r}_i$ and the positions of its poles are

$\mathbf{r}_{i,h} = \mathbf{r}_i + \mathbf{d}/2$ and $\mathbf{r}_{i,t} = \mathbf{r}_i - \mathbf{d}/2$, with $\mathbf{d} = d\left(\cos\left(\theta_{\mathrm{i}}\right), \sin\left(\theta_{\mathrm{i}}\right)\right)$; likewise for rotor $j$. Eq. 4.2 can be simplified by expressing the interaction potential of the rotors as a single anisotropic potential $\tilde{U}\left(\mathbf{r}, \theta\right)$, rather than as 2 identical isotropic potentials centered at points with opposite displacements from the center of the rotor:

$$\begin{aligned} \dot{\mathbf{r}}_i &= -\frac{1}{\gamma_t} \nabla_{\mathbf{r}_i} \sum_{j \neq i} \tilde{U}(\mathbf{r}_i - \mathbf{r}_j, \theta_i - \theta_j) + \boldsymbol{\xi}_{t,i} \\ \dot{\theta}_i &= \frac{2v_0}{d} - \frac{2}{\gamma_r d} \nabla_{\theta_i} \sum_{j \neq i} \tilde{U}(\mathbf{r}_i - \mathbf{r}_j, \theta_i - \theta_j) + \xi_{r,i} \end{aligned} \tag{4.3}$$

The terms $\{\boldsymbol{\xi}_t, \xi_r\}$ represent the translational and rotational thermal noise, respectively. They have Gaussian statistics with zero mean and uncorrelated variances, given by $\langle \xi_{ti\alpha}(t) \xi_{tj\beta}(0) \rangle = 2T\delta_{ij}\delta_{\alpha\beta}\delta(t)/\gamma_t^2$ and $\langle \xi_{ri}(t) \xi_{rj}(0) \rangle = 2T\delta_{ij}\delta(t)/(\gamma_r d/2)^2$.

### *4.2.1 Numerical Simulations*

In this chapter we use three different potentials for the interactions between particles, and we introduce them here. The first one is the Yukawa potential given by $U(\mathbf{r}) = (A/r)e^{-\kappa r}$ for $r < r_0$ where $r_0$ is the cutoff. The second potential we consider is the Weeks-Chandler-Anderson potential given by $v(\mathbf{r}) = 4v_0\big[(\sigma/|\mathbf{r}|)^{12} - (\sigma/|\mathbf{r}|)^6\big]\Theta(2^{1/6}\sigma - |\mathbf{r}|)$ [193]. The third potential we consider is the harmonic potential given by $U(\mathbf{r}) = A(1-r)^2$ for $r < r_0$, where $r_0$ is the cutoff. We will provide the parameter values used in these potentials in the main text and the Figure captions.

For Secs. 4.3-4.4 pertaining to strong interactions and relationship between rate of work and density correlations, the simulations are run in a two-dimensional box $10^2\sigma \times 10^2\sigma$ with periodic boundary conditions, where $\sigma = 1$ is the particle diameter. The time step for harmonic and WCA simulations is $\delta t = 10^{-4}$, and that for Yukawa simulations is $\delta t = 10^{-6}$; in all cases, the initial condition is homogeneous. The density is $\rho_0 = 0.5$ when the Yukawa potential is used, 0.75 when WCA is used, and 1 when the harmonic potential is used. The

pair correlation function was extracted over a range of $r \in [0, 3.2]$ in increments of $dr = 0.01$ for WCA particles, and $r \in [0, L/2]$ for Yukawa and harmonic particles, where $L$ denote here the system size. Between 5 and 25 trials are conducted for each set of parameters.

Whenever the harmonic potential is used, the equations of motion are integrated using a custom code of molecular dynamics, based on finite time difference. The harmonic systems are equilibrated for 500 units of simulation time, corresponding to at least $500\tau$ for all simulations, where $\tau$ is the persistence time of the active noise, and data is collected every 100 units for a duration of 1000 time steps.

Whenever the Yukawa or WCA potentials are used, the simulations are performed using the LAMMPS simulation package with a custom fix to allow for self-propulsion and custom code to measure pair correlation functions and dissipation in separate simulations. In the case of WCA, equilibration is performed for at least $50\tau$, pair correlation functions are measured for at least $100\tau$, and dissipation is measured for at least $25\tau$. In the case of Yukawa, equilibration is performed for $12\tau$, pair correlation functions are measured for at least $24\tau$, and dissipation is measured for $6\tau$.

In Sec. 4.4, calculation of $\tilde{I}$ from simulations is performed by numerically integrating the difference between the $h(r)$ and $h_{\mathrm{eq}}(r)$ histograms from simulations, multiplied by derivatives of the interparticle potential following Eq. 4.20. Calculation of $c_{\mathrm{eq}}(k)$ for theoretical predictions is done by numerically Fourier transforming the portion of the equilibrium $h_{\mathrm{eq}}(r)$ with $r \in [0, 16]$ (note that these simulations take place in a box with side length 40) to obtain $h_{\mathrm{eq}}(k)$, and then computing $c_{\mathrm{eq}}(k)$ using the Ornstein-Zernike relation (shown in Eq. 4.17 extended to non-equilibrium systems).

In Sec. 4.5, the simulations used to extract AOUP snapshots for machine learning were run in a two-dimensional box of size $25\sigma \times 25\sigma$ with periodic boundary conditions, using the custom code referenced previously. The systems were equilibrated for $500\tau$, and snapshots were saved every $\tau$ for a total simulation time of $20\tau$. Multiple trajectories were generated

so as to obtain the 6000 snapshots per each $T_{\mathrm{A}}$ required for training and testing.

The active rotor simulations used for machine learning in Sec. 4.5, as well as the simulations used to relate rate of work and structure of active rotors at the end of Sec. 4.4, were run with 250 rotors in a box with side lengths $L_X, L_Y = 50$ and periodic boundary conditions. Friction constants $\gamma_t$, $\gamma_r$ and temperature $T$ were set to 1. The duration of the simulations was 1000 units of time after an equilibration period of 250 units, with $dt = 4 \cdot 10^{-5}$. $h_R(\mathbf{r})$ was measured every 10 units of time, while the rate of work was tabulated at every timestep. Dynamics were implemented via the stochastic Runge-Kutta (SRK) algorithm, with checks to ensure that rotor poles did not move large distances in a single timestep.

### *4.2.2 Architecture of Continuous Convolutional Neural Network*

The network was built in keras using the Functional API [28]. The individual input to the network consists of interparticle distances. This choice was made because it greatly simplified the implementation of a custom layer in keras. To be more precise, our input consists of an array of size $N \times D$ that contains, for each of the $N$ particles, the distances to its nearest $D$ neighbors. For the simulations with the harmonic potential, $N = 625$ and $D = 20$, while for the simulations with the Yukawa potential, $N = 312$ and $D = 25$. The choice of $D$ was such that distances to around $r = 3$ ($r = 4$ for Yukawa) are mostly captured, but with certainty all distances up to the cutoff ($r = 1$ for harmonic and $r = 2.5$ for Yukawa) are captured. This choice ensures that the algorithm has access to all the distances that we expect are crucial to extracting the rate of work.

The continuous convolutional layer is built as a custom layer and performs the following. For each convolution center, in this case the position of active particle $\mathbf{r}_i$, the convolution operation consists of evaluating the sum $\sum_{j=1}^{D} F(|\mathbf{r}_i - \mathbf{r}_j|)$. Here, $F$ is a learnable function, the equivalent of a filter in traditional convolutional neural networks. We express this function in a basis of $N_G = 30$ ($N_G = 40$ for Yukawa) Gaussian functions centered between

$r = 0$ and $r = 3$ ($r = 4$ for Yukawa) and with a standard deviation of 0.05:

$$F(r) = \sum_{i=0}^{N_G-1} \beta_i e^{-(r-0.1i)^2/(2\cdot 0.05^2)}. \tag{4.4}$$

Finally, the output to the network consists of the rate of work $\dot{w}$, centered as

$$\dot{w}_{\mathrm{c}} = \dot{w} + \alpha\rho_0 \int \left[(\nabla U)^2 - T\nabla^2 U\right] g_{\mathrm{eq}} d\mathbf{r}, \tag{4.5}$$

and scaled as

$$\dot{w}' = \frac{\dot{w}_{\mathrm{c}}}{\max(\dot{w}_{\mathrm{c}}) - \min(\dot{w}_{\mathrm{c}})}, \tag{4.6}$$

where the maximum and minimum of the rate of work is taken over the entire training data set. Centering and scaling the rate of work improves the accuracy and speed of our network. The rate of work is calculated at each value of $T_{\mathrm{A}}$ as an average in the nonequilibrium steady state of Eq. 4.18. Hence, different snapshots at the same value of $T_{\mathrm{A}}$ will be mapped to the same output.

In conclusion, the machine performs the following mapping from the input (interparticle distances $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$) to the output (scaled rate of work):

$$u \sum_{i,j} \sum_{k=0}^{N_G-1} \beta_k e^{-(r_{ij}-0.1k)^2/(2\cdot 0.05^2)} + b = \dot{w}'. \tag{4.7}$$

We enforce particle indistinguishibility by setting a layer constraint that the weights $u$ are identical. The machine is tasked with finding the best $\{u, \beta_k, b\}$ to minimize the deviation between its predicted output, $\dot{w}_{\mathrm{pred}}$, and $\dot{w}'$. We choose to quantify this deviation through a loss function of the form

$$L(\{\dot{w}^a_{\mathrm{pred},i}\}) = \sum_a \left[ \left( \frac{\sum_i \dot{w}^a_{\mathrm{pred},i}}{N_a} \right) - \dot{w}'_a \right]^2, \tag{4.8}$$

where the $a$ indexes the $T_{\rm A}$ values used in the training process, and $N_a$ is the number of snapshots at that value of $T_{\rm A}$ in each training batch over which the loss function is calculated. Our choice of loss function ensures that we are mapping the average of the output of the network at each $T_{\rm A}$ to the rate of work at the same $T_{\rm A}$. Since the rate of work is indeed an average function of the configurations, this choice of loss function was the most suitable for training our model. In the large $N$ limit, we have the following relation:

$$u\sum_{i,j}F(r_{ij}) = Nu\rho_0\int F(r)g(\mathbf{r})d\mathbf{r}. \tag{4.9}$$

If the training is successful and the machine predicts outputs that are very close to $\dot{w}'$, then using Eq. 4.20, along with Eqs. 4.4-4.7, we recover:

$$\begin{aligned} &Nu\rho_0\int Fgd\mathbf{r} + b \\ &\quad = \frac{\alpha\rho_0}{\max(\dot{w}_{\rm c}) - \min(\dot{w}_{\rm c})}\int \big[(\nabla U)^2 - T\nabla^2 U\big]gd\mathbf{r}, \end{aligned} \tag{4.10}$$

yielding

$$\begin{aligned} &b = 0, \quad \gamma F = (\nabla U)^2 - T\nabla^2 U, \\ &\gamma = \frac{Nu}{\alpha}\big[\max(\dot{w}_{\rm c}) - \min(\dot{w}_{\rm c})\big]. \end{aligned} \tag{4.11}$$

The good agreement between $\gamma F$ and $(\nabla U)^2 - T\nabla^2 U$ shown in Fig. 4.7 confirms the relations in Eq. 4.11.

In producing the data of Fig. 4.7(a-b), we prepared 35000 snapshots using the harmonic potential and 30000 snapshots using the Yukawa potential. These correspond to 5000 snapshots per $T_{\rm A}$, where $T_{\rm A}$ ranged from 0 to 12 in increments of 2 for harmonic and from 0 to 30 in increments of 6 for Yukawa. The snapshots were divided into 80% train and 20% validation sets, and the network was trained until we observed convergence of the validation loss (between 200 and 300 epochs of training), using the adam optimizer in keras with batch size 512.

The test data reflected in Figs. 4.7(a-b) comes from independent simulations ran over a more fine-grained range of $T_{\rm A}$'s. Specifically, for harmonic we extracted 1000 configurations per $T_{\rm A}$ with $T_{\rm A}$ ranging over the integers between 0 and 12. For Yukawa we extracted 1000 configurations per $T_{\rm A}$ with $T_{\rm A}$ going between 0 and 30 in increments of 3.

Producing the data of Fig. 4.7(c-d) was similar to (a-b). We produced 5000 snapshots per $v_0$ to train and 1000 snapshots for the test $v_0$'s. The input to the network is an array that contains, for each of $N = 500$ total rotor heads, all the interparticle distances to the nearest $D = 25$ rotor head neighbors except the rotor head that's connected to it. We note that in the case of active rotors, we did not rescale the rate of work as shown in Eqs. 4.5 -4.6. This step was only required in the previous case to improve training and also infer the learned function more easily, but for the active rotors it was not required. Finally, another modification is that the learned function was expressed in a wider basis of $N_G = 60$ Gaussian functions centered between $r = 0$ and $r = 6$ and with a standard deviation of 0.05, to allow for greater distances between rotor heads to become relevant.

For the deep learning in Fig. 4.9, the learning process took 200-300 epochs with a batch size of 64. We used 5000 input arrays for the training and 1000 for the testing. The results are averaged over 10 independent runs where the validation loss successfully leveled off, and the error bars represent the standard deviation.

Codes for molecular dynamics and machine learning can be found at [179].

## 4.3 Nonequilibrium Mean Field Theory

We start by considering the effective dynamics of an active tracer embedded in a bath consisting of the other particles. To derive analytically the statistics of the tracer displacement, inspired by recent works [40, 39], a strategy is to rely on a mean-field approach by considering that interactions among the tracer and the bath are weak. This leads one to scale the interaction strength by a dimensionless factor $\varepsilon$, which can be regarded as a small parameter

for perturbative expansion. The equation of motion of the tracer position $\mathbf{r}_0(t)$ then reads

$$\dot{\mathbf{r}}_0 = \mathbf{f}_0 - \varepsilon \int \nabla_0 U(\mathbf{r}_0 - \mathbf{r}')\rho(\mathbf{r}', t)d\mathbf{r}' + \xi_0, \tag{4.12}$$

where the bath is described in terms of the density field $\rho(\mathbf{r}, t) = \sum_{i=1}^{N} \delta(\mathbf{r} - \mathbf{r}_i(t))$ where $N$ is the number of bath particles. The dynamics of the density field $\rho(\mathbf{r}, t)$, can be readily obtained following the procedure in [36] as:

$$\begin{aligned}\frac{\partial \rho(\mathbf{r}, t)}{\partial t} &= T\nabla^2\rho(\mathbf{r}, t) + \nabla \cdot \left[\sqrt{2\rho T}\Lambda(\mathbf{r}, t) - \mathbf{P}(\mathbf{r}, t)\right] \\ &+ \nabla \cdot \left(\rho \nabla \left[\int U(\mathbf{r} - \mathbf{r}')\rho(\mathbf{r}', t)d\mathbf{r}' + \varepsilon U(\mathbf{r} - \mathbf{r}_0)\right]\right),\end{aligned} \tag{4.13}$$

where $\mathbf{P}$ denotes here the polarization field $\mathbf{P}(\mathbf{r}, t) = \sum_i \mathbf{f}_i(t)\delta(\mathbf{r} - \mathbf{r}_i(t))$. The term $\mathbf{\Lambda}$ is a Gaussian white noise with zero mean and unit variance: $\langle \Lambda_\alpha(\mathbf{r}, t)\Lambda_\beta(\mathbf{r}', t')\rangle = \delta_{\alpha\beta}\delta(\mathbf{r} - \mathbf{r}')\delta(t - t')$. In principle, the dynamics in Eqs. 4.12-4.13 can be solved recursively to obtain the statistics of the density field $\rho(\mathbf{r}, t)$ and of the tracer position $\mathbf{r}_0$. Some of us already took this approach in [180, 60], using a perturbation in the weak interaction limit. In what follows, we extend this approach to characterize the system beyond the regime of weak interactions.

The structure of the system is determined by the two-point correlation of density $h$, defined by $\rho_0 h(\mathbf{r}) = (1/N)\sum_{i\neq j}\langle\delta(\mathbf{r} - \mathbf{r}_i + \mathbf{r}_j)\rangle - \rho_0$, where $\rho_0$ denotes the overall density. In the homogeneous state, where density correlations are evaluated by measuring the average number of particles away from *any* representative tracer, the Fourier transform $h(\mathbf{k}) = \int \mathrm{e}^{i\mathbf{k}\cdot\mathbf{r}} h(\mathbf{r})d\mathbf{r}$ can be written in terms of $\delta\rho = \rho - \rho_0$ as:

$$h(\mathbf{k}) = \frac{1}{\rho_0}\left\langle e^{i\mathbf{k}\cdot\mathbf{r}_0(t)}\delta\rho(\mathbf{k}, t)\right\rangle. \tag{4.14}$$

Our nonequilibrium mean-field theory is built as follows. As described in Appendix A3.1, we first linearize the dynamics (Eqs. 4.12-4.13) and obtain a solution for $\delta\rho$ in the Fourier

domain. Then, we construct an expansion in the coupling parameter $\varepsilon$ to compute the two-point correlation function $h(\mathbf{k})$. We keep terms up to first-order correction in $\varepsilon$, obtaining:

$$h(\mathbf{k}) = -\mathbf{k}^2 \varepsilon U(\mathbf{k}) \frac{G(\mathbf{k}) + T}{G(\mathbf{k})} \int_{-\infty}^{0} ds e^{\mathbf{k}^2((G(\mathbf{k})+T)s + R(s))}, \tag{4.15}$$

where $R(s) = T_{\mathrm{A}} s + T_{\mathrm{A}} \tau (1 - e^{s/\tau})$ and $G(\mathbf{k}) = T + \rho_0 U(\mathbf{k})$. This form is valid only in the weak-interacting regime. To go beyond this regime, we draw inspiration from equilibrium solvation theories [24]. In this context, the density around a tracer particle, which interacts strongly with its neighbors, is captured by considering the convolution between density correlation and direct correlation functions. Besides, linear response for weak interactions enforces that the same information is captured by convoluting density correlation function and interaction potential. Hence, this suggests that the substitution of $U(\mathbf{k})$ with $-Tc_{\mathrm{eq}}(\mathbf{k})$ effectively accounts for higher-order corrections at equilibrium.

To further motivate this substitution, we investigated the response of an equilibrium fluid to a time varying perturbation. Specifically, we simulated a system of particles interacting via the short-ranged but strongly repulsive harmonic potential (see Methods section) with a cutoff of 1 and $A = 64T$. We allowed this system to reach a steady state in the presence of a single fixed external potential $U(\tilde{\mathbf{r}})$ corresponding to a single extra particle fixed at position $\tilde{\mathbf{r}}$, removed this particle at $t = 0$, and measured $\rho(\mathbf{r} - \tilde{\mathbf{r}})$ over time as the system relaxed to a steady state. We Fourier transformed $\rho(\mathbf{r} - \tilde{\mathbf{r}})$ into $h(\mathbf{k}, t)$ and compared it with the predictions obtained both from the linearized dynamical evolution equation for the density [36], Eq. 4.13, and from the same equation but with $U(\mathbf{k})$ replaced by $-Tc_{\mathrm{eq}}(\mathbf{k})$ (Appendix A3.1). These results are described in Fig. 4.2(A). We found that the evolution equation with $U(\mathbf{k})$ replaced by $-Tc_{\mathrm{eq}}(\mathbf{k})$ yielded an accurate prediction for the decay of $h(\mathbf{k}, t)$ in the region of $k$-space around the primary peak of $h_{eq}(\mathbf{k})$, which corresponds to the first peak and dominant wavelength of $\rho(\mathbf{r})$. In contrast, predictions of $h(\mathbf{k}, t)$ made *without* using the substitution of $-Tc_{\mathrm{eq}}(\mathbf{k})$ for $U(\mathbf{k})$ were very poor in this region, as expected

for such strongly interacting particles where the mean-field approximation is expected to break down. Unsurprisingly, predictions with both methods are poor in the small-$k$ region corresponding to the structure of the system on large length scales. Changes in the system structure on the large length scales corresponding to such values of $k$ are connected to the compressibility of the active system.

These results motivate substituting $U(\mathbf{k})$ replaced by $-Tc_{\mathrm{eq}}(\mathbf{k})$ in our non-equilibrium mean field theory as a heuristic approach to account for higher-order corrections. Overall, our theory then leads to the following expression for the density correlations:

$$h(\mathbf{k}) = \mathbf{k}^2 \frac{\hat{G}(\mathbf{k}) + T}{\hat{G}(\mathbf{k})} T c_{\mathrm{eq}}(\mathbf{k}) \int_{-\infty}^{0} ds e^{\mathbf{k}^2[(\hat{G}(\mathbf{k})+T)s + R(s)]}, \tag{4.16}$$

where $\hat{G}(\mathbf{k}) = T\left(1 - \rho_0 c_{\mathrm{eq}}(\mathbf{k})\right)$. At equilibrium ($T_{\mathrm{A}} = 0$), Eq. 4.16 becomes the famed Ornstein-Zernike relation. Out of equilibrium ($T_{\mathrm{A}} \neq 0$), our prediction (Eq. 4.16) can be used to deduce the structure of the system, given by $h(\mathbf{k})$, based solely on measurements of the equilibrium structure, namely for a given $h_{\mathrm{eq}}(\mathbf{k})$ and $c_{\mathrm{eq}}(\mathbf{k})$. Note that the perturbation theory leading to Eq. 4.16 ignores the effect of the polarization term in Eq. 4.13. In a previous work, which provides an expression for $\dot{w}$ in a weakly interacting system [60], some of us have already reported that the numerical contribution from such terms is negligible in the range of systems and regimes examined there. We surmise that these contributions are small under the set of assumptions, approximations and regimes that we employ in the present paper, too. Our numerical results support this hypothesis.

To compare our mean-field prediction with numerical results, we introduce the *nonequilibrium direct correlation function*, denoted by $c$ and defined as:

$$c(\mathbf{k}) = \frac{h(\mathbf{k})}{1 + \rho_0 h(\mathbf{k})}. \tag{4.17}$$

This definition can be regarded as an extension of the Ornstein-Zernike relation for equilib-

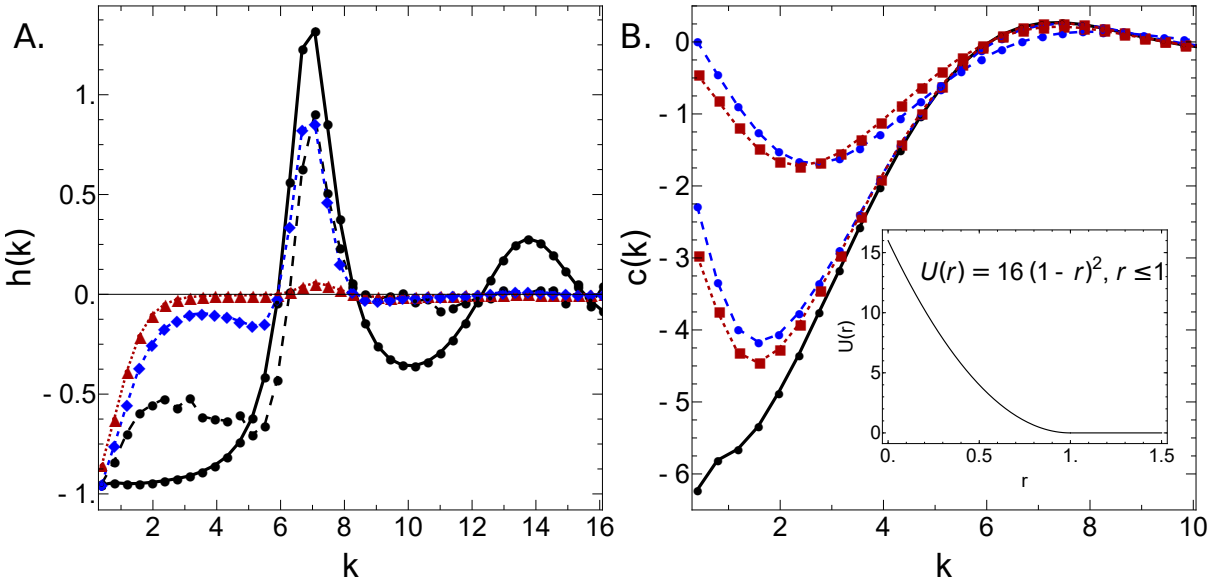



Figure 4.2: Mean-field theory for nonequilibrium structure: Prediction for the nonequilibrium direct correlation function $c$, defined in Eq. 4.17, as a function of wavenumber $k = |\mathbf{k}|$. (A) Relaxation of AOUPs predicted from $c_{eq}(\mathbf{k})$: prediction of change in $h(\mathbf{k})$ in response to a perturbing potential after $0.02\tau$ has elapsed for a system of AOUPs interacting via strong harmonic potentials. Except at small values of k, we find good agreement between measured $h(\mathbf{k})$ (dashed black line with circles) and $h(\mathbf{k})$ predicted from $c_{eq}(\mathbf{k})$ (dashed blue line with diamonds), while predictions using $U(\mathbf{k})$ (dashed red line with triangles) are very poor. The steady-state $c_{eq}(\mathbf{k})$ (solid black line) is shown as a reference. Parameters: $\rho_0 = 1.0$, $A = 64$, $\tau = 1.0$, $T = 1$. (B) System of harmonic particles driven by active forces. The predicted curves for $c$ (dashed red lines with squares) are compared with simulation results (dashed blue lines with circles) both near equilibrium ($T_{\mathrm{A}} = 2$, lower pair of dashed lines) and far from equilibrium ($T_{\mathrm{A}} = 40$, higher pair of dashed lines). The equilibrium reference $c_{\mathrm{eq}}$ measured numerically (solid black line), which is used as an input for the mean-field prediction, is shown as a reference. The inset displays the harmonic interaction potential between the AOUPs. The good agreement between predictions and simulations demonstrates that our mean-field theory captures well the deviation from equilibrium structure. In particular, it reproduces quantitatively the effective attraction at large wavelengths/small wavenumbers, which arises due to active forces. Interestingly, while the $c_{eq}\,(\mathbf{r})$-based prediction for the *relaxation of passive particles to a steady state* over short time scales in (A) is poor in the small-k regime, this does not mean that the $c_{eq}\,(\mathbf{r})$-based theory is poor at predicting the structure of *steady states of active particles* in the same small-k regime in (B). Parameters: $\rho_0 = 1.0$, $A = 16$, $\tau = 0.4$, $T = 1$.

rium liquids. Note that $c$ can no longer be related to any free-energy *a priori*. In Fig. 4.2(B), we plot the predicted $c$, as deduced from Eqs. 4.16-4.17, along with estimates obtained from simulations. Again, let us emphasize that the only input for our prediction is the equilibrium direct correlation function $c_{\rm eq}$.

## 4.4 Dissipation, Structure and Strong Interactions

We now use our mean-field theory to study connections between structure and dissipation of active liquids. Following standard definitions of stochastic thermodynamics [157, 155], the dissipation is defined as the rate of work done by the thermostat on particles: $\mathcal{J} = \sum_i \langle \dot{\mathbf{r}}_i \cdot (\dot{\mathbf{r}}_i - \boldsymbol{\xi}_i) \rangle$. It can be decomposed as a constant term, independent of interactions, and a contribution from interactions: $\mathcal{J} = N(T_{\rm A}/\tau - \dot{w})$, where we have introduced the *active work* per particle [180, 60, 93]:

$$\dot{w} = \frac{1}{N} \sum_{i,j \neq i} \left\langle \mathbf{f}_i \cdot \nabla_i U(\mathbf{r}_i - \mathbf{r}_j) \right\rangle. \tag{4.18}$$

The expression in Eq. 4.18 illustrates that, for a generic many-body active system, any particle-based evaluation of dissipation requires measuring the local polarization $\mathbf{f}_i$. This is notoriously difficult in experimental systems of isotropic active particles, for which the local driving direction is not encoded in the particle shape.

To demonstrate that the active work can be generically related to the static structure, we derive an expression for $\dot{w}$ by following our mean-field approach. As described in Appendix A3.2, the first term in this expansion leads to the following expression:

$$\begin{aligned} \dot{w} = \rho_0 T_{\rm A} \int \frac{d\mathbf{k}}{(2\pi)^d} \mathbf{k}^4 \frac{\hat{G}(\mathbf{k}) + T}{\hat{G}(\mathbf{k})} T c_{\rm eq}(\mathbf{k}) U(\mathbf{k}) \\ \times \int_{-\infty}^{0} ds e^{\mathbf{k}^2 [(\hat{G}(\mathbf{k}) + T)s + R(s)]} (1 - e^{s/\tau}), \end{aligned} \tag{4.19}$$

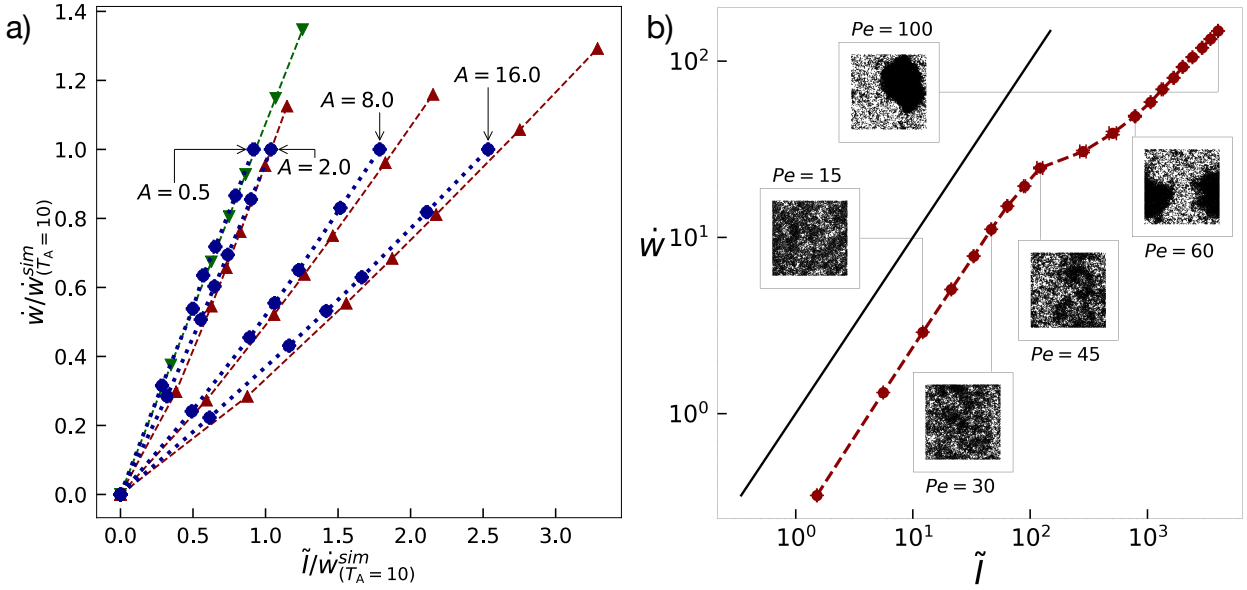



Figure 4.3: Connecting dissipation and structure. (a) Theory and simulation results, when displayed in parametric plot for the rate of the work $\dot{w}$ and the structural observable $\tilde{I}$ (see Eq. 4.20), lead to the linear relation $\dot{w} = \alpha \tilde{I}$, where the slope $\alpha$ varies with the interaction strength $A$. For weak interactions ($A = 0.5$), the slopes predicted by our mean-field theory (green), as described in the main text, and that measured numerically (blue) agree very well. For stronger interactions ($A > 0.5$), our mean-field theory leads to predicted slopes (red) that are close to that of simulations (blue). Overall, these results demonstrate that the dissipation-structure relation is robust across different values of $A$, and that our mean-field prediction performs well in capturing such a relationship. Parameters: Harmonic potential, $\rho_0 = 1$, $\tau = 1$, $T_{\rm A} \leq 10$, $T = 1$. (b) The linear relation between $\dot{w}$ and $\tilde{I}$, as measured numerically, holds up to values of the Péclet number Pe that are within 25% of that at which phase transition occurs. The insets show snapshots of the system at different values of Pe. The black solid line is a guideline corresponding to $\dot{w} = \tilde{I}$. Parameters: WCA potential, $\rho_0 = 0.75$, $\tau = 1/3$, $T = 1$, and $\mathrm{Pe} \equiv (\sigma/T)\sqrt{T_{\rm A}/\tau}$ with $\sigma = 1$.

where $d$ is the spatial dimension, and $\{R, \hat{G}\}$ are defined as in Eq. 4.15 and Eq. 4.16, respectively. In Ref. [180], some of us showed empirically that the rate of work $\dot{w}$ can be connected to the two-point density correlation function $h(\mathbf{r})$ through the relation:

$$\dot{w} = \alpha\rho_0 \int \left[(\nabla U)^2 - T\nabla^2 U\right](h - h_{\rm eq})d\mathbf{r} \equiv \alpha\tilde{I}, \tag{4.20}$$

where the factor $\alpha$ is a function of the interaction potential $U$, the overall density $\rho_0$, and the persistence time $\tau$. Crucially, $\alpha$ is independent of the Péclet number $\mathrm{Pe} \equiv (\sigma/T)\sqrt{T_{\rm A}/\tau}$, which compares the relative strength of the active drive with respect to thermal fluctuations, where $\sigma$ denotes here the typical interaction range and is taken as unity in what follows. The empirical connection in Eq. 4.20 was reported in the previous chapter for particles with strong interactions, using a Weeks-Chandler-Anderson (WCA) potential. We now demonstrate that the mean-field results in Eq. 4.16 and Eq. 4.19 provide an analytical justification for this connection.

In Fig. 4.3(a), we plot the estimates of $\dot{w}$ and $\tilde{I}$ obtained from simulations and from our analytical mean-field theory. We first use a short-ranged harmonic potential with interaction strength $A$, as given in the previous Section. For every $A$, the simulations are performed at various values of $T_{\rm A}$ within the range $[2, 10]$, with fixed persistence $\tau = 1$ and density $\rho_0 = 1$. The simulation results exhibit a linear connection between $\dot{w}$ and $\tilde{I}$ for all values of $A$ (blue lines). This observation generalizes the results from the previous chapter, showing that the connection between dissipation and structure is robust for a breadth of regimes, from weak to strong interactions.

In the weak interacting limit, namely at small $\varepsilon$, it is possible to derive a compact expression for $\tilde{I}$ in terms of the interaction potential $U$ (see Appendix A3.4), which we use to plot the theory curve at $A = 0.5T$, the green line in Fig. 4.3(a). For higher values of $A$ (red lines in Fig. 4.3(a)), we numerically integrate the expression of $\tilde{I}$ in Eq. 4.20 by substituting $h$ as given in Eq. 4.16, and by using $c_{\rm eq}$ measured in simulations. The analytical predictions

for $\dot{w}$ and $\tilde{I}$ deviate moderately from the simulation data, which might be expected given the errors in the predicted structure at short wavelengths, as discussed in the previous Section. Importantly, despite this slight deviation, our predictions are still able to reproduce very well the value of the slope $\alpha$ for each interaction strength $A$. This illustrates that our mean-field theory, designed to capture nonequilibrium liquid properties, accurately predicts the dissipation-structure relation inherent to active liquids.

The linear connection between $\dot{w}$ and $\tilde{I}$ eventually fails at larger values of $T_{\mathrm{A}}$ for harmonic potential (Fig. 4.4). We anticipate that this stems from the fact that such a potential allows for a complete overlap of particles at sufficiently strong driving. Considering liquids with a steeper repulsion among particles, such as either Yukawa or WCA potentials, this linear connection is still present within a broader regime of $T_{\mathrm{A}}$ (Fig. 4.4). In particular, we show in Fig. 4.3(b) that the relation between dissipation and structure persists even close to the phase transition at high Péclet number, thus highlighting again the robustness of this relationship.

To determine whether or not the connection between $\dot{w}$ and $\tilde{I}$ is strictly limited to AOUPs, we also considered actively driven rotors similar to those in [200] (see Methods). $\tilde{I}$ and $\dot{w}$ for these rotors could be defined analogously to how these quantities were defined for AOUPs (see Appendix A3.5) and were measured from simulations at multiple active rotation speeds for rotors of various lengths. While the mode of activity of these rotors is qualitatively different from AOUPs and our mean-field theory is not directly applicable to them, we nevertheless found that for all but the shortest active rotors, $\tilde{I}$ and $\dot{w}$ obeyed a linear relationship (Fig. 4.5). Even for this sub case, as we demonstrate in the next section, our machine learning results point to the existence of a connection between the configurational structure and dissipation. Together, these results suggests that the connection between dissipation and structure is robust within a sizable class of simple active matter systems.

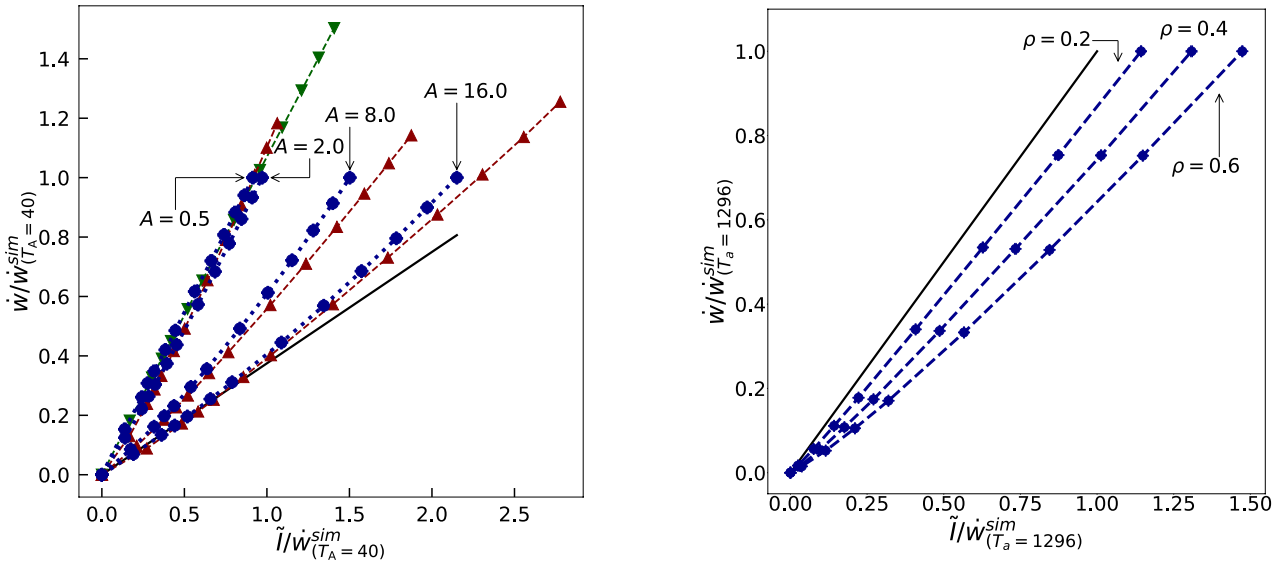



Figure 4.4: (Left) Simulation results for $\dot{w}$ vs $\tilde{I}$ for particles with the harmonic potential (dashed blue lines) at various amplitudes (labeled) scaled by the value of $\dot{w}$ at $T_{\rm A} = 40$. Theoretical predictions obtained using Eq. 5.135 are shown for $A = 0.5$ (dashed green, downward triangles). Theoretical predictions using $c_{\rm eq}(\mathbf{k})$ are shown for $A = 2, 8, 16$ (dashed red, upward triangles). This figure is an extension of Fig. 4.3(a) to larger values of $T_{\rm A}$. The solid black line suggests what a linear relationship between $\dot{w}$ and $\tilde{I}$ at A = 16 would look like if their ratio at $T_{\rm A} = 5$ held for all values of $T_{\rm A}$. Deviations of both theoretical and simulation results from this line illustrate the breakdown of the linear relationship as $T_{\rm A}$ increases. Parameters: harmonic potential, $\rho_0 = 1$, $\tau = 1$, $T_{\rm A} \in [0, 40]$. (Right) Simulation results for $\dot{w}$ vs $\tilde{I}$ for particles with the Yukawa interparticle potential (dashed blue lines) at various densities (labeled). The ratio between $\dot{w}$ vs $\tilde{I}$ is essentially constant for values of $T_{\rm A}$ ranging from weak to moderately strong driving, with an increasing value of the slope as $\rho$ increases. The solid black line corresponds to a slope of 1. Parameters: Yukawa potential, A = 50, $\kappa = 4.0$, $\tau = 1$, $T_{\rm A} \in [0, 1296]$.

## 4.5 Machine Learning Rate of Work

The results in Fig. 4.3 demonstrate that in certain parameter regimes, the dissipation rate can be inferred by analyzing *solely* the two-point density correlation which characterizes static configurations. A natural question is then whether it would be possible to train a machine learning algorithm to "learn" the dissipation rate from snapshots of active matter.

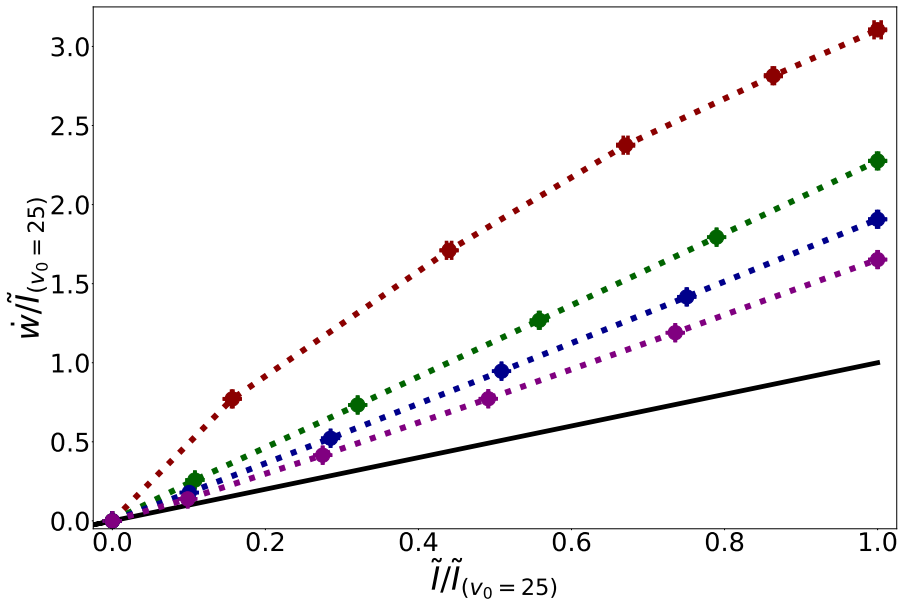



Figure 4.5: Simulation results for $\dot{w}$ vs $\tilde{I}$ for active rotors of various lengths (scaled by the value of $\tilde{I}$ at $v_0 = 25$ for rotors of a given length). Rotors of different lengths are represented by dotted lines of different colors: d = 1.50 (red, top dotted line), d = 2.50 (green, second dotted line), d = 3.50 (blue, third dotted line), d = 5.50 (purple, bottom dotted line). Solid black line corresponds to the relationship $\dot{w} = \tilde{I}$. Point particles on each half of each rotor interact through the Yukawa potential. Parameters: Yukawa potential, A = 50, $\kappa = 4.0$, $T = 1.0$, $v_0 \in [0, 25]$

The field of machine learning applications in chemistry, physics and materials science is a relatively new but flourishing one [20, 18, 152]. More closely related to our intended application, a lot of work has been done recently towards using machine learning in active matter [29]. Specifically, a variety of tools have been used to infer patterns, phase boundaries or equations of motions from images or movies of nonequilibrium systems [21, 199, 16, 71, 44]. Convolutional neural networks have also been recently used in Ref. [154] to show how the dissipation rate is connected to time-reversal symmetry violations.

Inspired by these recent successes, the purpose of this Section is to construct neural networks that can map snapshots of particle positions to the underlying dissipation induced by their dynamics. Our goal is to show that the dissipation of generic active systems can

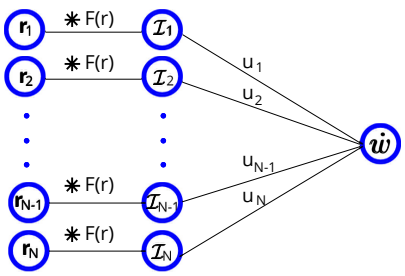



Figure 4.6: The machine learning architecture consists of two layers. First, a convolutional layer, which performs the operation $\mathbf{r}_i * F \equiv \sum_{j \neq i} F(|\mathbf{r}_i - \mathbf{r}_j|)$, where $j$ runs over a set number of nearest neighbors. The function $F$ is expressed in a basis of Gaussian functions with coefficients learned by the algorithm, see Methods section. The second layer then fully connects the convolutional layer $\mathcal{I}_i \equiv \mathbf{r}_i * F$ to the rate of work $\dot{w}$ through the weight $u = u_1 = u_2 = \cdots = u_N$ and bias $b$: $u \sum_i \mathcal{I}_i + b = \dot{w}'$, where $\dot{w}'$ is the scaled rate of work used as the output of the network.

be inferred by analyzing solely the *static* configurations of the system, without need of any underlying information about the details of the dynamics. Such a proof of principle, which highlights again the simple relation between dissipation and structure in active matter, may then form the basis for the development of feedback control processes, where the dissipation rate is tuned adaptively until the desired structure is achieved.

To this end, we design a network consisting of a continuous convolutional layer followed by a fully connected layer, as described in Fig. 4.6, that maps the input data given by position vectors $\{\mathbf{r}_i\}$ to the output data given by the scaled rate of work $\dot{w}'$, which is related to the physical work $\dot{w}$ through a simple deterministic transformation (see Methods Sec. 4.2.2). We perform this transformation in order to enhance accuracy and speed of learning. Due to particle indistinguishibility, we constrain the weights of the fully connected layer to be equal. The continuous convolutional layer is inspired by a recent work on rotationally and translationally invariant networks for atomic systems [177]. This layer effectively scans the neighbors of each particle, and it sums up the result of applying a learnable function $F(\mathbf{r})$ to each of a number of nearby neighbors, where $\mathbf{r}$ denotes here the position vector from the particle of interest to its neighbor, see Methods Sec. 4.2.2. This operation is similar to a

usual convolution on a grid, with the particle of interest being analogous to the convolution center, and the function $F$ to a filter.

Inspired by the relationship between dissipation and pair correlation functions in Eq. 4.20, we constrain the form of $F$ to be a function of the distance between particles $r = |\mathbf{r}|$. Recognizing that the relation between rate of work $\dot{w}$ and density correlations $h$ given in Eq. 4.20 is also analogous to a convolution, the function $F$, scaled by an overall deterministic factor $\gamma$ described in Methods Sec. 4.2.2, can be expected to be learned as $(\nabla U)^2 - T\nabla^2 U$.

In Figs. 4.7(a-b), we show the agreement between the values of $\dot{w}$ measured using the particle-based definition in Eq. 4.18, referred to as "exact" values, and that predicted from static configurations of the system using our neural network. We consider separately configurations taken from either harmonic or Yukawa AOUP particles over a finite range of $T_{\mathrm{A}}$. Each data point is averaged over 10 independent machine learning training cycles. The neural network is able to infer the rate of work with good accuracy, which confirms that static configurational data is indeed sufficient to evaluate dissipation. Note that, in the prediction phase, the neural network is supplied with configurations generated at values of $T_{\mathrm{A}}$ that are never used in the training phase.

We further demonstrate in Figs. 4.7(a-b) (inset plot) how our continuous convolution neural network is able to predict a form of the function $F$ that is remarkably close to the shape of the function $(\nabla U)^2 - T\nabla^2 U$. Our simple network, which is designed with the minimum number of elements to capture a relationship reminiscent of Eq. 4.20, succeeds in learning that same relationship with great accuracy. We stress again that our approach does not require any information regarding the microscopic polarization, and it also does not rely on measuring any current, at variance with [5, 68, 69, 103].

In Figs. 4.7(c-d), we perform similar machine learning experiments using active rotors with rotor head distance of 1.5 and 5.5, respectively. We vary $v_0$ over a finite range, and teach a network identical to that used for AOUPs to predict rate of work from datasets

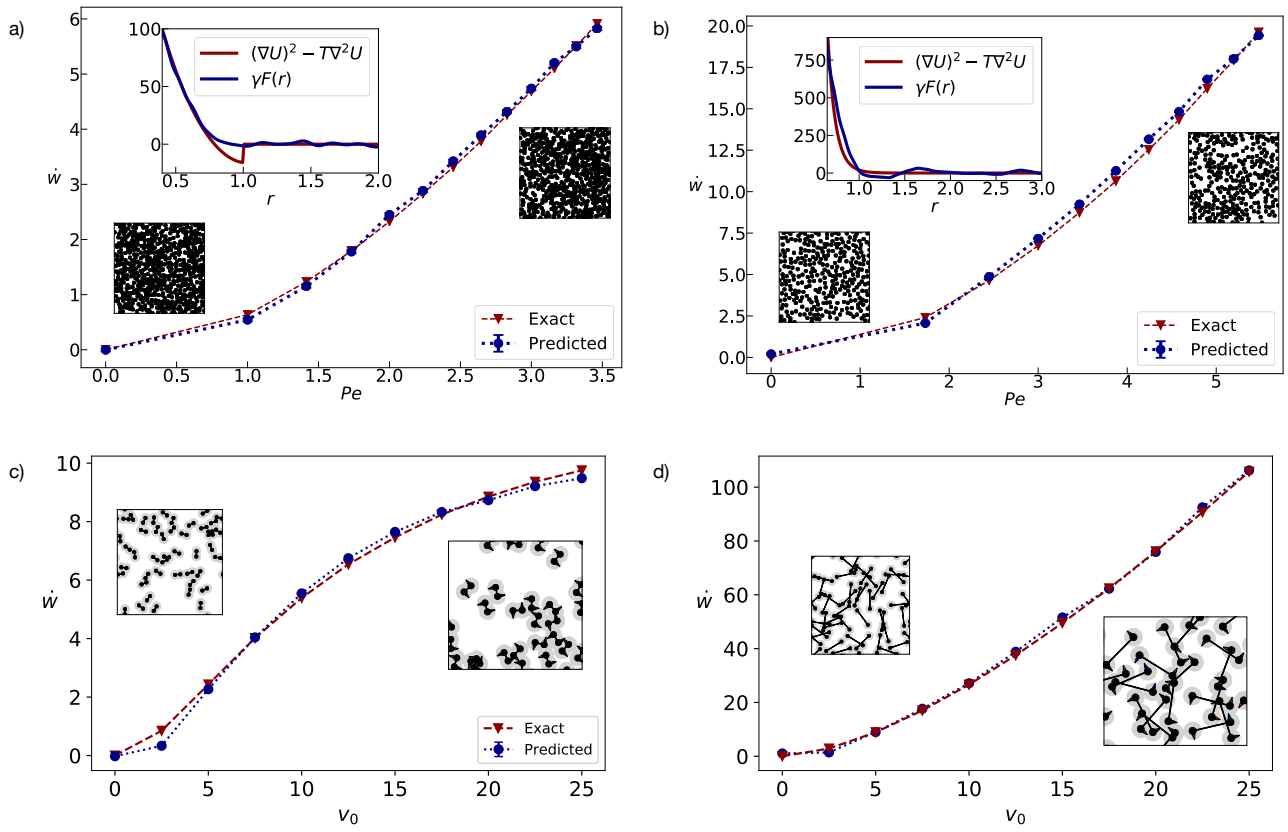



Figure 4.7: Inferring dissipation from static configurations. (a-b) The machine is trained using AOUP configurations generated at even values of $T_{\rm A} \in [0, 12]$, and the values of $\dot{w}$ are predicted at both odd and even values within this range. (c-d) The machine is trained using active rotors configurations generated at values of $v_0$ spaced every 5 in the interval $[0, 25]$, and the values of $\dot{w}$ are predicted at values spaced every 2.5 within this range. The agreement between predicted and exact values of $\dot{w}$ is excellent. This demonstrates that the machine learning algorithm is able to learn dissipation from static snapshots, without any prior information about the underlying dynamics, for both AOUPs and, remarkably, even for active rotors with short distances. (a-b Insets) We compare the learned function $F$ scaled by an overall deterministic factor $\gamma$, as described in Methods Sec. 4.2.2, with $(\nabla U)^2 - T\nabla^2 U$, showing that the algorithm reproduces the dissipation-structure relation given in Eq. 4.20 without any prior knowledge of such a relationship. Parameters for AOUPs: (a) Harmonic potential, $\rho_0 = 1$, $A = 8T$, $\tau = 1$, $T = 1$, (b) Yukawa potential, $\rho_0 = 0.5$, $A = 50T$, $\tau = 1$, $\kappa = 4$, $T = 1$. Insets show representative snapshots at $T_{\rm A} = \{0, 10\}$ (Harmonic), and $T_{\rm A} = \{0, 30\}$ (Yukawa). As in the main text, we use $\mathrm{Pe} \equiv (\sigma/T)\sqrt{T_{\rm A}/\tau}$ with $\sigma = 1$. Parameters for active rotors: Yukawa potential, $\rho_0 = 0.1$, $A = 50T$, $\kappa = 4$, $T = 1$, (c) $d = 1.50$, (d) $d = 5.50$. Insets show representative snapshots at $v_0 = \{5, 25\}$.

of purely static interparticle distances. The network is able to map the configurations to the rate of work even for the short rotor head distance of 1.5, where the empirical linear relationship between rate of work and pair correlation function breaks down.

In Fig. 4.7(d), we expect to be able to match the learned function $F(r)$ to a rescaled $(\nabla U)^2 - \nabla^2 U$. However, in Fig. 4.7(c), our results suggest that there exists a function $F'(r)$, different from $(\nabla U)^2 - \nabla^2 U$, that maps between structure and rate of work. This insight is novel and entirely discovered with the help of machine learning.

We note that the machine learning network we use could be allowed to learn more complex filters, for instance, by considering filters decomposed as products of radial functions and spherical harmonics [177]. However, when we allow the network to learn more complex filters in the upcoming Section, we do not achieve notable improvements.

## 4.6 Other Machine Learning Architectures for Active Matter

In order to better understand the success of our convolutional neural network approach applied a variety of active systems, we tested architectures other than the one in Fig. 4.6 to infer rate of work from snapshots of the rotors.

Until now, we had constrained the learned function to be radial. We also considered more general functions, expressed as sums of products of learnable radial functions and circular harmonics. In Fig. 4.8, we show the machine learning fit using a learnable function of the form $F(\mathbf{r}) = F_1(r) + F_2(r)\sin(\theta) + F_3(r)\cos(2\theta)$, where $\theta$ is the angle to the neighbor and the three radial functions are each learned as sums of 60 Gaussians. This is a decomposition in terms of the first three circular harmonics. We found that little is gained by moving past the simplest case of a radial function.

Lastly, since the design of our convolutional network cannot capture correlations beyond two-body, we were interested to see whether a deep neural network could learn rate of work directly from the positions of the rotors themselves. Based on the results in the previous

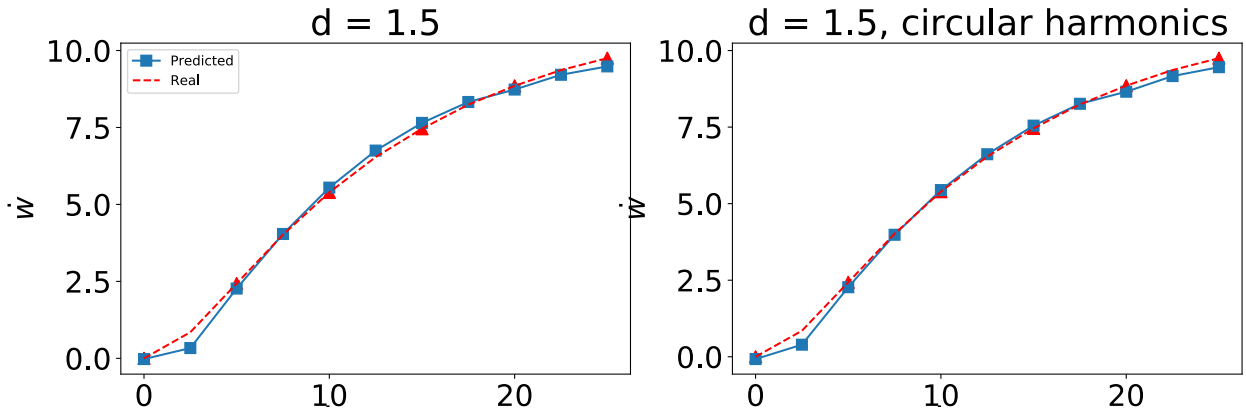



Figure 4.8: Results from machine learning runs using as input active rotor inter-particle distances and learning rate of work from them. The rotor length is $d = 1.5$. The blue square markers indicate the predicted rate of work after the network is successfully trained. The dashed red line shows the real curve of rate of work versus $v_0$, as computed from simulation, while the red triangles indicate the values of $v_0$ at which the network was trained. The training data was obtained for $v_0$ between 0 and 25 in increments of 5, and the testing data in increments of 2.5. For the figure on the right, we allowed the learned function $F(r)$ from Fig. 4.6 to have angular degrees of freedom by expressing it as $F_1(r) + F_2(r)\sin(\theta) + F_3(r)\cos(2\theta)$, where $\theta$ is the angle to the neighbor and the three radial functions are each learned as sums of Gaussians. The parameters are the same as those in Fig. 4.7(c).

chapter on the connection between rate of work and two- and three-body interactions in active fluids, we surmise that a similar connection exists in the case of the rotors. A deep neural network, functioning as a universal function approximator, ought to be able to capture it. We see in Fig. 4.9 that this is indeed possible but it is tricky to design a deep network that works as well as our simple convolutional network. Despite multiple architectures and attempts, the best result we could achieve is not as good of a fit as with the convolutional network.

We attempted various deep network designs. The most successful network (see Appendix Fig. A2.1) takes as input arrays of size $N \times 2D$, where $N = 250$ is the number of rotors and $D = 40$ is the number of nearest rotors we consider. On each of the $N$ row vectors, the input contains the $D$ distances and $D$ angles to the nearest $D$ rotors. We did not include

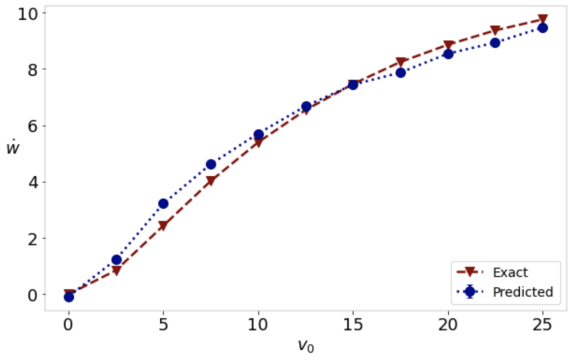



Figure 4.9: Result of training the architecture in Fig. A2.1(A) on a system of 250 active rotors with length $d = 1.5$. Parameters are the same as in Fig. 4.7(c).

the whole system size as an attempt to reduce the memory required to store and handle all coordinates. We did, however, consider more nearest neighbors (up to 100), but we ran into convergence problems and decreased performance.

The deep neural network we designed has 2 fully connected layers of size 16 and 32, respectively, followed by a dropout layer. The dropout layer randomly inactivates a fraction of neurons (in our case, 30%) to prevent overfitting. The next layer has 16 fully-connected neurons, followed by another dropout layer that inactivates 30% of neurons. The last layer of the deep network fully connects all the previous layers to a single neuron. However, since our input has $N = 250$ row vectors, each vector will result in an output neuron, which need to be flattened into a final vector. The final fully connected layer maps into a single value, the rate of work. Just like in the convolutional architecture used so far, we require that the weights mapping the flattened final layer of size $N = 250$ to the dissipation rate be equal, but this is not necessary. The neural network will learn well regardless, but by imposing equal weights, we can interpret what the network is doing. Namely, the deep network of size 16 x 32 x 16 learns a multi-dimensional function $F$ of the coordinates of a particle $i$'s neighbor, $F(\vec{x}_i) = F(r_{i,1}, r_{i,2}, \ldots, r_{i,D}, \theta_{i,1}, \theta_{i,2}, \ldots, \theta_{i,D})$ and finally, the flattening and final fully connected layer perform an average over all the particles of this learned function.

## 4.7 Conclusions

Our results demonstrate that the dissipation induced by driving forces in active liquids can be inferred generically by analyzing the static structure of the system. Although it is well-known that active forces affect the emerging structure [72], reliably predicting the nonequilibrium structure of active systems has remained largely an outstanding problem [172, 146, 195, 170]. In this work, we propose a mean-field theory to quantitatively predict the two-point density correlations even in systems with strong interactions, on which basis we show that there exists a robust relation between dissipation and structure.

Furthermore, we show that neural networks are able to accurately infer the dissipation by learning only from snapshots of positions. This result illustrates the promise that neural networks, when suitably trained, can capture generic connections between microscopic features and macroscopic variables [20, 18]. It also solves the outstanding problem of how to reliably evaluate dissipation based on time-independent and easily accessible data. Indeed, the local polarization of active particles is difficult to access in experimental settings, in particular for isotropic particles, so that obtaining an accurate particle-based measurement of dissipation is generally challenging [183, 58, 2]. Besides, measuring currents to bound dissipation, as proposed in [5, 68, 69, 103], has only limited applications to active systems without any obvious observable current, such as the one considered here.

Last but not least, we want to emphasize that a deep neural network could be designed to directly learn a multi-variate function of the particle coordinates that yields dissipation rate. We explore this avenue and indeed, we succeed in designing a deep neural network that infers dissipation from structure. Nonetheless, the performance is significantly lowered compared to our convolutional network, the learning process is not robust (some runs result in diverging or flat losses), and finding a successful architecture was a laborious process. Thus, we suggest that our approach here, where we combine some empirical knowledge with state-of-the art machine learning techniques such as convolutional layers, is a far superior way

to infer macroscopic quantities from microscopic data than traditional deep neural networks.

# CHAPTER 5

# FUTURE DIRECTIONS

I expect future work, in my group and elsewhere, to further generalize and exploit the relations between work, structure and diffusion elaborated on in this dissertation. In this concluding chapter I offer a non-exhaustive list of possible future avenues or systems where the ideas in this dissertation could prove fruitful.

The results in Sec. 3.5 open promising perspectives to evaluate dissipation simply from the structure of the system. Inspired by recent works [126, 31], one could also introduce a map of dissipation, directly related to the statistics of interaction forces, to resolve spatially where energy is released in the thermostat. Though the corresponding integrated map would not cover the total dissipation, it would already provide insightful information about locations of low and high dissipation with respect to a constant background set by the squared driving amplitude.

In Sec. 3.6, I motivate functional forms that are used to bias ensembles of trajectories away from equilibrium dynamics, towards a nonzero average amount of energy exchange with the heat bath. Since dynamical bias consists in favoring rare noise fluctuations, the corresponding dynamics effectively provides useful insights on how to promote atypical configurations with an external drive. In practice, the driving protocol should simply mimic the biased noise realizations. This line of thought has already been exploited for efficient sampling of the biased ensemble [129, 127, 15, 53], where control forces make rare events become typical. Moreover, since the analytic framework encompasses the case of a specific bias for each pair of particle, it could potentially be regarded as a fruitful route to promote the spontaneous self-assembly of complex structures at the cost of energy dissipation. For instance, inspired by recent works [123, 169], one might consider applying this approach to design energetic landscapes, in terms of the pair-specific bias parameters, which selectively stabilize some target molecules.

The idea of dynamical bias to modulate structure that originated in Sec. 3.6 is also currently being applied to cytoskeletal networks. Cytoskeletal networks are active systems that rely on complex component interactions to undergo structural changes in response to external stimuli. Recent work conducted by my colleagues Alexandra Lamtyugina and Yuqing Qiu has shown that dynamic biasing of energy dissipation in an actin-myosin cytoskeletal network induces a structural transformation (i.e. transitions between aster-like and bundle-like configurations) similar to the one observed when a property of myosin motors, their rigidity, is modulated. These findings could materialize into a novel non-equilibrium thermodynamic control principle for altering the properties of complex materials through modulation of the system's energy budget.

In Sec. 4.4, a novel relationship between structure and energy dissipation was uncovered. It would be interesting to explore whether such a relation extends to other types of active liquids, such as for instance liquids with aligning interactions among the particles [25], or with driving forces that sustain a permanent spinning of the particles [200, 164]. The machine learning analysis performed on the system of active rotors suggests that robust relations between dissipation and structure might be generalized to these contexts. Moreover, a similar dissipation-structure relation holds when the system is driven by external deterministic forces [180]. This suggests that dissipation can be potentially inferred from structure for a large class of nonequilibrium liquids, even beyond active matter.

Interestingly, the mean-field theory developed in Sec. 4.3 can be extended in a systematic manner beyond the present study. Indeed, it relies mostly on tools of liquid-state theory [24, 79, 36], which are agnostic to the details of the driving forces. Thus, I believe that this approach can serve as a basis for developing perturbation theories in generic nonequilibrium liquids [124]. This would open the door to anticipating how density correlations are modified by *any* type of driving forces, as a first step towards externally controlling the emerging structure with a specific drive [48, 132, 83].

Importantly, the work described in this dissertation culminates with a machine learning application in active matter, whereby a custom continuous convolutional neural network is trained to infer dissipation rate from purely static snapshots of active particles. Machine learning applications in physics, and specifically active matter, are flourishing and have a promising future [29, 44, 30, 154]. In particular, reinforcement learning has recently seen applications in this field [26, 65, 120, 45] due to its ability to navigate high-dimensional spaces of spatio-temporal activity patterns and quickly develop system-specific intuition. Of relevance is a recent study [50] where reinforcement learning is used in a Vicsek-like model to induce net transport in a specific direction, using a spotlight that induces self-propulsion locally. Yuqing Qiu, a colleague of mine, and I have been collaborating to implement a reinforcement learning architecture in a cytoskeleton actin-myosin network. As mentioned in a previous paragraph, it has already been shown that biasing the dynamics of this system using energy dissipation induces transitions between its two structural forms, aster and bundle. The goal of the reinforcement learning would be to accomplish this transition in a different way, by teaching a learning agent to systematically change pulsating control parameters such as binding rates and myosin rigidity in order to change the structure of the system. Combining insight from the biased dynamics work with the results of the reinforcement learning will take the research efforts in this dissertation to new horizons. It will allow for a comparison of the two ways of achieving structural transitions through trajectory modulation (by biasing with energy dissipation or reinforcement learning) in order to gather novel insight into how energy dissipation carves structure in complex nonequilibrium matter where system-specific intuition is hard to come by.

# APPENDICES

## A1 Appendix for Chapter 2

### *A1.1 Expression for the Effective Spring Constants in the Gaussian Representation*

In this section, we provide expressions for the effective spring constant $K_G$ using ideas from liquid state theory.

We begin by writing down a Hamiltonian appropriate for Gaussian density fluctuations in a liquid [24]:

$$\beta H = \frac{\beta}{2} \int d\mathbf{r} \int d\mathbf{r}' \delta\rho(\mathbf{r}) \chi^{-1}(\mathbf{r}, \mathbf{r}') \delta\rho(\mathbf{r}') \tag{5.1}$$

$\chi^{-1}$ is the inverse of the variance, given by:

$$\chi(|\mathbf{r} - \mathbf{r}'|) = \left\langle \delta\rho(\mathbf{r}) \delta\rho(\mathbf{r}') \right\rangle = \bar{\rho}\delta(\mathbf{r} - \mathbf{r}') + \bar{\rho}^2[g(|\mathbf{r} - \mathbf{r}'|) - 1] \tag{5.2}$$

where $g(r)$ is the pair correlation function and $\bar{\rho}$ is the density of the fluid.

We approximate the inverse as follows [187, 24]

$$\int \int \chi^{-1} \approx \frac{1}{\int \int \chi} \,. \tag{5.3}$$

To do so we need to compute:

$$\int d\mathbf{r} \int d\mathbf{r}' \chi(|\mathbf{r} - \mathbf{r}'|) = \int d\mathbf{r} \int d\mathbf{r}' (\bar{\rho}\delta(\mathbf{r} - \mathbf{r}') + \bar{\rho}^2[g(|\mathbf{r} - \mathbf{r}'|) - 1]). \tag{5.4}$$

The first term is:

$$\int d\mathbf{r} \int d\mathbf{r}' \bar{\rho}\delta(\mathbf{r} - \mathbf{r}') = \int d\mathbf{r} \bar{\rho} = V\bar{\rho} \tag{5.5}$$

To find the second term, we first perform a change of variable. The correlation depends

only on the magnitude $|\mathbf{r}-\mathbf{r}'|$. Let $\mathbf{r}''=\mathbf{r}-\mathbf{r}'$, then:

$$\int d\mathbf{r}\int d\mathbf{r}'\bar{\rho}^2[g(|\mathbf{r}-\mathbf{r}'|)-1]=\int d\mathbf{r}'\int d\mathbf{r}''\bar{\rho}^2[g(r'')-1] \tag{5.6}$$

$$=V\rho\int d\mathbf{r}\bar{\rho}[g(r)-1]. \tag{5.7}$$

From the zero frequency limit of the Ornstein-Zernicke equation we have the textbook result:

$$1+\bar{\rho}\int d\mathbf{r}[g(r)-1]=\frac{1}{1-\bar{\rho}\int d\mathbf{r}c(r)} \tag{5.8}$$

where $g(r)$ and $c(r)$ are the radial distribution function and direct pair correlation function, respectively. Using this:

$$\int d\mathbf{r}\int d\mathbf{r}'\chi(|\mathbf{r}-\mathbf{r}'|)=\frac{V\bar{\rho}}{1-\bar{\rho}\int d\mathbf{r}c(r)}\,. \tag{5.9}$$

We can now compute the inverse, $\int\int\chi^{-1}$:

$$\int\int\chi^{-1}=\frac{1-\bar{\rho}\int d\mathbf{r}c(r)}{V\bar{\rho}} \tag{5.10}$$

$$=\frac{1}{V\bar{\rho}}\int d\mathbf{r}\left[\delta(\mathbf{r}-\mathbf{r}')-\bar{\rho}c(r)\right] \tag{5.11}$$

$$=\int d\mathbf{r}\int d\mathbf{r}'\left[\frac{\delta(\mathbf{r}-\mathbf{r}')}{\bar{\rho}}-c(\mathbf{r},\mathbf{r}')\right] \tag{5.12}$$

Substituting this back in to the Hamiltonian and taking the functional derivative with respect to $\delta\rho(\mathbf{r})$ gives the effective force on $\delta\rho(\mathbf{r})$:

$$\frac{\delta H}{\delta(\delta\rho)}=\beta\left[\frac{\delta\rho(\mathbf{r})}{\bar{\rho}}-\int d\mathbf{r}'c(\mathbf{r},\mathbf{r}')\delta\rho(\mathbf{r}')\right] \tag{5.13}$$

Assuming that $\delta\rho$ is not spatially dependent, this looks like a linear restoring force:

$$\frac{\delta H}{\beta\delta(\delta\rho)} \approx \delta\rho\left[\frac{1}{\bar{\rho}} - \int d\mathbf{r}c(r)\right] \equiv K_G\delta\rho = F_G \tag{5.14}$$

where $K_G$ is the effective spring constant of the Gaussian density fluctuations. Further simplification can be achieved by using the Percus-Yevick closure [24] to write down an approximate expression for $c(r)$,

$$K_G = \left[\frac{1}{\bar{\rho}} - \int d\mathbf{r}c(r)\right] = \left[\frac{1}{\bar{\rho}} + \int d\mathbf{r}g(r)(e^{\beta V(r)} - 1)\right], \tag{5.15}$$

where $V(r)$ denotes the pairwise potential between particles. The $1/\bar{\rho}$ term in Eq. 5.15 is an *ideal gas* contribution.

### *A1.2 Proof $\langle\mathbf{F}^2\rangle = \langle\mathbf{F}^2\rangle_0$ For a Liquid in Equilibrium*

Assuming we have fixed a particle with position $\mathbf{r}_1$ at the origin, by Taylor expanding the total force $\mathbf{F}_1$ acting on this particle term by term we get the equation below in the $x$ direction and an analogous equation in the other directions:

$$F_{1,x}(t+\Delta t) = F_{1,x}(t) + \sum_{i=1}^{N}\frac{\partial F_{1,x}(t)}{\partial x_i}\dot{x}_i(t)\Delta t + \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{\partial^2 F_{1,x}(t)}{\partial x_i\partial x_j}\dot{x}_i(t)\dot{x}_j(t)(\Delta t)^2 + O(\Delta t^3) \tag{5.16}$$

Taking the dot product with the noise (in 2D to exemplify but easily generalizable to other dimensions) we obtain:

$$\lim_{\Delta t\to 0}\eta_1(t)\cdot\mathbf{F}_1(t+\Delta t) = \lim_{\Delta t\to 0}(\eta_{1,x}(t),\eta_{1,y}(t))\cdot(F_{1,x}(t+\Delta t),F_{1,y}(t+\Delta t)) \tag{5.17}$$

$$= \lim_{\Delta t\to 0}\Big(\eta_{1,x}(t)F_{1,x}(t)+\eta_{1,y}(t)F_{1,y}(t)+ \sum_{i=1}^{N}\eta_{1,x}(t)\frac{\partial F_{1,x}(t)}{\partial x_i}\dot{x}_i(t)+\eta_{1,y}(t)\frac{\partial F_{1,y}(t)}{\partial y_i}\dot{y}_i(t)\Delta t+O(\Delta t^2)\Big) \tag{5.18}$$

Terms of order $\Delta t^2$ or higher in Eq. 5.18 vanish as $\Delta t$ goes to 0. Using the Langevin equation $\dot{x}_i(t)=\frac{1}{\gamma}F_{i,x}(t)+\eta_{i,x}(t)$, it follows that:

$$\langle\mathbf{F}^2\rangle_0 = -\frac{\gamma}{2}\lim_{\Delta t\to 0}\eta(t)\cdot\mathbf{F}(t+\Delta t)$$

$$= -\frac{\gamma}{2}\lim_{\Delta t\to 0}\frac{\partial F_{1,x}(t)}{\partial x_1}\eta_{1,x}^2(t)+\frac{\partial F_{1,y}(t)}{\partial y_1}\eta_{1,y}^2(t)\Delta t \tag{5.19}$$

$$= \frac{\gamma}{2}\lim_{\Delta t\to 0}\left(\frac{\partial^2 U}{\partial x_1^2}\frac{dW_x^2}{\Delta t}+\frac{\partial^2 U}{\partial y_1^2}\frac{dW_y^2}{\Delta t}\right) \tag{5.20}$$

$$= K_B T_1^2 U \tag{5.21}$$

$$= K_B T\rho\int^2 u(\mathbf{r})g(\mathbf{r})d\mathbf{r} \tag{5.22}$$

In the above, we defined $\langle dW_x^2\rangle\equiv\langle(\eta_x\Delta t)^2\rangle = 2D_0\Delta t = \frac{2K_BT}{\gamma}\Delta t$.

Let us now expand the term $\mathbf{F}^2$. Again by focusing on a single particle with position $\mathbf{r}_1$ fixed at the origin and doing similar manipulations we obtain:

$$\mathbf{F}^2 = ({}_1U(\mathbf{r}))^2$$

$$= \rho\int(u(\mathbf{r}))^2 g(\mathbf{r})d\mathbf{r}$$

$$+\rho^2\int\int u(\mathbf{r})\cdot u(\mathbf{r}')g_3(\mathbf{r},\mathbf{r}')d\mathbf{r}d\mathbf{r}' \tag{5.23}$$

To aid us in evaluating Eq. 5.23, we assume the system is in equilibrium with a Boltzmann

distributed steady state, and use the Yvon-Born-Green hierarchy:

$$g(\mathbf{r})u(\mathbf{r}) = -K_B T g(\mathbf{r}) - \rho \int d\mathbf{r}' u(\mathbf{r}') g_3(\mathbf{r}, \mathbf{r}') \tag{5.24}$$

Plugging this expression into Eq. 5.23 cancels the term containing the three-point correlation function, yielding:

$$\mathbf{F}^2 = -K_B T \rho \int u(\mathbf{r}) \cdot g(\mathbf{r}) d\mathbf{r} \tag{5.25}$$

Integrating by parts and setting boundary terms to zero gives the final expression for $\langle \mathbf{F}^2 \rangle$:

$$\mathbf{F}^2 = K_B T \rho \int^2 u(\mathbf{r}) g(\mathbf{r}) d\mathbf{r} \tag{5.26}$$

The expression is equal to that in Eq. 5.22, completing the proof that $\langle \mathbf{F}^2 \rangle = \langle \mathbf{F}^2 \rangle_0$ at equilibrium.

### *A1.3 Expression for Diffusion from Perturbation Theory*

The expression for diffusion in terms of force-force and noise-force correlations is:

$$D - D_0 = \frac{1}{d} \int_0^\infty \left[ \left\langle \frac{\mathbf{F}(0) \cdot \mathbf{F}(t)}{\gamma^2} \right\rangle + \left\langle \frac{\boldsymbol{\eta}(0) \cdot \mathbf{F}(t)}{\gamma} \right\rangle \right] dt \tag{5.27}$$

To set up our perturbative approach to computing the quantities in Eq. 5.27, we start from the overdamped Langevin equation:

$$\dot{\mathbf{r}} = \frac{h}{\gamma} \tilde{\mathbf{F}}(\mathbf{r}) + \tilde{\eta}(t) = \frac{h}{\gamma} \left( \tilde{\mathbf{F}}_c(\mathbf{r}) + Pe \tilde{\mathbf{F}}_d(\mathbf{r}) \right) + \tilde{\eta}(t) \tag{5.28}$$

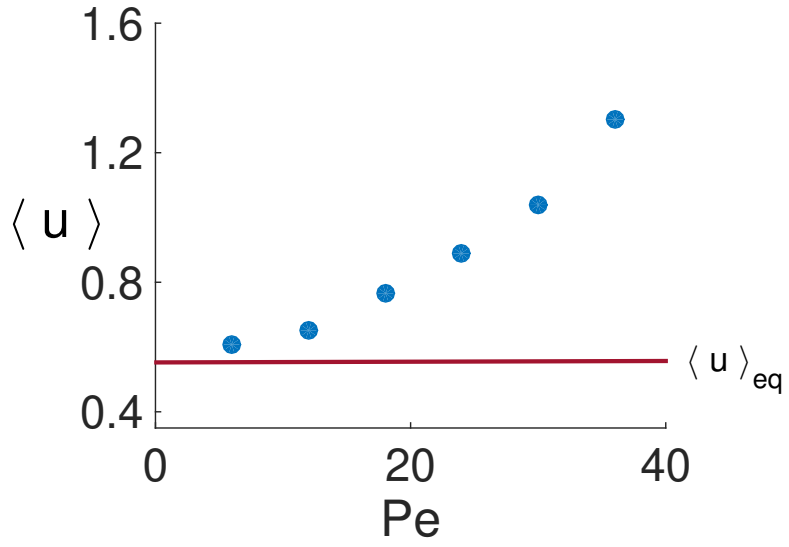



Figure A1.1: The effective energy felt by a tracer active particle in the full many body liquid scales as $Pe^2$. This energy is higher than the equilibrium average, showed as a red line, for all values of Pe sampled. The simulations were done at a density of 0.45, with 10% active particles, and $\tau = 20$.

where $\tilde{\mathbf{F}}_c(\mathbf{r}) = -\nabla\tilde{U}(\mathbf{r})$ is the conservative force and $\tilde{\mathbf{F}}_d(\mathbf{r}) = \nabla \times \tilde{\mathbf{A}}(\mathbf{r})$ the dissipative force. $\tilde{U}(\mathbf{r})$ and $\tilde{\mathbf{A}}(\mathbf{r})$ obey $\int \tilde{U}(\mathbf{r})d\mathbf{r} = 0$ and $\nabla \cdot \tilde{\mathbf{A}}(\mathbf{r}) = 0$. We will be working in the limit $Pe\,, h << 1$. The probability distribution for this system obeys the Fokker-Planck equation:

$$\frac{\partial P}{\partial t} = -h\vec{\nabla} \cdot \left[\tilde{\mathbf{F}}_c P\right] + D\nabla^2 P - hPe\vec{\nabla} \cdot \left[\tilde{\mathbf{F}}_d P\right]. \tag{5.29}$$

When $Pe = 0$, the system is in equilibrium and Eq. 5.29 is solved simply by the Boltzmann distribution $P_0 = \dfrac{e^{-\beta h U(\mathbf{r})}}{Z}$, where $Z$ is the partition function. We construct a perturbation theory in $Pe$ by considering the latter term in Eq. 5.29 as a perturbation to the equilibrium distribution. The first correction to the equilibrium distribution is then given by:

$$P_1 = Pe \sum_i \frac{|u_i\rangle\langle v_i|\tilde{\mathbf{F}}_d \cdot \nabla|P_0\rangle}{\lambda_i} \tag{5.30}$$

where the vectors $|u_i\rangle$ and $\langle v_i|$ are the right and left eigenvectors of $W_0$, respectively, and $\lambda_i$

are the eigenvalues. The total probability for our system is then given by:

$$P(\mathbf{r}) = \frac{e^{-\beta h U(\mathbf{r})}}{Z} + hPe \sum_i \frac{|u_i\rangle\langle v_i|\tilde{\mathbf{F}}_d \cdot \nabla |P_0\rangle}{\lambda_i} \tag{5.31}$$

To proceed with evaluating the expression for the diffusion coefficient perturbatively. Defining $\epsilon_1 \equiv h$ and $\epsilon_2 \equiv Peh$, we will evaluate terms to second order in $\epsilon_{1,2}$. In order for mixed third order terms such as $\epsilon_1^2\epsilon_2$ to be smaller than all quadratic we need $h < Pe$ in our setup. We will average the force-force correlations and the noise-force correlations only over the noise here. The averaging over the realizations of the force are implied at the end and shown as angular brackets in Section 2.7. We first consider the expression for $\int_0^\infty \langle \tilde{\mathbf{F}}(0) \cdot \tilde{\mathbf{F}}(t)\rangle dt$:

$$h^2 \int_0^\infty \langle \tilde{\mathbf{F}}(0) \cdot \tilde{\mathbf{F}}(t)\rangle dt = h^2 \int_0^\infty dt \int d\mathbf{r} \frac{P(\mathbf{r})}{Z} \int \frac{d\mathbf{q}}{(2\pi)^d} \int \frac{d\mathbf{q}'}{(2\pi)^d} \tilde{\mathbf{F}}(\mathbf{q}) \cdot \tilde{\mathbf{F}}(\mathbf{q}') e^{i(\mathbf{q}+\mathbf{q}')\cdot\mathbf{r}} \tag{5.32}$$

$$\langle e^{\int_0^t i\mathbf{q}'\cdot\tilde{\eta}(t')dt' + h/\gamma \int_0^t i\mathbf{q}'\cdot\tilde{\mathbf{F}}(t')dt'} \rangle$$

Keeping terms to order $\epsilon_{1,2}^2$ we have:

$$h^2 \int_0^\infty \langle \tilde{\mathbf{F}}(0) \cdot \tilde{\mathbf{F}}(t)\rangle dt \approx h^2 \frac{1}{V} \int \frac{d\mathbf{q}}{(2\pi)^d} \int \frac{d\mathbf{q}'}{(2\pi)^d} \tilde{\mathbf{F}}(\mathbf{q}) \cdot \tilde{\mathbf{F}}(\mathbf{q}') \delta(\mathbf{q}+\mathbf{q}') \int_0^\infty \langle e^{\int_0^t i\mathbf{q}'\cdot\tilde{\eta}(t')dt'} \rangle dt \tag{5.33}$$

$$= h^2 \frac{1}{V} \int \frac{d\mathbf{q}}{(2\pi)^d} \tilde{\mathbf{F}}(\mathbf{q}) \cdot \tilde{\mathbf{F}}(-\mathbf{q}) \int_0^\infty e^{-|\mathbf{q}|^2 D_0 t} dt \tag{5.34}$$

$$= h^2 \frac{1}{D_0 V} \int \frac{d\mathbf{q}}{(2\pi)^d} \frac{1}{|\mathbf{q}|^2} \big( \tilde{\mathbf{F}}_c(\mathbf{q}) \cdot \tilde{\mathbf{F}}_c(-\mathbf{q}) + 2Pe \tilde{\mathbf{F}}_c(\mathbf{q}) \cdot \tilde{\mathbf{F}}_d(-\mathbf{q}) +$$

$$Pe^2 \tilde{\mathbf{F}}_d(\mathbf{q}) \cdot \tilde{\mathbf{F}}_d(-\mathbf{q}) \big) \tag{5.35}$$

The noise-force correlation term is given by:

$$\int_0^\infty dt\langle\tilde{\eta}(0)\cdot\tilde{\mathbf{F}}(t)\rangle = \int_0^\infty dt\int d\mathbf{r}\frac{P(\mathbf{r})}{Z}\int\frac{d\mathbf{q}}{(2\pi)^d}\langle\tilde{\eta}(0)\cdot\tilde{\mathbf{F}}(\mathbf{q})e^{i\mathbf{q}\cdot\mathbf{r}}e^{\int_0^t i\mathbf{q}\cdot\tilde{\eta}(t')dt'+h/\gamma\int_0^t i\mathbf{q}\cdot\tilde{\mathbf{F}}(t')dt'}\rangle \tag{5.36}$$

The first term of order $\epsilon_{1,2}^2$, which we call $T_1$, comes from $P(\mathbf{r})$, which we expand in terms of $h$:

$$T_1 = \frac{h^2}{V}\int d\mathbf{r}(-\beta\tilde{U}(\mathbf{r}))\int\frac{d\mathbf{q}}{(2\pi)^d}e^{i\mathbf{q}\cdot\mathbf{r}}\int_0^\infty\langle\tilde{\eta}(0)\cdot\tilde{\mathbf{F}}(\mathbf{q})e^{\int_0^t i\mathbf{q}\cdot\tilde{\eta}(t')dt'}\rangle dt \tag{5.37}$$

$$= \frac{-\beta h^2}{V}\int\frac{d\mathbf{q}}{(2\pi)^d}\int\frac{d\mathbf{q}'}{(2\pi)^d}\tilde{U}(\mathbf{q}')\delta(\mathbf{q}+\mathbf{q}')\times$$

$$\int_0^\infty\langle\tilde{\eta}(0)\cdot\tilde{\mathbf{F}}(\mathbf{q})e^{\int_0^{\Delta t} i\mathbf{q}\cdot\tilde{\eta}(t')dt'}e^{\int_{\Delta t}^t i\mathbf{q}\cdot\tilde{\eta}(t')dt'}\rangle dt \tag{5.38}$$

$$= \frac{-2D_0\beta h^2}{V}\int\frac{d\mathbf{q}}{(2\pi)^d}\tilde{U}(-\mathbf{q})i\mathbf{q}\cdot\tilde{\mathbf{F}}(\mathbf{q})\int_0^\infty e^{-|\mathbf{q}|^2D_0t}dt \tag{5.39}$$

$$= \frac{-2\beta h^2}{V}\int\frac{d\mathbf{q}}{(2\pi)^d}\frac{1}{|\mathbf{q}|^2}\left(\tilde{\mathbf{F}}_c(-\mathbf{q})\right)\cdot\tilde{\mathbf{F}}(\mathbf{q}) \tag{5.40}$$

$$= -\frac{2\beta h^2}{V}\int\frac{d\mathbf{q}}{(2\pi)^d}\frac{1}{|\mathbf{q}|^{\mathbf{2}}}\left(\tilde{\mathbf{F}}_c(-\mathbf{q})\cdot\tilde{\mathbf{F}}_c(\mathbf{q})+Pe\tilde{\mathbf{F}}_c(-\mathbf{q})\cdot\tilde{\mathbf{F}}_d(\mathbf{q})\right) \tag{5.41}$$

The second term quadratic in $\epsilon_{1,2}$, $T_2$, comes from the force present in the exponential in Eq. 5.36. Ignoring constants, this term can be written as:

$$T_2 \propto h^2 \int d\mathbf{r} \int \frac{d\mathbf{q}}{(2\pi)^d} \int \frac{d\mathbf{q}'}{(2\pi)^d} e^{i(\mathbf{q}+\mathbf{q}')\cdot\mathbf{r}} i\mathbf{q}\cdot\tilde{\mathbf{F}}(\mathbf{q}')\times$$
$$\int_0^\infty dt \int_0^t dt' \langle \tilde{\mathbf{F}}(\mathbf{q})\cdot\tilde{\eta}(0) e^{\int_0^t i\mathbf{q}\cdot\tilde{\eta}(\tau)d\tau} e^{\int_0^{t'} i\mathbf{q}'\cdot\tilde{\eta}(\tau)d\tau}\rangle \tag{5.42}$$
$$= h^2 \lim_{\Delta t\to 0} \int \frac{d\mathbf{q}}{(2\pi)^d} i\mathbf{q}\cdot\tilde{\mathbf{F}}(-\mathbf{q}) \int_0^\infty dt \int_0^{\Delta t} dt' \langle \tilde{\mathbf{F}}(\mathbf{q})\cdot\tilde{\eta}(0) e^{\int_{t'}^t i\mathbf{q}\cdot\tilde{\eta}(\tau)d\tau}\rangle \tag{5.43}$$
$$= h^2 \lim_{\Delta t\to 0} \int \frac{d\mathbf{q}}{(2\pi)^d} i\mathbf{q}\cdot\tilde{\mathbf{F}}(-\mathbf{q}) \int_0^\infty dt \langle \tilde{\mathbf{F}}(\mathbf{q})\cdot\tilde{\eta}(0) e^{\int_0^t i\mathbf{q}\cdot\tilde{\eta}(\tau)d\tau}\rangle \Delta t \tag{5.44}$$
$$= 2h^2 \lim_{\Delta t\to 0} \int \frac{d\mathbf{q}}{(2\pi)^d} i\mathbf{q}\cdot\tilde{\mathbf{F}}(-\mathbf{q}) i\mathbf{q}\cdot\tilde{\mathbf{F}}(\mathbf{q}) \frac{1}{|\mathbf{q}|^2}\Delta t = 0 \tag{5.45}$$

The term linear in $h$, let's call it $T_3$, can be computed in a similar way, resulting in the following expression:

$$T_3 \propto h \int_0^\infty dt \int d\mathbf{r} \int \frac{d\mathbf{q}}{(2\pi)^d} \langle \tilde{\eta}(0)\cdot\tilde{\mathbf{F}}(\mathbf{q}) e^{i\mathbf{q}\cdot\mathbf{r}} e^{\int_0^t i\mathbf{q}\cdot\tilde{\eta}(t')dt'}\rangle \tag{5.46}$$
$$= 2h \int d\mathbf{r} \int \frac{d\mathbf{q}}{(2\pi)^d} \frac{1}{|\mathbf{q}|^{\mathbf{2}}} e^{i\mathbf{q}\cdot\mathbf{r}} i\mathbf{q}\cdot\left(\tilde{\mathbf{F}}_c(\mathbf{q}) + Pe\tilde{\mathbf{F}}_{\mathbf{d}}(\mathbf{q})\right) \tag{5.47}$$
$$= 2h\left(\int d\mathbf{r}\tilde{U}(\mathbf{r}) + Pe\int \frac{d\mathbf{q}}{(2\pi)^d} \frac{1}{|\mathbf{q}|^{\mathbf{2}}} i\mathbf{q}\cdot\tilde{\mathbf{F}}_{\mathbf{d}}(\mathbf{q})\right) \tag{5.48}$$
$$= 2h\left(\int d\mathbf{r}\tilde{U}(\mathbf{r}) + Pe\int \frac{d\mathbf{q}}{(2\pi)^d} \frac{1}{|\mathbf{q}|^{\mathbf{2}}} i\mathbf{q}\cdot i\mathbf{q}\times\tilde{\mathbf{A}}(\mathbf{q})\right) \tag{5.49}$$
$$= 2h\left(\int d\mathbf{r}\tilde{U}(\mathbf{r}) + Pe\int \frac{d\mathbf{q}}{(2\pi)^d} \frac{1}{|\mathbf{q}|^{\mathbf{2}}} i\mathbf{q}\times i\mathbf{q}\cdot\tilde{\mathbf{A}}(\mathbf{q})\right) = 0 \tag{5.50}$$

Plugging Eq. 1.57 and Eq. 5.41 into the expression Eq. 5.27, and using $\tilde{\mathbf{F}}_{\mathbf{c}}(\mathbf{q}) = -i\mathbf{q}\tilde{U}(\mathbf{q})$ and $\tilde{\mathbf{F}}_{\mathbf{d}}(\mathbf{q}) = i\mathbf{q}\times\tilde{\mathbf{A}}(\mathbf{q})$ yields the desired relationship:

$$D - D_0 = \frac{D_0 h^2}{(K_B T)^2 dV}\left(Pe^2\int |\tilde{\mathbf{A}}(\mathbf{r})|^2 d\mathbf{r} - \int \tilde{U}^2(\mathbf{r})d\mathbf{r}\right) \tag{5.51}$$

We can also use this perturbation theory to obtain expressions for the quantities $\langle\tilde{\mathbf{F}}^2\rangle$ and $\langle\tilde{\mathbf{F}}^2\rangle_0$:

$$\langle\tilde{\mathbf{F}}^2\rangle = \frac{h^2}{V}\int d\mathbf{r}\int\frac{d\mathbf{q}}{(2\pi)^d}\int\frac{d\mathbf{q}'}{(2\pi)^d}\tilde{\mathbf{F}}(\mathbf{q})\cdot\tilde{\mathbf{F}}(\mathbf{q}')e^{i(\mathbf{q}+\mathbf{q}')} \tag{5.52}$$

$$= \frac{h^2}{V}\int\frac{d\mathbf{q}}{(2\pi)^d}\left(\tilde{\mathbf{F}}_c(\mathbf{q})\cdot\tilde{\mathbf{F}}_c(-\mathbf{q}) + 2Pe\tilde{\mathbf{F}}_d(\mathbf{q})\cdot\tilde{\mathbf{F}}_{\mathbf{c}}(-\mathbf{q}) + Pe^2\tilde{\mathbf{F}}_d(\mathbf{q})\cdot\tilde{\mathbf{F}}_d(-\mathbf{q})\right) \tag{5.53}$$

$$\langle\tilde{\mathbf{F}}^2\rangle_0 = \frac{h^2}{V}\frac{\beta\gamma}{2}\int d\mathbf{r}\tilde{U}(\mathbf{r})\int\frac{d\mathbf{q}}{(2\pi)^d}e^{i\mathbf{q}\cdot\mathbf{r}}\langle e^{i\mathbf{q}\cdot\tilde{\eta}(0)\Delta t}\tilde{\eta}(0)\cdot\tilde{\mathbf{F}}(\mathbf{q})\rangle \tag{5.54}$$

$$= \frac{h^2}{V}\beta\gamma D_0\int\frac{d\mathbf{q}}{(2\pi)^d}\int\frac{d\mathbf{q}'}{(2\pi)^d}e^{i(\mathbf{q}+\mathbf{q}')\cdot\mathbf{r}}i\tilde{\mathbf{q}}\cdot\tilde{\mathbf{F}}(\mathbf{q})\tilde{U}(\mathbf{q}') \tag{5.55}$$

$$= \frac{h^2}{V}\int\frac{d\mathbf{q}}{(2\pi)^d}\tilde{\mathbf{F}}(\mathbf{q})\cdot\tilde{\mathbf{F}}_c(-\mathbf{q}) \tag{5.56}$$

$$= \frac{h^2}{V}\int\frac{d\mathbf{q}}{(2\pi)^d}\left(\tilde{\mathbf{F}}_c(\mathbf{q})\cdot\tilde{\mathbf{F}}_c(-\mathbf{q}) + Pe\tilde{\mathbf{F}}_d(\mathbf{q})\cdot\tilde{\mathbf{F}}_c(-\mathbf{q})\right) \tag{5.57}$$

Subtracting Eq. 5.57 from Eq. 5.53 yields the relationship for entropy dissipation:

$$\langle\tilde{\mathbf{F}}^2\rangle - \langle\tilde{\mathbf{F}}^2\rangle_0 = \frac{h^2}{V}\int\frac{d\mathbf{q}}{(2\pi)^d}\left(Pe\tilde{\mathbf{F}}_c(\mathbf{q})\cdot\tilde{\mathbf{F}}_d(-\mathbf{q}) + Pe^2\tilde{\mathbf{F}}_d(\mathbf{q})\cdot\tilde{\mathbf{F}}_d(-\mathbf{q})\right) \tag{5.58}$$

$$= \frac{h^2}{V}Pe^2\int\frac{d\mathbf{q}}{(2\pi)^d}\tilde{\mathbf{F}}_d(\mathbf{q})\cdot\tilde{\mathbf{F}}_d(-\mathbf{q}) \tag{5.59}$$

In the above we used the fact that $\int\frac{d\mathbf{q}}{(2\pi)^d}\tilde{\mathbf{F}}_c(\mathbf{q})\cdot\tilde{\mathbf{F}}_d(-\mathbf{q})$ can be rewritten as $\int d\mathbf{r}\tilde{\mathbf{F}}_c(\mathbf{r})\cdot\tilde{\mathbf{A}}(\mathbf{r}) = -\int d\mathbf{r}\tilde{\mathbf{A}}(\mathbf{r})\cdot\tilde{\mathbf{F}}_c(\mathbf{r}) = 0$. When $Pe = 0$, entropy production is zero and $\langle\tilde{\mathbf{F}}^2\rangle = \langle\tilde{\mathbf{F}}^2\rangle_0$.

We simulated a particle in a 2D periodic force landscape $F = (-h\cos(x) + hPe\cos(y),$ $-h\cos(y) - hPecos(x))$. We estimated $D_{eq}$ in such a system by running 5 simulations with $Pe = 0$ and averaging the resulting diffusion constants. Then we varied $Pe$ between 0.2 and 0.45, and computed the diffusion constant and the error in each case by averaging over

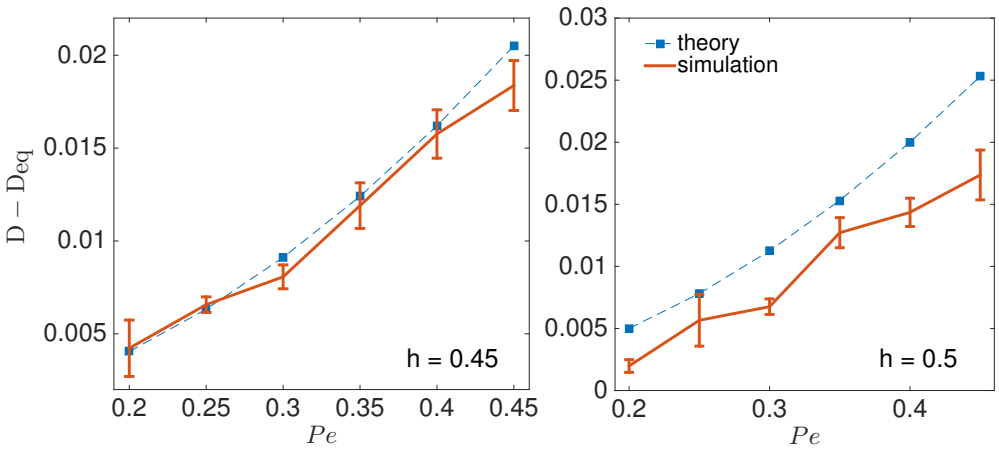



Figure A1.2: The correction to the diffusion constant due to the nonconservative force is plotted as a function of $Pe$ for a system characterized by a periodic force landscape. $D_{eq}$ is computed as the average diffusion constant for the system with $Pe = 0$. The data points are reasonably within error bars for $h = 0.45$, while there is some systematic deviation at $h = 0.5$.

5 simulations. In Fig. A1.2, we report how the deviation of the diffusion constant from $D_{eq}$ in our simulations compares to the expected correction due to $Pe$, which is $h^2Pe^2/2$. For this specific system, we set $\mathbf{F}_d \cdot \mathbf{F}_c = 0$ thus ensuring that $P_1 = 0$ (Eq. 5.30). Hence, although $h > Pe$ in these simulations, the correction to the diffusion constant arising from the nonequilibrium forces will scale like $Pe^2$ to leading order.

### *A1.4 Comparing our Diffusion Formula to other Analytical Expressions at Equilibrium*

At equilibrium, our prediction for the diffusion coefficient is given by:

$$D/D_0 = 1 - \frac{h^2}{(K_BT)^2 dV} \int U^2(\mathbf{r}) d\mathbf{r} \tag{5.60}$$

Here we compare our prediction to existing ones in the literature. In 1D, the analytical expression for the diffusion of a tagged particle in a periodic potential is due to [105] and reads: $D/D_0 = 1/\left(\langle e^{-\beta U}\rangle \langle e^{\beta U}\rangle\right)$, where the expectation value is taken to mean $\langle f(x)\rangle =$

$1/L \int_0^L f(x)dx$. In our perturbation theory, this expression becomes, to order $h^2$:

$$\frac{D}{D_0} = \frac{1}{\int_0^L e^{-\beta h U(x)} dx \int_0^L e^{\beta h U(x')} dx'} \tag{5.61}$$

$$= \frac{L}{L - h\beta \int_0^L U(x)dx + (h^2\beta^2/2) \int_0^L U^2(x)dx + \dots} \times \frac{L}{L + h\beta \int_0^L U(x')dx' + (h^2\beta^2/2) \int_0^L U^2(x')dx' + \dots} \tag{5.62}$$

$$\approx \frac{1}{1 + \frac{h^2\beta^2}{L} \int_0^L U^2(x)dx} \tag{5.63}$$

$$\approx 1 - \frac{h^2\beta^2}{L} \int_0^L U^2(x)dx \tag{5.64}$$

which is in agreement with our results. In 2D the corresponding formula $D/D_0 = 1/\langle e^{\beta U} \rangle$ is due to [35]. Expanding this expression to order $h^2$ and recognizing that $\int_0^L \int_0^L U(x,y)dxdy = 0$, we get:

$$\frac{D}{D_0} = \frac{L^2}{\int_0^L \int_0^L e^{\beta h U(x,y)} dxdy} \tag{5.65}$$

$$= \frac{L^2}{\int_0^L \int_0^L \left(1 + h\beta U(x,y) + (h^2\beta^2/2) U^2(x,y) + \dots\right) dxdy} \tag{5.66}$$

$$\approx \frac{1}{1 + \frac{h^2\beta^2}{2L^2} \int_0^L \int_0^L U^2(x,y)dxdy} \tag{5.67}$$

$$\approx 1 - \frac{h^2\beta^2}{2L^2} \int_0^L \int_0^L U^2(x,y)dxdy \tag{5.68}$$

This result is also in agreement with our prediction.

## *A1.5 Higher Order Terms in Expression for Diffusion*

We now show that higher order terms in $h$ do not diverge. Higher order terms of $\int_0^\infty \tilde{\mathbf{F}}(0) \cdot \tilde{\mathbf{F}}(t)dt$ can be expressed generally as:

$$H_{nm} \propto \frac{h^n}{(D_0)^{m-1}(2\pi)^{d(n-1)}} \int d\mathbf{q}_2 \ldots \int d\mathbf{q}_n \tilde{\mathbf{F}}\left(-\sum_{i=2}^{n} \mathbf{q}_i\right) \cdot \tilde{\mathbf{F}}(\mathbf{q}_2) \times \frac{i\mathbf{q}_2 \cdot \tilde{\mathbf{F}}(\mathbf{q}_3) \ldots i\mathbf{q}_{m-1} \cdot \tilde{\mathbf{F}}(\mathbf{q}_m)\tilde{U}(\mathbf{q}_{m+1}) \ldots \tilde{U}(\mathbf{q}_n)}{|\mathbf{q}_2|^2(|\mathbf{q}_2|^2 + |\mathbf{q}_3|^2) \ldots (|\mathbf{q}_2|^2 + |\mathbf{q}_3|^2 + \ldots + |\mathbf{q}_m|^2)} \tag{5.69}$$

where $n \geq 3$ and $n \geq m \geq 2$. This quantity scales with system size as $L^{(1-n)d+(2-m)+2(m-1)}$. When $m = n$, we get a maximum scaling with system size as $L^{(1-n)d+n}$. For $d = 1$, higher order terms in $n$ can diverge with system size. For $d = 2$, the scaling is $L^{2-n}$, which can diverge only for terms of order less than or equal to 2. (Therefore, as long as $\langle F \rangle = 0$ holds, then the quadratic terms are well behaved as described in Appendix A1.3). For d=3, the scaling is $L^{3-2n}$, which can not diverge since $n \geq 2$.

The higher order terms in $\int_0^\infty \eta(0) \cdot \tilde{\mathbf{F}}(t)dt$ can also be conveniently expressed as:

$$H_{nm} \propto \frac{2h^n}{(D_0)^{m-1}(2\pi)^{d(n-1)}} \int d\mathbf{q}_1 \ldots \int d\mathbf{q}_{n-1} i\mathbf{q}_1 \cdot \tilde{\mathbf{F}}(\mathbf{q}_1) \times \frac{i\mathbf{q}_1 \cdot \tilde{\mathbf{F}}(\mathbf{q}_2) \ldots i\mathbf{q}_{m-1} \cdot \tilde{\mathbf{F}}(\mathbf{q}_m)\tilde{U}(\mathbf{q}_{m+1}) \ldots \tilde{U}\left(-\sum_{i=1}^{n-1} \mathbf{q}_i\right)}{|\mathbf{q}_1|^2\left(|\mathbf{q}_1|^2 + |\mathbf{q}_2|^2\right) \ldots (|\mathbf{q}_1|^2 + |\mathbf{q}_2|^2 + \ldots + |\mathbf{q}_m|^2)} \tag{5.70}$$

where $n \geq 3$ and $n - 1 \geq m \geq 1$. Terms with $m = n$ in this case are zero.

This quantity scales with system size as $L^{(1-n)d+(-m)+2m}$. When $m = n$, we obtain the maximum scaling with system size but these terms are null so the higher order terms cannot diverge.

### *A1.6 Composition-dependent Diffusion Can Lead to Phase Separation*

We will start by showing that a simple coarse-graining of our many body system with 1/2 driven and 1/2 undriven particles can lead to phase separation. Namely, let us track a probe volume $V_0$ and define a parameter $f = \frac{N_R - N_B}{N_0}$, where $N_R$ and $N_B$ are the numbers of red and blue particles, respectively, and $N_0 = \rho V_0$ is the average number of particles in the probe volume. We can coarse-grain the system by positing that the particles in this probe volume do not interact with each other in any way, but solely diffuse with a diffusion constant $D(f)$ which depends on the instantaneous composition f. Since the only timescale relevant to the system is related to this diffusion, we can expect that, for small deviations around $f = 0$, the composition f obeys an equation of motion of the form below:

$$\frac{df}{dt} = -\alpha D(f) f + \sqrt{\beta D(f)}\eta \tag{5.71}$$

where $\alpha$ and $\beta$ are phenomenological constants, $\eta$ is white noise with unit variance and the expression is interpreted in the Itó sens. For now we will demonstrate how such an equation of motion can lead to instabilities, and further ahead in this section we will carefully provide a framework for predicting the form of the parameters $\alpha$ and $\beta$.

By ignoring the linear part of the diffusion coefficient in the many-body system as a function of $Pe$, we can write $D(f) \equiv D_{eq} + (1 - f^2) b(\tau) Pe^2$, to second order in the parameter $f$. This form satisfies the constraints that the diffusion increases quadratically with the driving forces, and decreases as we shift away from a mixed composition to a demixed composition of either just red or just blue particles. Plugging this form of $D(f)$ in Eq. 5.71 and solving the Fokker-Planck equation associated with such an equation of motion yields the solution:

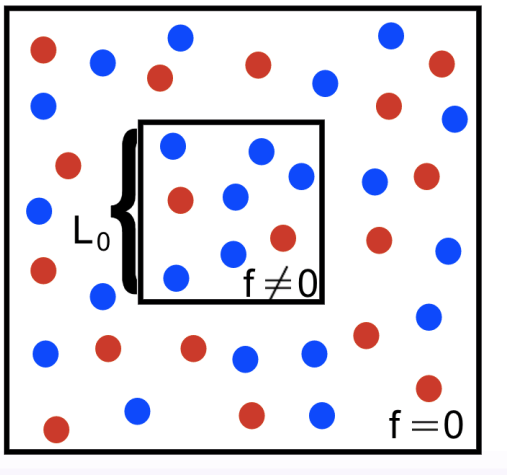



Figure A1.3: The setup of our composition calculations involves a small square probe volume of dimension $L_0$ with demixed composition ($f \neq 0$) in a mixed particle bath (f=0).

$$P(f) = C_0 \frac{e^{-\frac{\alpha}{2\beta}f^2}}{D_{eq} + (1-f^2)b(\tau)Pe^2} \tag{5.72}$$

where $C_0$ is the normalization constant.

Taylor expanding this expression, we obtain that $P(f)$ stops being peaked at $f = 0$ (and hence instabilities appear) whenever $\frac{\alpha}{2\beta} < \frac{b(\tau)Pe^2}{D_{eq} + b(\tau)Pe^2}$. This framework allows us to predict the critical $Pe^*$ at which phase separation occurs in terms of the constants $\alpha$, $\beta$, $D_{eq}$ and $b(\tau)$. In the next paragraphs we will derive formulas for the parameters $\alpha$ and $\beta$, and show that the condition we find is attainable in our simulations, to further confirm our theory is reasonable.

To recall, we will attempt to formally derive an equation of motion for the variable $f = \frac{N_R - N_B}{N_0}$. We will compute the statistics of the variable $\Delta f$, the change of the composition in a small amount of time, given that the composition is $f$ at $t = 0$ and the number of particles in the probe volume is fixed at $N_0 = \rho L_0^2$. We imagine the probe volume exchanges particles with a bath that has a mixed composition ($f = 0$) (Fig. A1.3).

We first compute $\Delta f$ by separating it into two parts, a part due to particles diffusing out

of the box, $\Delta f_{\text{out}}$ and a part from bath particles diffusing in, $\Delta f_{\text{in}}$.

We obtain:

$$\langle \Delta f_{\text{out}} \rangle = \frac{\langle -\Delta N_{R,\text{out}} + \Delta N_{B,\text{out}} \rangle}{N_0} \tag{5.73}$$

$$= \frac{\left\langle -\Delta N_{\text{out}} \frac{1+f}{2} + \Delta N_{\text{out}} \frac{1-f}{2} \right\rangle}{\langle N \rangle} \tag{5.74}$$

$$= -f \frac{\langle \Delta N_{\text{out}} \rangle}{N_0} \tag{5.75}$$

and

$$\langle \Delta f_{\text{in}} \rangle = \frac{\langle \Delta N_{R,\text{in}} - \Delta N_{B,\text{in}} \rangle}{N_0} \tag{5.76}$$

$$= \frac{\left\langle \Delta N_{\text{in}} \frac{1}{2} - \Delta N_{\text{in}} \frac{1}{2} \right\rangle}{N_0} \tag{5.77}$$

$$= 0 \tag{5.78}$$

Putting these two together, we obtain:

$$\langle \Delta f \rangle = -f \frac{\langle \Delta N_{\text{out}} \rangle}{N_0} \tag{5.79}$$

Similarly, we can compute the variance of $\Delta f$ as:

$$\langle \Delta f^2 \rangle = \frac{\langle \Delta N_{\text{out}}^2 \rangle}{N_0^2} + \frac{\langle \Delta N_{\text{in}}^2 \rangle}{N_0^2} \tag{5.80}$$

In the derivations of $\langle \Delta f \rangle$ and $\langle \Delta f^2 \rangle$ we assumed that the time step is small enough that essentially only on the order of unit changes occur in the number of particles. In this limit the probability that the particles leaving a box with composition $f$ are all red/blue is $\frac{1 \pm f}{2}$.

To compute the quantity $\langle \Delta N_{\text{out}} \rangle$ we imagine a box surrounded by vacuum. At time t=0, the probability for a particle to be located inside the box is uniform, and is equal to $P(\mathbf{x}, 0) = \frac{1}{L_0^2}$. We imagine at the boundaries the probability decays to zero over a distance $\sigma$ on the order of particle size. The particles can diffuse out of the box with a diffusion coefficient equal to $D$, the effective diffusion of particles in the many body liquid. After a timestep $\Delta t$, the change in the number of particles can be computed as:

$$\langle \Delta N_{\text{out}} \rangle = N_0 \int_{\text{vol}} \frac{dP(\mathbf{x}, 0)}{dt} dV \Delta t \tag{5.81}$$

$$= N_0 \int_{\text{vol}} \nabla \cdot D \nabla P(\mathbf{x}, 0) dV \Delta t \tag{5.82}$$

$$= N_0 \int_{\text{bdd}} D \nabla P(\mathbf{x}, 0) \cdot \vec{n} dS \Delta t \tag{5.83}$$

$$= N_0 D \Delta t \frac{1}{L_0^2 \sigma} 4 L_0 \tag{5.84}$$

$$= 4 \sqrt{N_0 \rho} D \Delta t \frac{1}{\sigma} \tag{5.85}$$

To compute the variance $\langle \Delta N_{\text{out}}^2 \rangle$, we consider flux of particles in and out of the probe volume and obtain:

$$\langle \Delta N_{\text{out}}^2 \rangle = 2 \rho \sqrt{N_0} D \Delta t \tag{5.86}$$

and

$$\langle \Delta N_{\text{in}}^2 \rangle = 2 \rho \sqrt{N_0} D (f = 0) \Delta t \tag{5.87}$$

Combining Eqs. 5.79, 5.80, 5.85, 5.86 and 5.87, and allowing the diffusion of the liquid to change as a function of composition, we write the final equation of motion for the composition

as:

$$\frac{df}{dt} = -4\sqrt{\frac{\rho}{N_0}}\frac{fD(f)}{\sigma} + \frac{\sqrt{2\rho(D(f)+D(0))}}{\sqrt{N_0^3}}\eta \tag{5.88}$$

where $\eta$ is a Gaussian noise of zero mean and unit variance.

The Fokker-Planck equation resulting from Eq. 5.88, which we interpret in the Itô sense, is:

$$\frac{dP}{dt} = \frac{4}{\sigma}\sqrt{\frac{\rho}{N_0}}\frac{d(D(f)fP)}{df} + \frac{2\rho}{\sqrt{N_0^3}}\frac{d^2((D(f)+D(0))P)}{df^2} \tag{5.89}$$

In order to derive the condition for phase separation, we will ignore the linear part of the diffusion coefficient as a function of $Pe$, and write $D(f) \equiv D_{eq} + (1-f^2)b(\tau)Pe^2$. Setting the flux associated with Eq. 5.89 to zero, we obtain the solution:

$$P(f) = C_0 e^{-\left(\frac{N_0}{\sqrt{\rho}\sigma}f^2\right)}\left(2D(0) - f^2 b(\tau)Pe^2\right)^{-\beta} \tag{5.90}$$

$$= \frac{C_0}{(2D(0))^\beta}\left(1 + \left(\frac{\beta b(\tau)Pe^2}{2D(0)} - \frac{N_0}{\sqrt{\rho}\sigma}\right)f^2 + \mathcal{O}(f^4)\right) \tag{5.91}$$

where $C_0$ is the normalization, $\beta = \dfrac{D(0)\frac{N_0}{\sqrt{\rho}\sigma}}{b(\tau)Pe^2} + 1$, and in going from the first line to the second line we employed a Taylor expansion.

Let's call $\dfrac{N_0}{\sqrt{\rho}\sigma} = \lambda$, the constant mentioned in Eq. 2.21. When $\lambda > \dfrac{b(\tau)Pe^2}{D_{eq} + b(\tau)Pe^2}$, the steady state solution is peaked at $f = 0$, meaning the system will be mixed. However, when $\lambda < \dfrac{b(\tau)Pe^2}{D_{eq} + b(\tau)Pe^2}$, the steady state solution has a minimum at $f = 0$. Thus, our analysis shows that a composition dependent diffusion can lead to phase separation.

In our simulations, we obtained that the condition for phase separation is roughly $\dfrac{b(\tau)Pe^2}{D_{eq} + b(\tau)Pe^2} > 0.5$ . In other words, $Pe^2 b(\tau) > D_{eq}$ for a phase transition.

By comparison, since $\sigma = 1$ and $\rho = 0.45$, by setting $L_0 = \sigma$ our predicted condition

for a phase transition is $\frac{b(\tau)Pe^2}{D_{eq}+b(\tau)Pe^2} > 0.67$. In other words, our simple estimates for the phenomenological constants in the composition fluctuation equation predict a phase transition when $Pe^2 b(\tau) > 2D_{eq}$. Given how simplistic our phenomenological composition fluctuation equation is, this prediction is reasonable. For a choice of $L_0 = 0.9\sigma$, the prediction from our theory matches numerical results.

## A2 Appendix for Chapter 3

### *A2.1 Dissipation and Diffusion*

This Appendix is devoted to the derivation of the dissipation rate $\mathcal{J}$ and the diffusion coefficient $D$ of a driven tracer, as defined in Sec. 3.2. We employ a perturbative treatment at weak interactions, originally introduced for a particle driven at constant force in [40, 39]. To this aim, the tracer-bath interaction potential $v$ is scaled by a small dimensionless parameter $h \ll 1$ in what follows. Besides, we focus on the regime of dilute tracers, so that interactions among them, either direct or mediated by the bath, can be safely neglected.

The dynamic action associated with the tracer dynamics (Eq. 3.11-3.12) follows from standard path integral methods [115, 34]. It can be separated into contributions from the free tracer motion and from interactions, respectively denoted by $\mathcal{A}_0$ and $\mathcal{A}_{\rm int}$:

$$\begin{aligned}
\mathcal{A}_0 &= \int \bar{\mathbf{r}}_0 \cdot \left[\mathrm{i}(\dot{\mathbf{r}}_0 - \mathbf{F}_{\rm d}/\gamma) + D_0 \bar{\mathbf{r}}_0\right] \mathrm{d}t, \\
\mathcal{A}_{\rm int} &= \frac{h^2}{\gamma} \int \frac{\mathrm{d}\mathbf{q}}{(2\pi)^d} |\mathbf{q}|^2 |v(\mathbf{q})|^2 \int_{-\infty}^{\infty} \mathrm{d}s \int_{-\infty}^{s} \mathrm{d}u \\
&\quad \times \mathrm{e}^{-D_{\rm G}|\mathbf{q}|^2 K(\mathbf{q})(s-u) + \mathrm{i}\mathbf{q}\cdot[\mathbf{r}_0(s) - \mathbf{r}_0(u)]} \\
&\quad \times \bar{\mathbf{r}}_0(s) \cdot \left[\frac{\bar{\mathbf{r}}_0(u)}{\gamma K(\mathbf{q})} - \frac{\mathbf{q}}{\gamma_{\rm G}}\right],
\end{aligned} \tag{5.92}$$

where $D_0 = T/\gamma$ is the tracer diffusion coefficient in the absence of interactions ($v = 0$), and $\bar{\mathbf{r}}_0$ is the process conjugated with the tracer position $\mathbf{r}_0$. For weak interactions $h \ll 1$,

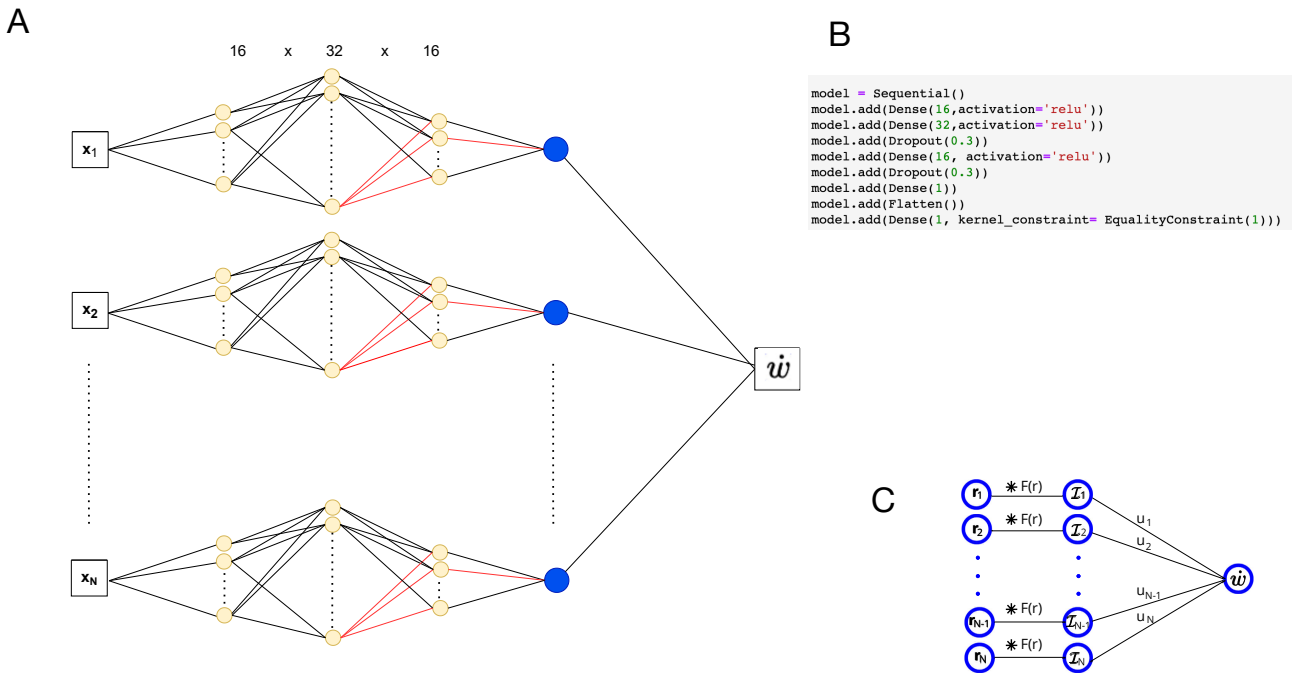



Figure A2.1: A) The architecture used in the deep neural network to predict rate of work from the center of mass data of the rotors. The first three fully-connected layers are of size 32, 64 and 16, respectively, and the weights and biases learned are the same for all input vectors $[r_{i,1}, r_{i,2}, \ldots, r_{i,D}, \theta_{i,1}, \theta_{i,2}, \ldots, \theta_{i,D}]$, where $D = 40$ is the number of neighbors included for each rotor. After each layer, except for the first two layers, we apply a dropout layer that inactivates 30% of the neurons, which is graphically shown as red links between neurons. These links are chosen randomly each epoch and their weights are set to zero during training. The fourth layer is of size 1, leading to the blue neurons which are flattened into a single vector and fully connected to the output, which is the rate of work. The weights of the final layer are set to be equal. B) Snippet of code to construct the network in the popular Python package keras. The kernel constraint "EqualityConstraint" was built separately and is available with the rest of the code in Ref. [179]. C) The convolutional architecture used in the main text (same figure as Fig. 4.6 in the main text)

any average value can be then expanded in terms of $h$ as $\langle\cdot\rangle = \langle\cdot\rangle_0 - h^2\langle\mathcal{A}_{\rm int}\cdot\rangle_0 + \mathcal{O}(h^4)$, where $\langle\cdot\rangle_0$ is the average taken with respect to $\mathcal{A}_0$ only. As a result, determining the first correction from interactions in any observable amounts to computing the corresponding average $\langle\mathcal{A}_{\rm int}\cdot\rangle_0$.

Considering the dissipation rate per particle $\mathcal{J}/N = \langle\dot{\mathbf{r}}_0\rangle\cdot\mathbf{F}_{\rm d}$, the leading order is $\langle\dot{\mathbf{r}}_0\rangle_0\cdot\mathbf{F}_{\rm d} = |\mathbf{F}_{\rm d}|^2/\gamma = f^2/\gamma$, and the first correction reads $-h^2\langle\mathcal{A}_{\rm int}\dot{\mathbf{r}}_0\rangle_0\cdot\mathbf{F}_{\rm d}$. Given the explicit form of $\mathcal{A}_{\rm int}$ in Eq. 5.92, the correlations of interest are

$$\begin{aligned}
&\Big\langle\dot{\mathbf{r}}_0(t)\big[\mathbf{q}\cdot\bar{\mathbf{r}}_0(s)\big]\mathrm{e}^{\mathrm{i}\mathbf{q}\cdot[\mathbf{r}_0(s)-\mathbf{r}_0(u)]}\Big\rangle_0\\
&\qquad = \mathrm{i}\mathbf{q}\delta(t-s)\mathrm{e}^{-D_0|\mathbf{q}|^2(t-u)+\frac{\mathrm{i}\mathbf{q}}{\gamma}\cdot\int_u^t\mathbf{F}_{\rm d}(w)\mathrm{d}w},\\
&\Big\langle\dot{\mathbf{r}}_0(t)\big[\bar{\mathbf{r}}_0(u)\cdot\bar{\mathbf{r}}_0(s)\big]\mathrm{e}^{\mathrm{i}\mathbf{q}\cdot[\mathbf{r}_0(s)-\mathbf{r}_0(u)]}\Big\rangle_0\\
&\qquad = -\mathrm{i}\mathbf{q}\delta(t-s)\mathrm{e}^{-D_0|\mathbf{q}|^2(t-u)+\frac{\mathrm{i}\mathbf{q}}{\gamma}\cdot\int_u^t\mathbf{F}_{\rm d}(w)\mathrm{d}w},
\end{aligned} \tag{5.93}$$

where we have used that the tracer statistics is Gaussian in the absence of interactions, following [40, 39]. From this result, we get

$$\begin{aligned}
&\mathcal{J} - f^2/\gamma\\
&= \frac{Nh^2}{d\gamma^2}\int\frac{\mathrm{d}\mathbf{q}}{(2\pi)^d}\mathrm{i}\mathbf{q}\cdot\mathbf{F}_{\rm d}(t)|\mathbf{q}|^2|v(\mathbf{q})|^2\frac{D_0+D_{\rm G}K(\mathbf{q})}{D_0K(\mathbf{q})}\\
&\quad\times\int_{-\infty}^t\mathrm{d}u\mathrm{e}^{-|\mathbf{q}|^2[D_0+D_{\rm G}K(\mathbf{q})](t-u)+\frac{\mathrm{i}\mathbf{q}}{\gamma}\cdot\int_u^t\mathbf{F}_{\rm d}(w)\mathrm{d}w}\\
&\quad+\mathcal{O}(h^4),
\end{aligned} \tag{5.94}$$

where we have used $\gamma_{\rm G} = \gamma/\rho_0$ and $D_{\rm G} = \rho_0 D_0$. Expanding at small $f$, we deduce

$$\begin{aligned}
&\mathcal{J} - f^2/\gamma \\
&= -\frac{Nh^2}{d\gamma^3} \int \frac{\mathrm{d}\mathbf{q}}{(2\pi)^d} |\mathbf{q}|^4 |v(\mathbf{q})|^2 \frac{D_0 + D_{\rm G} K(\mathbf{q})}{D_0 K(\mathbf{q})} \\
&\quad \times \int_{-\infty}^{t} \mathrm{d}u \mathrm{e}^{-|\mathbf{q}|^2 [D_0 + D_{\rm G} K(\mathbf{q})](t-u)} \int_{u}^{t} \mathrm{d}w \mathbf{F}_{\rm d}(t) \cdot \mathbf{F}_{\rm d}(w) \\
&\quad + \mathcal{O}(h^4, f^4).
\end{aligned} \tag{5.95}$$

Substituting the explicit expression of the deterministic drive in Eq. 3.2 into Eq. 5.95, and then integrating over $u$ and $w$, we obtain

$$\begin{aligned}
&\mathcal{J} - f^2/\gamma \\
&= -\frac{N(hf)^2}{d\gamma^3} \int \frac{\mathrm{d}\mathbf{q}}{(2\pi)^d} \frac{|\mathbf{q}|^4 |v(\mathbf{q})|^2}{|\mathbf{q}|^4 \big[D_0 + D_{\rm G} K(\mathbf{q})\big]^2 + \omega^2} \\
&\quad \times \frac{D_0 + D_{\rm G} K(\mathbf{q})}{D_0 K(\mathbf{q})} + \mathcal{O}(h^4, f^4).
\end{aligned} \tag{5.96}$$

For the case of active drive with correlations (Eq. 3.3), we exploit the equivalence with a disordered drive detailed in Sec. 3.2. Substituting the explicit drive (Eq. 3.4) in Eq. 5.95, and then averaging over disorder in the limit of many oscillators ($n \gg 1$), we get

$$\begin{aligned}
&\mathcal{J} - f^2/\gamma \\
&= -\frac{N(hf)^2}{d\gamma^3} \int \frac{\mathrm{d}\mathbf{q}\mathrm{d}\omega'}{(2\pi)^{d+1}} \frac{|\mathbf{q}|^4 |v(\mathbf{q})|^2 \phi(\omega')}{|\mathbf{q}|^4 \big[D_0 + D_{\rm G} K(\mathbf{q})\big]^2 + (\omega')^2} \\
&\quad \times \frac{D_0 + D_{\rm G} K(\mathbf{q})}{D_0 K(\mathbf{q})} + \mathcal{O}(h^4, f^4),
\end{aligned} \tag{5.97}$$

where $\phi(\omega') = 2\tau/\big[1+(\omega'\tau)^2\big]$, yielding

$$\begin{aligned}
&\mathcal{J} - f^2/\gamma \\
&= -\frac{N\tau(hf)^2}{d\gamma^3}\int \frac{\mathrm{d}\mathbf{q}}{(2\pi)^d}\frac{|\mathbf{q}|^2|v(\mathbf{q})|^2}{D_0K(\mathbf{q})} \\
&\quad\times \frac{1}{\tau|\mathbf{q}|^2\big[D_0+D_\mathrm{G}K(\mathbf{q})\big]+1} + \mathcal{O}(h^4,f^4).
\end{aligned} \tag{5.98}$$

The asymptotic results for the rate of work $\dot{w} = f^2/\gamma - \mathcal{J}$, presented in Sec. 3.4 for both deterministic and active drives, follow directly.

We now turn to deriving the diffusion coefficient $D$. It is defined in terms of the mean-squared displacement (MSD) $\langle\Delta\mathbf{r}_0^2(t)\rangle = \big\langle\big[\langle\mathbf{r}_0(t)\rangle - \mathbf{r}_0(t)\big]^2\big\rangle$ as $D = \lim\limits_{t\to\infty}\langle\Delta\mathbf{r}_0^2(t)\rangle/2dt$. At leading order, the MSD reads $\langle\Delta\mathbf{r}_0^2(t)\rangle_0 = 2dD_0t$. To obtain the first order, we need to compute the following correlations

$$\begin{aligned}
&\Big\langle\Delta\mathbf{r}_0^2(t)\big[\mathbf{q}\cdot\bar{\mathbf{r}}_0(s)\big]\mathrm{e}^{\mathrm{i}\mathbf{q}\cdot[\mathbf{r}_0(s)-\mathbf{r}_0(u)]}\Big\rangle_0 \\
&\qquad= -4(D_0/\gamma)|\mathbf{q}|^2(s-u)\Theta(t-s) \\
&\qquad\quad\times \mathrm{e}^{-D_0|\mathbf{q}|^2(t-u)+\frac{\mathrm{i}\mathbf{q}}{\gamma}\cdot\int_u^t\mathbf{F}_\mathrm{d}(w)\mathrm{d}w}, \\
&\Big\langle\Delta\mathbf{r}_0^2(t)\big[\bar{\mathbf{r}}_0(u)\cdot\bar{\mathbf{r}}_0(s)\big]\mathrm{e}^{\mathrm{i}\mathbf{q}\cdot[\mathbf{r}_0(s)-\mathbf{r}_0(u)]}\Big\rangle_0 \\
&\qquad= (2/\gamma^2)\Theta(t-s)\big[2D_0|\mathbf{q}|^2(s-u)-1\big] \\
&\qquad\quad\times \mathrm{e}^{-D_0|\mathbf{q}|^2(t-u)+\frac{\mathrm{i}\mathbf{q}}{\gamma}\cdot\int_u^t\mathbf{F}_\mathrm{d}(w)\mathrm{d}w},
\end{aligned} \tag{5.99}$$

where we have used again that $\mathcal{A}_0$ is Gaussian in terms of $\bar{\mathbf{r}}_0$, yielding

$$\begin{aligned}
&\left\langle \Delta\mathbf{r}_0^2(t)\right\rangle - 2dD_0t \\
&= \frac{2h^2}{\gamma^2}\int \frac{\mathrm{d}\mathbf{q}}{(2\pi)^d}\frac{|\mathbf{q}|^2|v(\mathbf{q})|^2}{K(\mathbf{q})} \\
&\quad\times \int_{-\infty}^{t}\mathrm{d}s\int_{-\infty}^{s}\mathrm{d}u\big\{2|\mathbf{q}|^2\big[D_0+D_{\mathrm{G}}K(\mathbf{q})\big](s-u)-1\big\} \\
&\quad\times \mathrm{e}^{-|\mathbf{q}|^2[D_0+D_{\mathrm{G}}K(\mathbf{q})](s-u)+\frac{\mathrm{i}\mathbf{q}}{\gamma}\cdot\int_u^s\mathbf{F}_{\mathrm{d}}(w)\mathrm{d}w}+\mathcal{O}(h^4).
\end{aligned} \tag{5.100}$$

Expanding at small $f$, we get

$$\begin{aligned}
&\left\langle \Delta\mathbf{r}_0^2(t)\right\rangle - 2dD_{\mathrm{eq}}t \\
&= -\frac{2h^2}{\gamma^4}\int \frac{\mathrm{d}\mathbf{q}}{(2\pi)^d}\frac{|\mathbf{q}|^4|v(\mathbf{q})|^2}{K(\mathbf{q})} \\
&\quad\times \int_{-\infty}^{t}\mathrm{d}s\int_{-\infty}^{s}\mathrm{d}u\big\{2|\mathbf{q}|^2\big[D_0+D_{\mathrm{G}}K(\mathbf{q})\big](s-u)-1\big\} \\
&\quad\times \mathrm{e}^{-|\mathbf{q}|^2[D_0+D_{\mathrm{G}}K(\mathbf{q})](s-u)}\int_u^s\mathrm{d}w_1\mathrm{d}w_2\mathbf{F}_{\mathrm{d}}(w_1)\cdot\mathbf{F}_{\mathrm{d}}(w_2) \\
&\quad+\mathcal{O}(h^4,f^4),
\end{aligned} \tag{5.101}$$

where $D_{\mathrm{eq}}$ refers to the diffusion coefficient in the absence of driving force ($f=0$). For the deterministic drive (Eq. 3.2), the explicit time integrations give

$$\begin{aligned}
D-D_{\mathrm{eq}} = \frac{(hf)^2}{d\gamma^4}\int \frac{\mathrm{d}\mathbf{q}}{(2\pi)^d}&\frac{|\mathbf{q}|^2|v(\mathbf{q})|^2}{K(\mathbf{q})\big[D_0+D_{\mathrm{G}}K(\mathbf{q})\big]} \\
&\times\frac{5|\mathbf{q}|^4\big[D_0+D_{\mathrm{G}}K(\mathbf{q})\big]^2+\omega^2}{\big\{|\mathbf{q}|^4\big[D_0+D_{\mathrm{G}}K(\mathbf{q})\big]^2+\omega^2\big\}^2} \\
&+\mathcal{O}(h^4,f^4).
\end{aligned} \tag{5.102}$$

Using the mapping in Sec. 3.2 for the case of active drive with correlations (Eq. 3.3), we

deduce

$$
\begin{aligned}
D - D_{\mathrm{eq}} = \frac{(hf)^2}{d\gamma^4} \int \frac{\mathrm{d}\mathbf{q}\mathrm{d}\omega'}{(2\pi)^{d+1}} \frac{|\mathbf{q}|^2 |v(\mathbf{q})|^2 \phi(\omega')}{K(\mathbf{q})\big[D_0 + D_{\mathrm{G}} K(\mathbf{q})\big]} \\
\times \frac{5|\mathbf{q}|^4 \big[D_0 + D_{\mathrm{G}} K(\mathbf{q})\big]^2 + (\omega')^2}{\big\{|\mathbf{q}|^4 \big[D_0 + D_{\mathrm{G}} K(\mathbf{q})\big]^2 + (\omega')^2\big\}^2} \\
+ \mathcal{O}(h^4, f^4),
\end{aligned}
\tag{5.103}
$$

where again $\phi(\omega') = 2\tau/\big[1 + (\omega'\tau)^2\big]$, yielding

$$
\begin{aligned}
D - D_{\mathrm{eq}} = \frac{\tau(hf)^2}{d\gamma^4} \int \frac{\mathrm{d}\mathbf{q}}{(2\pi)^d} \frac{|v(\mathbf{q})|^2}{K(\mathbf{q})\big[D_0 + D_{\mathrm{G}} K(\mathbf{q})\big]^2} \\
\times \frac{5\tau|\mathbf{q}|^2 \big[D_0 + D_{\mathrm{G}} K(\mathbf{q})\big] + 3}{\big\{\tau|\mathbf{q}|^2 \big[D_0 + D_{\mathrm{G}} K(\mathbf{q})\big] + 1\big\}^2} \\
+ \mathcal{O}(h^4, f^4).
\end{aligned}
\tag{5.104}
$$

Finally, we obtain the expressions in the asymptotic regimes, as reported in Sec. 3.4 for both deterministic and active drives.

## *A2.2 Equivalence of Biased Ensembles*

In this Appendix, we demonstrate the equivalence between specific dynamical biased ensembles. First, we consider ensembles related to the equilibrium dynamics (Eq. 3.28). Ensemble (a) corresponds to biasing with the factor $\exp\big[\sum_{i,j} \kappa_{ij} \int_0^t \mathcal{E}_{ij}(s)\mathrm{d}s\big]$ in the path probability, where

$$
\mathcal{E}_{ij} = \frac{1}{\gamma T} \sum_k \big[T\nabla_k - \nabla_k V\big] \cdot \nabla_k A(\mathbf{r}_i - \mathbf{r}_j). \tag{5.105}
$$

Ensemble (b) is associated with the first-order auxiliary dynamics, whose potential reads $V + 2\sum_{i,j} \kappa_{ij} A(\mathbf{r}_i - \mathbf{r}_j)$, biased with $\exp\big[\int_0^t \varepsilon'(s)\mathrm{d}s\big]$ where

$$
\varepsilon' = \frac{1}{\gamma T} \sum_k \Big[\sum_{i,j} \kappa_{ij} \nabla_k A(\mathbf{r}_i(s) - \mathbf{r}_j(s))\Big]^2. \tag{5.106}
$$

Obtaining the equivalence between (a) and (b) amounts to showing that their path probabilities are similar. The corresponding dynamic actions, denoted by $\mathcal{A}^{(\sigma)}(t) = \sum_k \int_0^t \mathbb{A}_k^{(\sigma)}(s)\mathrm{d}s$ for $\sigma = \{\mathrm{a}, \mathrm{b}\}$, are given by

$$\begin{aligned}\mathbb{A}_k^{(\mathrm{a})} = &\frac{1}{4\gamma T}\big[\gamma\dot{\mathbf{r}}_k + \nabla_k V\big]^2 - \frac{1}{2\gamma}\nabla_k^2 V \\ &- \frac{1}{\gamma T}\sum_{i,j}\kappa_{ij}\big[T\nabla_k - \nabla_k V\big]\cdot\nabla_k A(\mathbf{r}_i - \mathbf{r}_j),\end{aligned} \tag{5.107}$$

and

$$\begin{aligned}\mathbb{A}_k^{(\mathrm{b})} = &\frac{1}{4\gamma T}\Big[\gamma\dot{\mathbf{r}}_k + \nabla_k V + 2\sum_{i,j}\kappa_{ij}\nabla_k A(\mathbf{r}_i - \mathbf{r}_j)\Big]^2 \\ &- \frac{1}{2\gamma}\nabla_k^2\Big[V + 2\sum_{i,j}\kappa_{ij}A(\mathbf{r}_i - \mathbf{r}_j)\Big] \\ &- \frac{1}{\gamma T}\Big[\sum_{i,j}\kappa_{ij}\nabla_k A(\mathbf{r}_i - \mathbf{r}_j)\Big]^2.\end{aligned} \tag{5.108}$$

Expanding $\mathbb{A}_k^{(\mathrm{b})}$ in Eq. 5.108 and comparing with $\mathbb{A}_k^{(\mathrm{b})}$ in Eq. 5.107, it appears that $\mathcal{A}^{(\mathrm{a})}$ and $\mathcal{A}^{(\mathrm{b})}$ are indeed equal up to a boundary term proportional to $\sum_{i,j}\kappa_{ij}\big[A(\mathbf{r}_i(t) - \mathbf{r}_j(t)) - A(\mathbf{r}_i(0) - \mathbf{r}_j(0))\big]$ which can be neglected at large $t$: this establishes the equivalence between ensembles (a) and (b).

We now turn to demonstrate the equivalence between two ensembles related to the Vicsek-like dynamics (Eq. 3.39). Ensemble (c) is biased with the factor $\exp\big[\kappa\int_0^t \mathcal{E}_\theta(s)\mathrm{d}s\big]$, where $\mathcal{E}_\theta$ is defined in Eq. 3.40. Ensemble (d) corresponds to the first-order auxiliary dynamics, with torque given by $(1+\kappa)\mathcal{T}$, biased with $\exp\big[\int_0^t \varepsilon_\theta(s)\mathrm{d}s\big]$ where

$$\varepsilon_\theta = \frac{(\kappa\mu_\mathrm{r})^2}{4D_\mathrm{r}}\sum_{i,j,k}\mathcal{T}(\theta_i - \theta_k, \mathbf{r}_i - \mathbf{r}_k)\mathcal{T}(\theta_i - \theta_j, \mathbf{r}_i - \mathbf{r}_j). \tag{5.109}$$

The dynamic actions for each ensemble, denoted by $\mathcal{B}^{(\sigma)}(t) = \sum_i \int_0^t \mathbb{B}_i^{(\sigma)}(s)\mathrm{d}s$ for $\sigma = \{\mathrm{c}, \mathrm{d}\}$,

are given by

$$\mathbb{B}_i^{(\mathrm{c})} = \frac{1}{4D_\mathrm{r}}\Big[\dot{\theta}_i - \mu_\mathrm{r}\sum_j \mathcal{T}(\theta_i-\theta_j,\mathbf{r}_i-\mathbf{r}_j)\Big]^2 + \frac{\mu_\mathrm{r}}{2}\frac{\partial}{\partial\theta_i}\sum_j \mathcal{T}(\theta_i-\theta_j,\mathbf{r}_i-\mathbf{r}_j) + \frac{\kappa\mu_\mathrm{r}}{2}\sum_j\left[\frac{\partial}{\partial\theta_i}+\frac{\mu_\mathrm{r}}{D_\mathrm{r}}\sum_k \mathcal{T}(\theta_i-\theta_k,\mathbf{r}_i-\mathbf{r}_k)\right]\mathcal{T}(\theta_i-\theta_j,\mathbf{r}_i-\mathbf{r}_j), \tag{5.110}$$

and

$$\mathbb{B}_i^{(\mathrm{d})} = \frac{1}{4D_\mathrm{r}}\Big[\dot{\theta}_i - (1+\kappa)\mu_\mathrm{r}\sum_j \mathcal{T}(\theta_i-\theta_j,\mathbf{r}_i-\mathbf{r}_j)\Big]^2 + \frac{(1+\kappa)\mu_\mathrm{r}}{2}\frac{\partial}{\partial\theta_i}\sum_j \mathcal{T}(\theta_i-\theta_j,\mathbf{r}_i-\mathbf{r}_j) - \frac{(\kappa\mu_\mathrm{r})^2}{4D_\mathrm{r}}\sum_{j,k}\mathcal{T}(\theta_i-\theta_k,\mathbf{r}_i-\mathbf{r}_k)\mathcal{T}(\theta_i-\theta_j,\mathbf{r}_i-\mathbf{r}_j). \tag{5.111}$$

Expanding $\mathbb{B}_i^{(\mathrm{d})}$ in Eq. 5.110 and comparing with $\mathbb{B}_i^{(\mathrm{c})}$ in Eq. 5.111, it appears that $\mathcal{B}^{(\mathrm{c})}$ and $\mathcal{B}^{(\mathrm{d})}$ only differ by a term proportional to $\sum_{i,j}\dot{\theta}_i\mathcal{T}(\theta_i-\theta_j,\mathbf{r}_i-\mathbf{r}_j)$ which can safely be neglected at large $t$, thus proving the equivalence between ensembles (c) and (d).

## A3 Appendix for Chapter 4

### *A3.1 Mean-field Theory*

To build our mean-field theory, we begin with Eq. 4.13 in the main text, and we do not consider the polarization term further. Our choice is justified in the low-activity limit and, beyond that, supported by the results in Ref. [60]. In this previous work, the formulas for efficiency and mobility were obtained by setting two-point polarization correlators to zero (see Eqs. 8-9 in Appendix A of Ref. [60]), and the results agree with simulation data very well even at strong driving.

By ignoring polarization and by linearizing the dynamics of the density field $\rho$ around the overall density $\rho_0$, we arrive at closed-form equation of motion for $\delta\rho = \rho - \rho_0$. This linear approximation holds for weak interparticle potentials, so that any local density fluctuation is

small compared to the density of the liquid. The solution for $\delta\rho(\mathbf{k},t) = \int[\rho(\mathbf{r},t)-\rho_0]e^{-i\mathbf{k}\cdot\mathbf{r}}d\mathbf{r}$ follows readily as

$$\delta\rho(\mathbf{k},t) = \int_{-\infty}^{t} ds e^{-\mathbf{k}^2 G(\mathbf{k})(t-s)} \left(-\mathbf{k}^2\rho_0\varepsilon U(\mathbf{k})e^{-i\mathbf{k}\cdot\mathbf{r}_0(s)} + i\mathbf{k}\cdot\sqrt{2\rho_0 T}\Lambda(\mathbf{k},s)\right), \quad (5.112)$$

where $G(\mathbf{k}) = T + \rho_0 U(\mathbf{k})$, and $\mathbf{\Lambda}$ is a zero-mean Gaussian white noise with correlations

$$\langle\Lambda_\alpha(\mathbf{k},s)\Lambda_\beta(\mathbf{k}',s')\rangle = (2\pi)^d\delta_{\alpha\beta}\delta(s-s')\delta(\mathbf{k}+\mathbf{k}'), \quad (5.113)$$

where $d$ is the spatial dimension.

## *A3.2 Rate of Work*

Substituting the field equation of motion, given in Eq. 5.112, into the rate of work per particle, defined as

$$\begin{aligned}\dot{w} &= \frac{1}{N}\sum_{i,j\neq i}\left\langle \mathbf{f}_i\cdot\nabla_i\varepsilon U(\mathbf{r}_i-\mathbf{r}_j)\right\rangle \\ &= -\int\frac{d\mathbf{k}}{(2\pi)^d}\Big\langle \mathbf{f}_0(t)\cdot i\mathbf{k}\varepsilon U(\mathbf{k})\,\delta\rho(\mathbf{k},t)e^{i\mathbf{k}\cdot\mathbf{r}_0(t)}\Big\rangle, \end{aligned} \quad (5.114)$$

where we have used $U(\mathbf{k}) = U(-\mathbf{k})$, gives

$$\begin{aligned}\dot{w} = \rho_0\int\frac{d\mathbf{k}}{(2\pi)^d}\mathbf{k}^2(\varepsilon U(\mathbf{k}))^2\int_{-\infty}^{0} ds e^{\mathbf{k}^2G(\mathbf{k})s}\Big\langle i\mathbf{k}\cdot\mathbf{f}_0(0)e^{i\mathbf{k}\cdot(\mathbf{r}_0(0)-\mathbf{r}_0(s))}\Big\rangle \\ + \sqrt{2\rho_0 T}\int\frac{d\mathbf{k}}{(2\pi)^d}\mathbf{k}^2\varepsilon U(\mathbf{k})\int_{-\infty}^{0} ds e^{\mathbf{k}^2G(\mathbf{k})s}\left\langle e^{i\mathbf{k}\cdot\mathbf{r}_0(0)}\mathbf{f}_0(0)\cdot\Lambda(\mathbf{k},s)\right\rangle. \end{aligned} \quad (5.115)$$

From the tracer dynamics, see Eq. (2) in the main text, we deduce

$$\mathbf{r}_0(0) = \int_{-\infty}^{0}[\mathbf{f}_0(x)+\xi_0(x)]dx + \varepsilon\int\frac{d\mathbf{k}'}{(2\pi)^d}i\mathbf{k}'U(\mathbf{k}')\int_{-\infty}^{0} ds'\,\delta\rho(\mathbf{k}',s')e^{i\mathbf{k}\cdot\mathbf{r}_0(s')}, \quad (5.116)$$

from which, after expanding with respect to the parameter $\varepsilon$, follows

$$e^{i\mathbf{k}\cdot\mathbf{r}_0(0)} = e^{i\mathbf{k}\cdot\int_{-\infty}^{0}[\mathbf{f}_0(x)+\xi_0(x)]dx}\times \left[1-\varepsilon\int\frac{d\mathbf{k}'}{(2\pi)^d}\mathbf{k}\cdot\mathbf{k}'U(\mathbf{k}')\int_{-\infty}^{0}ds'\,\delta\rho(\mathbf{k}',s')e^{i\mathbf{k}'\cdot\int_{-\infty}^{s'}[\mathbf{f}_0(x)+\xi_0(x)]dx}+\mathcal{O}(\varepsilon^2)\right]. \tag{5.117}$$

Using the expression for $\delta\rho$ given in Eq. 5.112, we obtain

$$e^{i\mathbf{k}\cdot\mathbf{r}_0(0)} = e^{i\mathbf{k}\cdot\int_{-\infty}^{0}[\mathbf{f}_0(x)+\xi_0(x)]dx}\left[1-\varepsilon\int\frac{d\mathbf{k}'}{(2\pi)^d}(\mathbf{k}\cdot\mathbf{k}')U(\mathbf{k}')\int_{-\infty}^{0}ds'e^{i\mathbf{k}'\cdot\int_{-\infty}^{s'}[\mathbf{f}_0(x)+\xi_0(x)]dx}\times \int_{-\infty}^{s'}ds''e^{-\mathbf{k}'^2G(\mathbf{k}')(s'-s'')}i\mathbf{k}'\cdot\sqrt{2T\rho_0}\Lambda(\mathbf{k}',s'')+\mathcal{O}(\varepsilon^2)\right]. \tag{5.118}$$

Substituting Eq. 5.118 in Eq. 5.115, and truncating at second order, we then get

$$\begin{aligned}\dot{w} = {}& \rho_0\int\frac{d\mathbf{k}}{(2\pi)^d}\mathbf{k}^2(\varepsilon U(\mathbf{k}))^2\int_{-\infty}^{0}ds e^{\mathbf{k}^2G(\mathbf{k})s}\left\langle i\mathbf{k}\cdot\mathbf{f}_0(0)e^{i\mathbf{k}\cdot\int_{s}^{0}[\mathbf{f}_0(x)+\xi_0(x)]dx}\right\rangle \\ & -2\rho_0T\int\frac{d\mathbf{k}d\mathbf{k}'}{(2\pi)^{2d}}\mathbf{k}^2(\mathbf{k}\cdot\mathbf{k}')\varepsilon^2U(\mathbf{k})U(\mathbf{k}')\int_{-\infty}^{0}ds'e^{-\mathbf{k}'^2G(\mathbf{k}')s'} \\ & \times\left\langle i\mathbf{k}'\cdot\mathbf{f}_0(0)e^{i\mathbf{k}\cdot\int_{-\infty}^{0}[\mathbf{f}_0(x)+\xi_0(x)]dx+i\mathbf{k}'\cdot\int_{-\infty}^{s'}[\mathbf{f}_0(x)+\xi_0(x)]dx}\right\rangle \\ & \times\int_{-\infty}^{0}ds\int_{-\infty}^{s'}ds''e^{\mathbf{k}^2G(\mathbf{k})s+\mathbf{k}'^2G(\mathbf{k}')s''}\langle\Lambda_\alpha(\mathbf{k},s)\Lambda_\alpha(\mathbf{k}',s'')\rangle,\end{aligned} \tag{5.119}$$

where we have used that $\Lambda$, $\xi_0$, and $\mathbf{f}_0$ are independent. According to Wick's theorem, we can now write for the white noise

$$\left\langle e^{i\mathbf{k}\cdot\int_{s}^{0}\xi_0(x)dx}\right\rangle = e^{\mathbf{k}^2Ts}. \tag{5.120}$$

Starting from the time correlations of the active forces:

$$\left\langle f_{i\alpha}(t)f_{j\beta}(0)\right\rangle = \frac{T_{\mathrm{A}}}{\tau}\delta_{ij}\delta_{\alpha\beta}e^{-|t|/\tau}, \tag{5.121}$$

we can derive the following

$$\left\langle \int_s^0 f_{0\alpha}(0) f_{0\alpha}(x) dx \right\rangle = T_{\rm A}(1 - e^{s/\tau}), \tag{5.122}$$

and

$$\begin{aligned}\left\langle \int_s^0 \int_s^0 f_{0\alpha}(x) f_{0\alpha}(x') dx dx' \right\rangle &= \frac{T_{\rm A}}{\tau} \left( \int_s^0 \int_x^0 e^{-(x'-x)/\tau} dx' dx + \int_s^0 \int_s^x e^{-(x-x')/\tau} dx' dx \right) \\ &= -2\big[T_{\rm A} s + T_{\rm A}\tau(1 - e^{s/\tau})\big] \equiv -2R(s).\end{aligned} \tag{5.123}$$

According to Wick's theorem, we can now write two expressions that are needed to evaluate the quantities we are interested in:

$$\left\langle e^{i\mathbf{k}\cdot\int_s^0 \mathbf{f}_0(x) dx} \right\rangle = e^{\mathbf{k}^2 R(s)}. \tag{5.124}$$

and

$$\begin{aligned}\left\langle i\mathbf{k}\cdot\mathbf{f}_0(0) e^{i\mathbf{k}\int_s^0 \mathbf{f}_0(x) dx} \right\rangle &= -\mathbf{k}^2 \left\langle \int_s^0 f_{0\alpha}(0) f_{0\alpha}(x) dx \right\rangle e^{\mathbf{k}^2 R(s)} \\ &= -\mathbf{k}^2 e^{\mathbf{k}^2 R(s)} T_{\rm A}(1 - e^{s/\tau}).\end{aligned} \tag{5.125}$$

Going back to the expression for the rate of rate, we can collapse the delta functions in $\mathbf{k}+\mathbf{k}'$ and $s - s''$, and we use Eqs. (5.113, 5.120, 5.125) to get

$$\begin{aligned}\dot{w} = -\rho_0 T_{\rm A} \int \frac{d\mathbf{k}}{(2\pi)^d} \mathbf{k}^4 (\varepsilon U(\mathbf{k}))^2 \Bigg[ \int_{-\infty}^0 ds e^{\mathbf{k}^2((G(\mathbf{k})+T)s + R(s))}(1 - e^{s/\tau}) \\ + \frac{T}{G(\mathbf{k})} \int_{-\infty}^0 ds' e^{\mathbf{k}^2((G(\mathbf{k})+T)s' + R(s'))}(1 - e^{s'/\tau}) \Bigg].\end{aligned} \tag{5.126}$$

Therefore, the general formula for rate of work is

$$\dot{w} = -\rho_0 T_{\mathrm{A}} \int \frac{d\mathbf{k}}{(2\pi)^d} \mathbf{k}^4 (\varepsilon U(\mathbf{k}))^2 \frac{G(\mathbf{k}) + T}{G(\mathbf{k})} \int_{-\infty}^{0} ds e^{\mathbf{k}^2((G(\mathbf{k})+T)s + R(s))} (1 - e^{s/\tau}). \quad (5.127)$$

To get Eq. 4.19 in the main text, we set $\varepsilon = 1$ into the equation above (valid for weak potentials), but to extend validity beyond weak interaction regimes we substitute one of the two $U(\mathbf{k})$ in Eq. 4.19, the one that originated from the field equation of motion, by $-Tc_{\mathrm{eq}}(\mathbf{k})$. We also perform the same substitution in $G(\mathbf{k})$, yielding $G(\mathbf{k}) = T(1 - \rho_0 c_{\mathrm{eq}}(\mathbf{k}))$.

## *A3.3 Pair Correlation Function*

Similarly to the derivation of the previous result, we substitute the field equation of motion, Eq. 5.112, into

$$\rho_0 h(\mathbf{k}) = \left\langle e^{i\mathbf{k}\cdot\mathbf{r}_0} \delta\rho(\mathbf{k}) \right\rangle, \quad (5.128)$$

yielding

$$\rho_0 h(\mathbf{k}) = \left\langle e^{i\mathbf{k}\cdot\mathbf{r}_0(0)} \int_{-\infty}^{0} ds e^{\mathbf{k}^2 G(\mathbf{k}) s} \left[ -\rho_0 \mathbf{k}^2 \varepsilon U(\mathbf{k}) e^{-i\mathbf{k}\cdot\mathbf{r}_0(s)} + i\mathbf{k} \cdot \sqrt{2\rho_0 T} \Lambda(\mathbf{k}, s) \right] \right\rangle. \quad (5.129)$$

Substituting the tracer dynamics, Eq. 5.116, and expanding to first order in $\varepsilon$, we get

$$\begin{aligned}
\rho_0 h(\mathbf{k}) = & -\rho_0 \mathbf{k}^2 \varepsilon U(\mathbf{k}) \int_{-\infty}^{0} ds e^{\mathbf{k}^2((G(\mathbf{k})+T)s + R(s))} \\
& - 2\rho_0 T \int \frac{d\mathbf{k}'}{(2\pi)^d} (\mathbf{k}\cdot\mathbf{k}')^2 \varepsilon U(\mathbf{k}') \int_{-\infty}^{0} ds' e^{-\mathbf{k}'^2 G(\mathbf{k}') s'} \\
& \times \left\langle e^{i\mathbf{k}\cdot\int_{-\infty}^{0} [\mathbf{f}_0(x) + \xi_0(x)] dx + i\mathbf{k}'\cdot\int_{-\infty}^{s'} [\mathbf{f}_0(x) + \xi_0(x)] dx} \right\rangle \\
& \times \int_{-\infty}^{0} ds \int_{-\infty}^{s'} ds'' e^{\mathbf{k}^2 G(\mathbf{k}) s + \mathbf{k}'^2 G(\mathbf{k}') s''} \langle \Lambda_\alpha(\mathbf{k}, s) \Lambda_\alpha(\mathbf{k}', s'') \rangle. \quad (5.130)
\end{aligned}$$

Collapsing noise correlation functions and using Eqs. (5.113, 5.120, 5.124), we obtain the final form of the pair correlation function in Eq. 4.15 in Sec. 4.3.

In order to obtain Eq. 4.16 in Sec.4.3 for the pair correlation function, the equation above is handled similarly to the equation for rate of work.

## *A3.4 $\tilde{I}$ in the Weak Interaction Limit*

To obtain the form of $\tilde{I}$ in the weak-interaction limit, we can start from Eq. (26) in Ref. [180] for an active liquid:

$$\begin{aligned} \dot{w} &= \rho_0 \int g(\mathbf{r})[\nabla U(\mathbf{r})^2 - T\nabla^2 U(\mathbf{r})]d\mathbf{r} + \rho_0^2 \iint g_3(\mathbf{r},\mathbf{r}')[\nabla U(\mathbf{r})\cdot\nabla U(\mathbf{r}')]d\mathbf{r}d\mathbf{r}', \\ g(\mathbf{r}) &= h(\mathbf{r}) + 1, \quad \rho_0^2\, g_3(\mathbf{r},\mathbf{r}') = \frac{1}{N^2}\sum_{i\neq j\neq k} \big\langle \delta(\mathbf{r}-\mathbf{r}_i+\mathbf{r}_j)\delta(\mathbf{r}'-\mathbf{r}_i+\mathbf{r}_k)\big\rangle. \end{aligned} \tag{5.131}$$

Using the integrated form of the Yvon-Born-Green relation [79], valid for an equilibrium liquid and given by

$$0 = \int g_{\rm eq}(\mathbf{r})[\nabla U(\mathbf{r})^2 - T\nabla^2 U(\mathbf{r})]d\mathbf{r} + \rho_0 \iint g_3^{\rm eq}(\mathbf{r},\mathbf{r}')[\nabla U(\mathbf{r})\cdot\nabla U(\mathbf{r}')]d\mathbf{r}d\mathbf{r}', \tag{5.132}$$

we can write Eq. 5.131 as

$$\frac{\dot{w}}{\rho_0} = \int (h-h_{\rm eq})(\mathbf{r})[\nabla U(\mathbf{r})^2 - T\nabla^2 U(\mathbf{r})]d\mathbf{r} + \rho_0 \iint [g_3(\mathbf{r},\mathbf{r}') - g_3^{\rm eq}(\mathbf{r},\mathbf{r}')][\nabla U(\mathbf{r})\cdot\nabla U(\mathbf{r}')]d\mathbf{r}d\mathbf{r}' \tag{5.133}$$

We proceed by employing the simple approximation $g_3(\mathbf{r},\mathbf{r}') \approx g(\mathbf{r})g(\mathbf{r}')g(\mathbf{r}-\mathbf{r}') = (1 + [g(\mathbf{r})-1])(1+[g(\mathbf{r}')-1])(1+[g(\mathbf{r}-\mathbf{r}')-1]) \approx -2 + g(\mathbf{r}) + g(\mathbf{r}') + g(\mathbf{r}-\mathbf{r}')$, where we have assumed that deviations of the two-body correlation function $g$ from unity can be neglected

beyond first order in the regime of weak interactions, to derive

$$\frac{\dot{w}}{\rho_0} = \int (h-h_{\rm eq})(\mathbf{r})[\nabla U(\mathbf{r})^2 - T\nabla^2 U(\mathbf{r})]d\mathbf{r} + \rho_0 \iint [g(\mathbf{r}-\mathbf{r}') - g_{\rm eq}(\mathbf{r}-\mathbf{r}')][\nabla U(\mathbf{r})\cdot\nabla U(\mathbf{r}')]d\mathbf{r}d\mathbf{r}' \tag{5.134}$$

Working in the Fourier domain, and returning to our problem where we have set up perturbation in $\varepsilon$, we get

$$\tilde{I} = \dot{w} - \rho_0^2 \int d\mathbf{k}\mathbf{k}^2 U(\mathbf{k})^2 (h-h_{\rm eq})(\mathbf{k}). \tag{5.135}$$

In order to plot the theory curve in Fig. 2 of the main text at low potential strength $A$, we simply substitute Eqs. 4.15 and 5.127 into Eq. 5.135. Note that, although we do not use the substitution of $U(\mathbf{k})$ by $-Tc_{\rm eq}(\mathbf{k})$ in this weak interaction regime, our prediction still captures well the slope in Fig. 2 of the main text. Instead, we use the explicit expression of $U(\mathbf{k})$ for the harmonic potential in 2D, given by $U(\mathbf{k}) = (2\pi A/\mathbf{k}^2)\big[\pi(J_1(|\mathbf{k}|)H_0(|\mathbf{k}|) - J_0(|\mathbf{k}|)H_1(|\mathbf{k}|)) - 2J_2(|\mathbf{k}|)\big]$, where $J_n$ and $H_n$ are respectively the Bessel and Struve functions of order $n$.

At low density, Eq. 5.135 yields $\tilde{I} = \dot{w}$, so that the slope in parametric plot of $\dot{w}$ vs. $\tilde{I}$ converges to 1, which is independent of both the type of interactions and the persistence time $\tau$. At higher density, there still remains a linear relation between $\dot{w}$ and $\tilde{I}$, see Fig. 4.3 and also Fig. 4.4 (right) of the main text, so that the second term in Eq. 5.135 effectively leads to renormalize the value of the slope.

### *A3.5 Rate of Work and $\tilde{I}$ in Active Rotors*

The active force driving the counterclockwise rotation of a rotor moves each pole of the rotor a distance $v_0 dt$ in the infinitesimal time interval $dt$. It can correspondingly be thought of as a *translational* force acting on each pole of the rotor, pushing it against conservative forces resulting from interactions with neighboring rotor poles in a manner analogous to the colored noise driving force in Eq. 4.18 of AOUPs. The active work per rotor therefore has

the same form as Eq. 4.18, with the AOUP self-propulsion velocity $\mathbf{f}_i$ being replaced with $v_0$ multiplied by the orientation vector $\hat{\mathbf{u}}$ of the rotor pole, summed over both rotor poles:

$$\dot{w} = \frac{1}{N} \sum_{i,j \neq i} \sum_{k,l} v_0 \langle \hat{\mathbf{u}}_k \cdot \nabla_{i,k} U(\mathbf{r}_{i,k} - \mathbf{r}_{j,l}) \rangle. \tag{5.136}$$

Here, $\hat{\mathbf{u}}_h = (\cos(\theta + \pi/2), \sin(\theta + \pi/2)) = (\sin(\theta), -\cos(\theta))$. Similarly, $\hat{\mathbf{u}}_t = (\cos(\theta + 3\pi/2), \sin(\theta + 3\pi/2)) = (-\sin(\theta), \cos(\theta))$. To define $\tilde{I}$ for the active rotors, we again think of each rotor as 2 isotropic point particles interacting with neighboring particles. As with AOUPs, their motion leads to a buildup of particle density at their leading edge. The resulting change in the pair correlation function of the AOUPs is essentially what we connect to their rate of active work. We therefore use a rotor-pole-based pair correlation function in our definition of $\tilde{I}$ for active rotors:

$$h_R(\mathbf{r}) = \frac{1}{N} \sum_{i,j \neq i} \sum_{k,l} \left\langle \delta(\mathbf{r} - \mathbf{r}_{i,k} + \mathbf{r}_{j,l}) \right\rangle \tag{5.137}$$

$$\tilde{I}_R = 2\rho_0 \int \left[ (\nabla U)^2 - T\nabla^2 U \right] (h_R - h_{R,\mathrm{eq}}) d\mathbf{r} \tag{5.138}$$

Note that the definition of $h_R$ does not count the opposite pole of the same rotor in measuring the distribution of rotor poles, and that $\rho_0$ is the numerical density of rotors, which is half of the density of rotor poles. We emphasize that while we can simply define $\dot{w}$ and $\tilde{I}$ for active rotors, we did not establish the existence of a close theoretical connection between these quantities as we did for AOUPs, and *a priori* had no strong reason to assume that $\tilde{I}$ would be a measure of the system structure that obeyed a fixed relationship with $\dot{w}$. Indeed, we observed that it is *not* a useful measure of the structure for the smallest rotors, although it is reasonable for larger rotors.

# REFERENCES


[1] W. W. Ahmed et al. "Active Mechanics Reveal Molecular-Scale Force Kinetics in Living Oocytes". *Biophys. J.*, 114 (2018).

[2] Takayuki Ariga, Michio Tomishige, and Daisuke Mizuno. "Nonequilibrium Energetics of Molecular Motor Kinesin". *Phys. Rev. Lett.*, 121 (2018).

[3] Nicolas Bain and Denis Bartolo. "Dynamic response and hydrodynamics of polarized crowds". *Science*, 363.6422 (2019).

[4] M. Ballerini et al. "Interaction ruling animal collective behavior depends on topological rather than metric distance: Evidence from a field study". *Proc. Natl. Acad. Sci. USA*, 105.4 (2008).

[5] Andre C. Barato and Udo Seifert. "Thermodynamic Uncertainty Relation for Biomolecular Processes". *Phys. Rev. Lett.*, 114.15 (2015).

[6] Urna Basu, Christian Maes, and Karel Netočný. "Statistical forces from close-to-equilibrium media". *New J. Phys.*, 17.11 (2015).

[7] Christopher Battle et al. "Broken detailed balance at mesoscopic scales in active biological systems". *Science*, 351.6268 (2016).

[8] Clemens Bechinger et al. "Active particles in complex and crowded environments". *Rev. Mod. Phys.*, 88 (2016).

[9] O. Bénichou et al. "Nonlinear response and emerging nonequilibrium microstructures for biased diffusion in confined crowded environments". *Phys. Rev. E*, 93 (2016).

[10] Thibault Bertrand et al. "Optimized Diffusion of Run-and-Tumble Particles in Crowded Environments". *Phys. Rev. Lett.*, 120 (2018).

[11] Julian Bialké et al. "Negative Interfacial Tension in Phase-Separated Active Brownian Particles". *Phys. Rev. Lett.*, 115.9 (2015).

[12] Gili Bisker and Jeremy L. England. "Nonequilibrium associative retrieval of multiple stored self-assembly targets". *Proc. Natl. Acad. Sci. USA*, 115.45 (2018).

[13] Laurent Blanchoin et al. "Actin dynamics, architecture, and mechanics in cell motility". *Physiol. Rev.*, 94.1 (2014).

[14] Thierry Bodineau, Vivien Lecomte, and Cristina Toninelli. "Finite Size Scaling of the Dynamical Free-Energy in a Kinetically Constrained Model". *J. Stat. Phys.*, 147.1 (2012).

[15] Tobias Brewer et al. "Efficient characterisation of large deviations using population dynamics". *J. Stat. Mech.*, 2018.5 (2018).

[16] David B. Brückner, Pierre Ronceray, and Chase P. Broedersz. "Inferring the Dynamics of Underdamped Stochastic Systems". *Phys. Rev. Lett.*, 125.5 (2020).

[17] Eric W. Burkholder and John F. Brady. "Tracer diffusion in active suspensions". *Phys. Rev. E*, 95 (2017).

[18] Keith T. Butler et al. "Machine learning for molecular and materials science". *Nature*, 559.7715 (2018).

[19] F. Cagnetta et al. "Large Fluctuations and Dynamic Phase Transition in a System of Self-Propelled Particles". *Phys. Rev. Lett.*, 119.15 (2017).

[20] Giuseppe Carleo et al. "Machine learning and the physical sciences". *Rev. Mod. Phys.*, 91.4 (2019).

[21] C. Casert et al. "Interpretable machine learning for inferring the phase boundaries in a nonequilibrium system". *Phys. Rev. E*, 99.2 (2019).

[22] Michael E. Cates and Julien Tailleur. "Motility-Induced Phase Separation". *Annu. Rev. CMP*, 6.1 (2015).

[23] Andrea Cavagna and Irene Giardina. "Bird Flocks as Condensed Matter". *Ann. Rev. Condens. Matter Phys.*, 5.1 (2014).

[24] David Chandler. "Gaussian field model of fluids with an application to polymeric fluids". *Phys. Rev. E*, 48.4 (1993).

[25] H. Chaté. "Dry Aligning Dilute Active Matter". *Annu. Rev. Condens. Matter Phys.*, 11 (2020).

[26] Shriram Chennakesavalu and Grant M. Rotskoff. "Cooperative multi-agent reinforcement learning for high-dimensional nonequilibrium control". *ArXiv e-prints* (2021).

[27] Raphaël Chetrite and Hugo Touchette. "Nonequilibrium Microcanonical and Canonical Ensembles and Their Equivalence". *Phys. Rev. Lett.*, 111 (2013).

[28] François Chollet et al. *Keras*. 2015.

[29] Frank Cichos et al. "Machine learning for active matter". *Nat. Mach. Intell.*, 2.2 (2020).

[30] Jonathan Colen et al. "Machine learning active-nematic hydrodynamics". *Proc. Natl. Acad. Sci. USA*, 118.10 (2021).

[31] Emanuele Crosato, Mikhail Prokopenko, and Richard E. Spinney. "Irreversibility and emergent structure in active matter". *Phys. Rev. E*, 100.4 (2019).

[32] Lennart Dabelow, Stefano Bo, and Ralf Eichhorn. "Irreversibility in Active Matter Systems: Fluctuation Theorem and Mutual Information". *Phys. Rev. X*, 9 (2019).

[33] Kinjal Dasbiswas, Kranthi K. Mandadapu, and Suriyanarayanan Vaikuntanathan. "Topological localization in out-of-equilibrium dissipative systems". *Proc. Natl. Acad. Sci. USA*, 115.39 (2018).

[34] C. De Dominicis. "A Lagrangian version of Halperin-Hohenberg-Ma models for the dynamics of critical phenomena". *Lett. nuovo cimento (1971-1985)*, 12.15 (1975).

[35] David Dean, I. Drummond, and Ronald Horgan. "Effective transport properties for diffusion in random media". *J. Stat. Mech. Theory Exp.*, 2007 (2007).

[36] David S Dean. "Langevin equation for the density of a system of interacting Langevin processes". *J. Phys. A.: Math. Gen.*, 29.24 (1996).

[37] Stephen J DeCamp et al. "Orientational order of motile defects in active nematics". *Nat. Mat.*, 14.11 (2015).

[38] Vincent Démery. "Mean-field microrheology of a very soft colloidal suspension: Inertia induces shear thickening". *Phys. Rev. E*, 91 (2015).

[39] Vincent Démery, Olivier Bénichou, and Hugo Jacquin. "Generalized Langevin equations for a driven tracer in dense soft colloids: construction and applications". *New J. Phys.*, 16.5 (2014).

[40] Vincent Démery and David S. Dean. "Perturbative path-integral study of active- and passive-tracer diffusion in fluctuating fields". *Phys. Rev. E*, 84 (2011).

[41] Vincent Démery and Étienne Fodor. "Driven probe under harmonic confinement in a colloidal bath". *J. Stat. Mech.*, 2019.3 (2019).

[42] Christopher Dombrowski et al. "Self-Concentration and Large-Scale Coherence in Bacterial Dynamics". *Phys. Rev. Lett.*, 93.9 (2004).

[43] J. L. Doob. "The Brownian Movement and Stochastic Equations". *Ann. Math.*, 43.2 (1942).

[44] Austin R. Dulaney and John F. Brady. "Machine Learning for Phase Behavior in Active Matter Systems". *ArXiv e-prints* (2020).

[45] Mihir Durve, Fernando Peruani, and Antonio Celani. "Learning to flock through reinforcement". *Phys. Rev. E*, 102.1 (2020).

[46] Timothy Ekeh, Michael E. Cates, and Étienne Fodor. "Thermodynamic cycles with active matter". *Phys. Rev. E*, 102 (2020).

[47] J Elgeti, R G Winkler, and G Gompper. "Physics of microswimmers–single particle motion and collective behavior: a review". *Rep. Prog. Phys.*, 78.5 (2015).

[48] Jeremy L. England. "Dissipative adaptation in driven self-assembly". *Nat. Nano.*, 10 (2015).

[49] R. M. L. Evans. "Rules for Transition Rates in Nonequilibrium Steady States". *Phys. Rev. Lett.*, 92.15 (2004).

[50] Martin J. Falk et al. "Learning to control active matter". *Phys. Rev. Research*, 3.3 (2021).

[51] T. F. F. Farage, P. Krinninger, and J. M. Brader. "Effective interactions in active Brownian suspensions". *Phys. Rev. E*, 91 (2015).

[52] F. D. C. Farrell et al. "Pattern Formation in Self-Propelled Particles with Density-Dependent Motility". *Phys. Rev. Lett.*, 108 (2012).

[53] Grégoire Ferré and Hugo Touchette. "Adaptive Sampling of Large Deviations". *J. Stat. Phys.*, 172.6 (2018).

[54] Yaouen Fily and M. Cristina Marchetti. "Athermal Phase Separation of Self-Propelled Particles with No Alignment". *Phys. Rev. Lett.*, 108 (2012).

[55] Daniel S Fisher. "Random walks in random environments". *Phys. Rev. A*, 30.2 (1984).

[56] Daniel A. Fletcher and Phillip L. Geissler. "Active Biological Materials". *Annu. Rev. Phys. Chem.*, 60.1 (2009).

[57] É. Fodor et al. "Activity-driven fluctuations in living cells". *Europhys. Lett.*, 110.4 (2015).

[58] É. Fodor et al. "Nonequilibrium dissipation in living oocytes". *Europhys. Lett.*, 116.3 (2016).

[59] Étienne Fodor and M. C Marchetti. "The statistical physics of active matter: From self-catalytic colloids to living cells". *Physica A*, 504 (2018).

[60] Étienne Fodor, Takahiro Nemoto, and Suriyanarayanan Vaikuntanathan. "Dissipation controls transport and phase transitions in active fluids: Mobility, diffusion and biased ensembles". *New J. Phys.*, 22 (2020).

[61] Étienne Fodor et al. "How Far from Equilibrium Is Active Matter?" *Phys. Rev. Lett.*, 117.3 (2016).

[62] Christopher J. Fullerton and Robert L. Jack. "Dynamical phase transitions in supercooled liquids: Interpreting measurements of dynamical activity". *J. Chem. Phys.*, 138.22 (2013).

[63] M. L. Gardel et al. "Elastic Behavior of Cross-Linked and Bundled Actin Networks". *Science*, 304.5675 (2004).

[64] J. P. Garrahan et al. "Dynamical First-Order Phase Transition in Kinetically Constrained Models of Glasses". *Phys. Rev. Lett.*, 98.19 (2007).

[65] Mattia Gazzola, Babak Hejazialhosseini, and Petros Koumoutsakos. "Reinforcement learning and wavelet adapted vortex methods for simulations of self-propelled swimmers". *SIAM J. Sci. Comput.*, 36.3 (2014).

[66] Cristian Giardinà, Jorge Kurchan, and Luca Peliti. "Direct Evaluation of Large-Deviation Functions". *Phys. Rev. Lett.*, 96.12 (2006).

[67] Todd R Gingrich, Grant M Rotskoff, and Jordan M Horowitz. "Inferring dissipation from current fluctuations". *J. Phys. A: Math. Theor.*, 50.18 (2017).

[68] Todd R. Gingrich et al. "Dissipation Bounds All Steady-State Current Fluctuations". *Phys. Rev. Lett.*, 116 (2016).

[69] J. Gladrow et al. "Broken Detailed Balance of Filament Dynamics in Active Networks". *Phys. Rev. Lett.*, 116 (2016).

[70] F. S. Gnesotto et al. "Broken detailed balance and non-equilibrium dynamics in living systems: a review". *Rep. Prog. Phys.*, 81.6 (2018).

[71] Federico S. Gnesotto et al. "Learning the non-equilibrium dynamics of Brownian movies". *Nat. Commun.*, 11.1 (2020).

[72] Gerhard Gompper et al. "The 2020 motile active matter roadmap". *J. Phys.: Condens. Matter*, 32.19 (2020).

[73] Trevor GrandPre et al. "Entropy production fluctuations encode collective behavior in active matter". *Phys. Rev. E*, 103.1 (2021).

[74] Michael Grünwald et al. "Exploiting non-equilibrium phase separation for self-assembly". *Soft Matter*, 12.5 (2016).

[75] Jules Guioth and Eric Bertin. "Lack of an equation of state for the nonequilibrium chemical potential of gases of active particles in contact". *J. Chem. Phys.*, 150 (2019).

[76] Ming Han et al. “Effective temperature concept evaluated in an active colloid mixture”. *Proc. Natl. Acad. Sci. USA*, 114.29 (2017).

[77] Benjamin Hancock and Aparna Baskaran. “Statistical mechanics and hydrodynamics of self-propelled hard spheres”. *J. Stat. Mech. Theory Exp.*, 2017.3 (2017).

[78] Peter Hanggi. “Driven Quantum Systems”. *Quantum Transp. Dissipation*. Ed. by T Dittrich et al. Wiley-VCH, 1998. Chap. 5 pp. 249–286.

[79] Jean-Pierre Hansen and Ian R. McDonald. *Theory of Simple Liquids (Fourth Edition)*. Oxford: Academic Press, 2013.

[80] Takahiro Harada and Shin-ichi Sasa. “Energy dissipation and violation of the fluctuation-response relation in nonequilibrium Langevin systems”. *Phys. Rev. E*, 73 (2006).

[81] Takahiro Harada and Shin-ichi Sasa. “Equality Connecting Energy Dissipation with a Violation of the Fluctuation-Response Relation”. *Phys. Rev. Lett.*, 95 (2005).

[82] Lester O. Hedges et al. “Dynamic Order-Disorder in Atomistic Models of Structural Glass Formers”. *Science*, 323.5919 (2009).

[83] Daniel Hexner, Andrea J. Liu, and Sidney R. Nagel. “Periodic training of creeping solids”. *Proc. Natl. Acad. Sci. USA*, 117.50 (2020).

[84] Jordan Horowitz and Christopher Jarzynski. “Illustrative example of the relationship between dissipation and relative entropy”. *Phys. Rev. E*, 79.2 (2009).

[85] G. Hummer et al. “An information theory model of hydrophobic interactions.” *Proc. Natl. Acad. Sci.*, 93.17 (1996).

[86] Pablo I. Hurtado and Pedro L. Garrido. “Test of the Additivity Principle for Current Fluctuations in a Model of Heat Conduction”. *Phys. Rev. Lett.*, 102.25 (2009).

[87] Pierre Illien et al. “Nonequilibrium Fluctuations and Enhanced Diffusion of a Driven Particle in a Dense Environment”. *Phys. Rev. Lett.*, 120 (2018).

[88] R. L. Jack and P. Sollich. “Effective interactions and large deviations in stochastic processes”. *Eur. Phys. J. Special Topics*, 224.12 (2015).

[89] Robert L Jack and Peter Sollich. “Large deviations of the dynamical activity in the East model: analysing structure in biased trajectories”. *J. Phys. A: Math. Theor.*, 47.1 (2014).

[90] Robert L. Jack and Peter Sollich. “Large Deviations and Ensembles of Trajectories in Stochastic Models”. *Prog. Theor. Phys. Supp.*, 184 (2010).

[91] Robert L. Jack, Ian R. Thompson, and Peter Sollich. "Hyperuniformity and Phase Separation in Biased Ensembles of Trajectories for Diffusive Systems". *Phys. Rev. Lett.*, 114 (2015).

[92] Abhijeet Joshi et al. "The interplay between activity and filament flexibility determines the emergent properties of active nematics". *Soft Matter*, 15 (2019).

[93] Clara del Junco, Laura Tociu, and Suriyanarayanan Vaikuntanathan. "Energy dissipation and fluctuations in a driven liquid". *Proc. Natl. Acad. Sci. USA*, 115.14 (2018).

[94] Katherine Klymko et al. "Rare behavior of growth processes via umbrella sampling of trajectories". *Phys. Rev. E*, 97.3 (2018).

[95] Teruhisa S. Komatsu and Naoko Nakagawa. "Expression for the Stationary Distribution in Nonequilibrium Steady States". *Phys. Rev. Lett.*, 100 (2008).

[96] Matthias Krüger and David S. Dean. "A Gaussian theory for fluctuations in simple liquids". *J. Chem. Phys.*, 146.13 (2017).

[97] Nitin Kumar et al. "Flocking at a distance in active granular matter". *Nat. Commun.*, 5.4688 (2014).

[98] Felix Kümmel et al. "Circular Motion of Asymmetric Self-Propelling Particles". *Phys. Rev. Lett.*, 110.19 (2013).

[99] T. Laffargue et al. "Large deviations of Lyapunov exponents". *J. Phys. A*, 46.25 (2013).

[100] Ganhui Lan et al. "The energy–speed–accuracy trade-off in sensory adaptation". *Nat. Phys.*, 8.5 (2012).

[101] B. Lander et al. "Noninvasive measurement of dissipation in colloidal systems". *Phys. Rev. E*, 86 (2012).

[102] Pushkar P. Lele, Basarab G. Hosu, and Howard C. Berg. "Dynamics of mechanosensing in the bacterial flagellar motor". *Proc. Natl. Acad. Sci. USA*, 110.29 (2013).

[103] J. Li et al. "Quantifying dissipation using fluctuating currents". *Nat. Commun.*, 10 (2019).

[104] Zhenghan Liao, William T. M. Irvine, and Suriyanarayanan Vaikuntanathan. "Rectification in Nonequilibrium Parity Violating Metamaterials". *Phys. Rev. X*, 10 (2020).

[105] Shneior Lifson and Julius L. Jackson. "On the Self-Diffusion of Ions in a Polyelectrolyte Solution". *J. Chem. Phys.*, 36.9 (1962).

[106] David T. Limmer and David Chandler. "Theory of amorphous ices". *Proc. Natl. Acad Sci. USA*, 111.26 (2014).

[107] Christian Maes and Karel Netočný. "Rigorous meaning of McLennan ensembles". *J. Math. Phys.*, 51.1 (2010).

[108] Christian Maes, Karel Netočný, and Bidzina M. Shergelashvili. "Nonequilibrium relation between potential and stationary distribution for driven diffusion". *Phys. Rev. E*, 80 (2009).

[109] C. Maggi et al. "Multidimensional stationary probability distribution for interacting active particles". *Sci. Rep.*, 5 (2015).

[110] Satya N. Majumdar and Alan J. Bray. "Large-deviation functions for nonlinear functionals of a Gaussian stationary Markov process". *Phys. Rev. E*, 65 (2002).

[111] Dibyendu Mandal, Katherine Klymko, and Michael R. DeWeese. "Entropy Production and Fluctuation Theorems for Active Matter". *Phys. Rev. Lett.*, 119 (2017).

[112] M. C. Marchetti et al. "Hydrodynamics of soft active matter". *Rev. Mod. Phys.*, 85 (2013).

[113] Umberto Marini Bettolo Marconi and Claudio Maggi. "Towards a statistical mechanical theory of active fluids". *Soft Matter*, 11.45 (2015).

[114] D. Martin et al. "Extracting maximum power from active colloidal heat engines". *Europhys. Lett.*, 121.6 (2018).

[115] P. C. Martin, E. D. Siggia, and H. A. Rose. "Statistical Dynamics of Classical Systems". *Phys. Rev. A*, 8 (1973).

[116] I. A. Martınez et al. "Inferring broken detailed balance in the absence of observable currents". *Nat. Commun.*, 10 (2019).

[117] Samuel R. McCandlish, Aparna Baskaran, and Michael F. Hagan. "Spontaneous segregation of self-propelled particles with different motilities". *Soft Matter*, 8.8 (2012).

[118] James A. McLennan. "Statistical Mechanics of the Steady State". *Phys. Rev.*, 115 (1959).

[119] Daisuke Mizuno et al. "Nonequilibrium Mechanics of Active Cytoskeletal Networks". *Science*, 315.5810 (2007).

[120] S. Muiños-Landin et al. "Reinforcement learning with artificial microswimmers". *Sci. Robot.*, 6.52 (2021).

[121] Michael Murrell et al. "Forcing cells into shape: the mechanics of actomyosin contractility". *Nat. Rev. Mol. Cell Biol.*, 16.8 (2015).

[122] Arvind Murugan and Suriyanarayanan Vaikuntanathan. "Topologically protected modes in non-equilibrium stochastic systems". *Nat. Commun.*, 8 (2017).

[123] Arvind Murugan et al. "Multifarious assembly mixtures: Systems allowing retrieval of diverse stored structures". *Proc. Natl. Acad. Sci. USA*, 112.1 (2015).

[124] Saroj Kumar Nandi. "Understanding the approximations of mode-coupling theory for sheared steady states of colloids". *Phys. Rev. E*, 92.4 (2015).

[125] Vijay Narayan, Sriram Ramaswamy, and Narayanan Menon. "Long-lived giant number fluctuations in a swarming granular nematic." *Science*, 317.5834 (2007).

[126] Cesare Nardini et al. "Entropy Production in Field Theories without Time-Reversal Symmetry: Quantifying the Non-Equilibrium Character of Active Matter". *Phys. Rev. X*, 7 (2017).

[127] Takahiro Nemoto, Robert L. Jack, and Vivien Lecomte. "Finite-Size Scaling of a First-Order Dynamical Phase Transition: Adaptive Population Dynamics and an Effective Model". *Phys. Rev. Lett.*, 118.11 (2017).

[128] Takahiro Nemoto et al. "Optimizing active work: Dynamical phase transitions, collective motion, and jamming". *Phys. Rev. E*, 99 (2019).

[129] Takahiro Nemoto et al. "Population-dynamics method with a multicanonical feedback control". *Phys. Rev. E*, 93.6 (2016).

[130] H. P. Nguyen Nguyen et al. "Emergent Collective Phenomena in a Mixture of Hard Shapes through Active Rotation". *Phys. Rev. Lett.*, 112 (2014).

[131] M. Nguyen and S. Vaikuntanathan. "Dissipation induced transitions in two-dimensional elastic membranes". *ArXiv e-prints* (2019).

[132] Michael Nguyen and Suriyanarayanan Vaikuntanathan. "Design principles for nonequilibrium self-assembly." *Proc. Natl. Acad. Sci. USA*, 113.50 (2016).

[133] Pelerine Tsobgni Nyawo and Hugo Touchette. "Dynamical phase transition in drifted Brownian motion". *Phys. Rev. E*, 98 (2018).

[134] J. Palacci et al. "Living Crystals of Light-Activated Colloidal Surfers". *Science*, 339.6122 (2013).

[135] Alexander P. Petroff, Xiao-Lun Wu, and Albert Libchaber. "Fast-Moving Bacteria Self-Organize into Active Two-Dimensional Crystals of Rotating Cells". *Phys. Rev. Lett.*, 114.15 (2015).

[136] Patrick Pietzonka and Udo Seifert. "Entropy production of active particles and for particles in active baths". *J. Phys. A: Math. Theor.*, 51.1 (2018).

[137] Patrick Pietzonka et al. "Autonomous Engines Driven by Active Matter: Energetics and Design Principles". *Phys. Rev. X*, 9 (2019).

[138] E. Pitard, V. Lecomte, and F. van Wijland. "Dynamic transition in an atomic glass former: A molecular-dynamics evidence". *Europhys. Lett.*, 96.5 (2011).

[139] Steve Plimpton. "Fast Parallel Algorithms for Short-Range Molecular Dynamics". *J. Comput. Phys.*, 117.1 (1995).

[140] V. Popkov and G. M. Schütz. "Transition Probabilities and Dynamic Structure Function in the ASEP Conditioned on Strong Flux". *J. Stat. Phys.*, 142.3 (2011).

[141] Vladislav Popkov, Gunter M Schütz, and Damien Simon. "ASEP on a ring conditioned on enhanced flux". *J. Stat. Mech.*, 2010.10 (2010).

[142] Karel Proesmans and Bernard Derrida. "Large-deviation theory for a Brownian particle on a ring: a WKB approach". *J. Stat. Mech.*, 2019.2 (2019).

[143] Ushnish Ray, Garnet Kin-Lic Chan, and David T. Limmer. "Exact Fluctuations of Nonequilibrium Steady States from Approximate Auxiliary Dynamics". *Phys. Rev. Lett.*, 120.21 (2018).

[144] Gabriel S. Redner, Michael F. Hagan, and Aparna Baskaran. "Structure and Dynamics of a Phase-Separating Active Colloidal Fluid". *Phys. Rev. Lett.*, 110 (2013).

[145] Gabriel S. Redner et al. "Classical Nucleation Theory Description of Active Colloid Assembly". *Phys. Rev. Lett.*, 117.14 (2016).

[146] Markus Rein and Thomas Speck. "Applicability of effective pair potentials for active Brownian particles". *Eur. Phys. J. E*, 39.9 (2016).

[147] David Richard, Hartmut Löwen, and Thomas Speck. "Nucleation pathway and kinetics of phase-separating active Brownian particles". *Soft Matter*, 12.24 (2016).

[148] Édgar Roldán et al. "Quantifying entropy production in active fluctuations of the hair-cell bundle from time irreversibility and uncertainty relations". *ArXiv e-prints* (2021).

[149] J.S. Rowlinson and B. Widom. *Molecular Theory of Capillarity*. Oxford: Clarendon Press, 1982.

[150] Syeda Sabrina et al. "Coarsening dynamics of binary liquids with active rotation." *Soft Matter*, 11.43 (2015).

[151] Tim Sanchez et al. "Spontaneous motion in hierarchically assembled active matter". *Nature*, 491.7424 (2012).

[152] Philippe Schwaller et al. "Molecular Transformer: A Model for Uncertainty-Calibrated Chemical Reaction Prediction". *ACS Cent. Sci.*, 5.9 (2019).

[153] Daniel S. Seara, Benjamin B. Machta, and Michael P. Murrell. "Irreversibility in dynamical phases and transitions". *Nat. Commun.*, 12 (2021).

[154] Alireza Seif, Mohammad Hafezi, and Christopher Jarzynski. "Machine learning the thermodynamic arrow of time". *Nat. Phys.*, 17 (2021).

[155] Udo Seifert. "Stochastic thermodynamics, fluctuation theorems and molecular machines". *Rep. Prog. Phys.*, 75.12 (2012).

[156] Ken Sekimoto. "Kinetic Characterization of Heat Bath and the Energetics of Thermal Ratchet Models". *J. Phys. Soc. Japan*, 66.5 (1997).

[157] Ken Sekimoto. "Langevin Equation and Thermodynamics". *Prog. Theor. Phys. Supp.*, 130 (1998).

[158] Suraj Shankar and M. Cristina Marchetti. "Hidden entropy production and work fluctuations in an ideal active gas". *Phys. Rev. E*, 98 (2018).

[159] Marina Soares e Silva et al. "Active multistage coarsening of actin networks driven by myosin motors". *Proc. Natl. Acad. Sci. USA*, 108.23 (2011).

[160] A. P. Solon et al. "Pressure is not a state function for generic active fluids". *Nat. Phys.*, 11.8 (2015).

[161] Alexandre P Solon et al. "Generalized thermodynamics of motility-induced phase separation: phase equilibria, Laplace pressure, and change of ensembles". *New J. Phys.*, 20.7 (2018).

[162] Alexandre P. Solon et al. "Generalized thermodynamics of phase equilibria in scalar active matter". *Phys. Rev. E*, 97.2 (2018).

[163] Alexandre P. Solon et al. "Pressure and Phase Equilibria in Interacting Active Brownian Spheres". *Phys. Rev. Lett.*, 114 (2015).

[164] Anton Souslov et al. “Topological Waves in Fluids with Odd Viscosity”. *Phys. Rev. Lett.*, 122 (2019).

[165] Thomas Speck, Alex Malins, and C. Patrick Royall. “First-Order Phase Transition in a Model Glass Former: Coupling of Local Structure and Dynamics”. *Phys. Rev. Lett.*, 109.19 (2012).

[166] Thomas Speck et al. “Dynamical mean-field theory and weakly non-linear analysis for the phase separation of active Brownian particles”. *J. Chem. Phys.*, 142.22 (2015).

[167] Thomas Speck et al. “Effective cahn-hilliard equation for the phase separation of active brownian particles”. *Phys. Rev. Lett.*, 112.21 (2014).

[168] Joakim Stenhammar et al. “Role of Correlations in the Collective Behavior of Microswimmer Suspensions”. *Phys. Rev. Lett.*, 119 (2017).

[169] Menachem Stern, Matthew B. Pinson, and Arvind Murugan. “The Complexity of Folding Self-Folding Origami”. *Phys. Rev. X*, 7 (2017).

[170] Grzegorz Szamel. “Mode-coupling theory for the steady-state dynamics of active Brownian particles”. *J. Chem. Phys.*, 150.12 (2019).

[171] Grzegorz Szamel. “Self-propelled particle in an external potential: Existence of an effective temperature”. *Phys. Rev. E*, 90 (2014).

[172] Grzegorz Szamel, Elijah Flenner, and Ludovic Berthier. “Glassy dynamics of athermal self-propelled particles: Computer simulations and a nonequilibrium microscopic theory”. *Phys. Rev. E*, 91.6 (2015).

[173] J. Tailleur and M. E. Cates. “Statistical Mechanics of Interacting Run-and-Tumble Bacteria”. *Phys. Rev. Lett.*, 100.21 (2008).

[174] Julien Tailleur and Jorge Kurchan. “Probing rare physical trajectories with Lyapunov weighted dynamics”. *Nat. Phys.*, 3.3 (2007).

[175] S. C. Takatori and J. F. Brady. “Towards a thermodynamics of active matter”. *Phys. Rev. E*, 91 (2015).

[176] S. C. Takatori, W. Yan, and J. F. Brady. “Swim Pressure: Stress Generation in Active Matter”. *Phys. Rev. Lett.*, 113 (2014).

[177] Nathaniel Thomas et al. “Tensor Field Networks: Rotation- and Translation-Equivariant Neural Networks for 3D Point Clouds”. *ArXiv e-prints* (2018).

[178] Nicolás Tizón-Escamilla, Vivien Lecomte, and Eric Bertin. "Effective driven dynamics for one-dimensional conditioned Langevin processes in the weak-noise limit". *J. Stat. Mech.*, 2019.1 (2019).

[179] Laura Tociu. *Github*.

[180] Laura Tociu et al. "How Dissipation Constrains Fluctuations in Nonequilibrium Liquids: Diffusion, Structure, and Biased Interactions". *Phys. Rev. X*, 9.4 (2019).

[181] Laura Tociu et al. "Inferring dissipation from static structure in active matter". *ArXiv e-prints* (2020).

[182] Hugo Touchette. "The large deviation approach to statistical mechanics". *Phys. Rep.*, 478.1 (2009).

[183] Shoichi Toyabe et al. "Nonequilibrium Energetics of a Single $\mathbf{F}_1$-ATPase Molecule". *Phys. Rev. Lett.*, 104 (2010).

[184] J. C. Tsai et al. "A chiral granular gas". *Phys. Rev. Lett.*, 94.21 (2005).

[185] Pelerine Tsobgni Nyawo and Hugo Touchette. "Large deviations of the current for driven periodic diffusions". *Phys. Rev. E*, 94 (2016).

[186] H. Turlier et al. "Equilibrium physics breakdown reveals the active nature of red blood cell flickering". *Nat. Phys.*, 12.513 (2016).

[187] Suriyanarayanan Vaikuntanathan and Phillip L. Geissler. "Putting water on a lattice: The importance of long wavelength density fluctuations in theories of hydrophobic and interfacial phenomena". *Phys. Rev. Lett.*, 112.2 (2014).

[188] Suriyanarayanan Vaikuntanathan et al. "Necessity of capillary modes in a minimal model of nanoscale hydrophobic solvation". *Proc. Natl. Acad. Sci. USA* (2016).

[189] Tamás Vicsek et al. "Novel Type of Phase Transition in a System of Self-Driven Particles". *Phys. Rev. Lett.*, 75 (1995).

[190] Teun Vissers, Alfons van Blaaderen, and Arnout Imhof. "Band formation in mixtures of oppositely charged colloids driven by an ac electric field." *Phys. Rev. Lett.*, 106.22 (2011).

[191] Fangbin Wang et al. "Non-equilibrium effect in the allosteric regulation of the bacterial flagellar switch". *Nat. Phys.*, 13.7 (2017).

[192] Siyang Wang and Marek W. Urban. "Self-healing polymers". *Nat. Rev. Mater.*, 5.8 (2020).

[193] John D. Weeks, David Chandler, and Hans C. Andersen. "Role of Repulsive Forces in Determining the Equilibrium Structure of Simple Liquids". *J. Chem. Phys.*, 54.12 (1971).

[194] René Wittmann et al. "Effective equilibrium states in the colored-noise model for active matter I. Pairwise forces in the Fox and unified colored noise approximations". *J. Stat. Mech.*, 2017.11 (2017).

[195] René Wittmann et al. "Effective equilibrium states in the colored-noise model for active matter II. A unified framework for phase equilibria, structure and mechanical properties". *J. Stat. Mech.*, 2017.11 (2017).

[196] Adam Wysocki and Hartmut Löwen. "Oscillatory driven colloidal binary mixtures: Axial segregation versus laning". *Phys. Rev. E*, 79.4 (2009).

[197] J. Yan, S. C. Bae, and S. Granick. "Rotating crystals of magnetic Janus colloids". *Soft Matter*, 11.1 (2015).

[198] R. Zakine et al. "Surface Tensions between Active Fluids and Solid Interfaces: Bare vs Dressed". *Phys. Rev. Lett.*, 124 (2020).

[199] Hongbo Zhao et al. "Learning the Physics of Pattern Formation from Images". *Phys. Rev. Lett.*, 124.6 (2020).

[200] Benjamin C. van Zuiden et al. "Spatiotemporal order and emergent edge currents in active spinner materials". *Proc. Natl. Acad. Sci. USA*, 113.46 (2016).