Variational sampling and optimal design of rare nonequilibrium molecular dynamics

by

Avishek Das

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Chemistry

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor David T. Limmer, Chair
Professor Phillip L. Geissler
Professor Ehud Altman

Spring 2022

Variational sampling and optimal design of rare nonequilibrium molecular dynamics

Copyright 2022
by
Avishek Das

Abstract

Variational sampling and optimal design of rare nonequilibrium molecular dynamics

by

Avishek Das

Doctor of Philosophy in Chemistry

University of California, Berkeley

Professor David T. Limmer, Chair

Nonequilibrium driving can independently tune the structure and dynamics of molecular and colloidal systems, resulting in novel assembly and transport phenomena that are not found in equilibrium. Exploring nonequilibrium dynamics through numerical simulations is crucial towards understanding the working mechanisms of functional inorganic or biological materials. However, behavior of dynamical observables in such systems at experimental timescales is often dominated by statistically rare fluctuations that are poorly sampled in simulated trajectories. Rare event sampling algorithms that assume a Boltzmann distribution of configurations and detailed balance for dynamics cannot be applied out of thermal equilibrium, thus limiting available techniques that can efficiently sample nonequilibrium rare events.

In this thesis we have developed novel variational algorithms for the sampling and design of rare nonequilibrium molecular dynamics trajectories by application of an optimized driving force. This approach relies on the equivalence of a trajectory ensemble conditioned on a rare event to occur to an ensemble driven with the optimal force where the rare event occurs typically. For systems with many interacting degrees of freedom, we numerically learn the optimal force within arbitrary basis sets by statistically estimating explicit gradients of trajectory probability. This method allows us to efficiently compute large deviation functions of dynamical observables in nonequilibrium steady-states, and to automate the inverse design of self-assembling colloids and molecular machines for desired structure and dynamics. We have finally augmented our approach with reinforcement learning techniques, resulting in a new paradigm to efficiently compute nonequilibrium reaction rates, Variational Path Sampling. Our approach of using optimized forces to improve sampling of nonequilibrium trajectories is versatile and can give access to rare reactive fluctuations and dynamical phases that cannot be sampled otherwise.

To

Ashok and Didi

Contents

Contents ii

1 Introduction 1
1.1 Nonequilibrium rare events in physics and chemistry 1
1.2 Langevin dynamics and trajectory ensembles 3
1.3 Steady-state rare events: large deviation theory 6
1.4 Barrier crossing events: reaction rate theory 8
1.5 Existing sampling techniques for rare nonequilibrium trajectories 11

2 Variational optimization of steady-state trajectories 14
2.1 Variational principle from Girsanov reweighting 15
2.2 Optimization algorithm with explicit gradients 17
2.3 Corrections of systematic errors 20
2.4 Current uctuations in non-interacting overdamped/underdamped system . . 23
2.5 Activity uctuations in interacting overdamped system 29
2.6 Current uctuations in interacting overdamped system 36
2.7 Conclusion 39

3 Inverse design of nonequilibrium colloidal self-assembly 41
3.1 Coarse-grained model for DNA-labeled colloids 43
3.2 Automated inverse design algorithm 45
3.3 Design principles for yields of rigid and nonrigid clusters 52
3.4 Design principles for yields of microphase separation 59
3.5 Design principles for reactive ux in steady-states 62
3.6 Optimal design principles for molecular machines 65
3.7 Conclusion 68

4 Reinforcement Learning of rare di usive dynamics 70
4.1 Trajectory ensemble Formalism 71
4.2 Gradient optimization for nite time constrained dynamics 73
4.3 Discrete timestep implementations of nite time algorithms 79
4.4 Rare uctuations in nite time 86

4.5 Alternative CGF estimates 95
4.6 Errors in gradient estimates 97
4.7 Gradient optimization for in nite time dynamics 98
4.8 Rare uctuations in the long time limit 101
4.9 Conclusion 107

5 Direct evaluation of rare events from variational path sampling 110
5.1 Variational principle for rate computation 111
5.2 Optimized control force for passive dimer 114
5.3 Active driving in dimer isomerization 117
5.4 Protocol for learning optimal force 119
5.5 Unbiased reactive events from VPS 122
5.6 Comparison with Forward Flux Sampling 123
5.7 Dissipative rate bounds 125
5.8 Nonequilibrium solvation structure 126
5.9 Conclusion 128

Bibliography 129

Acknowledgments

The past ve years have been simultaneously the most challenging and happiest in my life so far. I owe my identity as a person and a researcher today to the patience, love and support from many, many people who believed in me, inspired me and dreamed for me. I am grateful to be able to thank them here.

First and foremost, I will thank my adviser David Limmer without whom this dissertation would not exist. I have learnt most of my `Chandlerian statistical mechanics' from his brilliant instruction and from teaching and working with him. His endless patience and willingness to help with hands-on research, deep technical intuition and foresight, scienti c creativity, and relentless enthusiasm about all things chemistry have made working with him extraordinary fun. David has actively encouraged my exploration of truly fundamental questions in statistical mechanics by shielding me from non-academic logistical constraints and giving me a lot of independence, for which I am thankful. He has diligently taught me how to systematize the process of abstract scienti c inquiry into careful experimentation, analysis, inference and publication, in short, how to be a practicing scientist. I very much look forward to many more years of enjoyable scienti c discussions with David in the future.

Many other scientists and students at Berkeley and at other universities contributed greatly to my development as a researcher. I am thankful to my collaborators Naomi Ginsberg, Dmitri Talapin and Sam Teitelbaum for our many interesting conversations about designing and detecting the self-assembly of charged nanocrystals. I have learnt an incredible amount and had a lot of fun from discussing this problem with Christian Tanner, James Utterback, Vivian Wall, Igor Coropceanu and Josh Portner. I have also greatly enjoyed exploring the connections between reinforcement learning and importance sampling in my collaboration with Juan Garrahan and Dominic Rose. The wealth of knowledge I have gained from them on the statistical physics of machine learning paved the way for some of the most crucial developments in this dissertation. I am also grateful to Robert Jack and Hugo Touchette for encouragingly supporting my research on large deviation sampling in its early stages, engaging with it over the years and providing many valuable comments all through. Similarly I am indebted to Phillip Geissler and Eric Neuscamman for discussions on the physics of trajectory sampling and variational algorithms, and to Phill and Ehud Altman for a critical reading of this dissertation. Not only have I intellectually and scientifically bene ted from all my interactions with these leading researchers, but they have been inspirational role-models to me on how to practice fundamental science.

Getting to this stage today has taken many leaps of faith from past mentors who trusted and indulged my academic potential. I would like to thank my parents rst for bringing me up in an intellectual and inquisitive environment and for nurturing in me a love for books and knowledge. I am thankful for the memorable courses by my undergraduate professors E. D. Jemmis, Vijay Shenoy and D. D. Sarma, all of whom showed me how fascinating the study of `more is di erent' in the chemistry and physics of condensed matter could be. I also owe my sincere gratitude to Prof. Gopalan Rajaraman from IIT Bombay, with whom I did my rst theoretical chemistry project and fell in love with the power and reductionist

sophistication that computational methods provide. He was an excellent mentor who granted me research independence even while I was an undergraduate, and with whom I felt the thrill of discovering something fundamental and novel for the rst time.

Aside from research, teaching has been a large component of my PhD at Berkeley and has been one of my most enjoyable academic vocations. I would like to thank Michelle Douskey for a very interesting course on chemistry pedagogy that impacted me a lot as a teacher. I must thank Birgitta Whaley for a memorable co-instructing experience in graduate quantum mechanics that led to my full development as a teacher in physical chemistry. I am fortunate to have had extremely kind and talented students over the years who have since become close friends and collaborators. The desire to continue teaching such excellent students even after ful ling my teaching requirement is what partially contributed to my involvement in the development of the UC Berkeley Math Bootcamp, which has been incredibly rewarding and ful lling. I would like to thank my co-instructors at the bootcamp, Rachel Clune, Dipti Jasrasaria, Orion Cohen and Elliot Rossomme, for having made the bootcamp such a special experience. I must also thank Anne Baranger and Brice Yates for guiding us through the exploration of learning outcomes of the bootcamp and for the many insightful discussions on equity and inclusion in the College of Chemistry. Teaching at UC Berkeley is a gift that has kept on giving and I am con dent that the many professional and personal friendships I made with my ex-students will last long.

Ful lling friendships are what have helped me cultivate my professional and personal identity and thrive through the di cult times in my doctoral journey. I am incredibly fortunate to have had interacted with so many wonderful friends who have time and again surprised me with their unending love and support. My biggest con dantes on the ups and downs of graduate school have been the past and present members of the Limmer group. I would like to thank Laura Scal , who was my rst friend in the group, and Addison Schile, who taught me how to make nice presentations. I would like to thank Trevor GrandPre, Chloe Gao and Amael Obliger for teaching me a lot of statistical mechanics in my rst few years. I thank Frank Gao for having been a rock of emotional support for me within the group. I must thank Ben Kuznets-Speck and Aditya Singh for the many valuable research conversations we have had and our continuing collaborations. I would like to thank Mirza Galib, Sam Niblett, Nirvaan Khera, Yoonjae Park, Michelle Anderson, Tom Fay, Phillip Helms, Anthony Poggioli, Jin Moon, Sam Oaks-Leaf, Artur Lyssenko and Joseph Slivka, with each of whom I have spent many fun, memorable evenings that I always looked forward to from a stressful day at work. I am thankful to the broader Pitzer center community as well for the solidarity and camaraderie that makes spending time here so rewarding. I thank Dipti Jasrasaria, my best friend in my cohort and a wonderful and generous human being I am fortunate to know. I thank my friends Arpan, Subhayan, Sridevi, Shriya, Naren, Alistair, Ananth, Sabyasachi, Milind and Sarban for being my lifeline through this journey, for making the pandemic tolerable with our numerous Zoom hangouts, and for creating many beautiful memories with me across the country. I have indescribable gratitude for Maria Sutter, who has helped me heal from chronic abuse and trauma, and has helped me for the rst time hope for having an actually happy life for myself. I thank Jane Scherr for having been a wonderful

host and housemate during my time here in Berkeley, and for being such an inspirational role-model for identifying what the important things in life are. I thank my comrades in our academic student workers' unions UAW-2865 and UAW-SRU for having ensured strong workplace contracts that have given me the luxury of being a ful lled researcher in the rst place. And I am thankful for the many wonderful queer friends I have made in the bay area, who have had enormous roles in shaping me as the human being I am today.

Finally, I would like to thank the most important people in my life, who have showed me the meaning of unconditional love. My sister Jayita and brother-in-law Kaushik have been my pillars of support through the many arduous times we have together gone through. Words fail to describe how lucky I feel to have had them as supportive family. And I thank Ashok, my partner and my best friend, for being the sweetest, funniest, most extraordinary person I have ever known and for being there for me every single day through grad school. Some of my happiest moments in the past few years I have spent with him, either virtually or in-person. I am incredibly excited to see both of us grow together in the coming years.

All of the above still feel like massive understatements on the amount of help and support I have actually had from people. Kindness, love and inspiration from friends and strangers have sustained me through graduate school that resulted in this dissertation. I sincerely hope I am able to pay it forward.

Chapter 1

Introduction

1.1 Nonequilibrium rare events in physics and chemistry

The past few decades have seen rapid growth in the capability of computer simulations to address questions about the microscopic dynamics of diverse chemical and biological systems across multiple length and time scales. Computer simulations have been used to elucidate reaction pathways for physical processes like dissolution of a solute [1], for deciphering reaction mechanisms of atomic and molecular rearrangements in solvents [2], to understand energy, charge and mass transport through electrochemical cells, micro uidic devices and crystalline and amorphous solids [3, 4, 5, 6, 7], and to explore assembly and ordering phenomena in inorganic and biological materials [8, 9, 10]. Even in classical limits where the e ects of quantum coherence are negligible, dynamics in a condensed phase is often collective, involving interactions of a given system of interest with thermal, mechanical or chemical baths that can inject noise and memory into the system's dynamics [11]. The complexity of such dynamics at the nanoscale makes any analytical treatment prohibitive, and hence conclusions about reaction mechanisms, transport properties and inverse design of materials are often drawn from computer simulations. Nanoscale dynamical information obtained from computer simulations are also extensively used to build coarse-grained e ective models of materials, allowing a reduced description at longer length and time scales [12]. But inferring such high-level description from the microscopic dynamics of molecules faces the crucial challenge of sampling rare dynamical events.

Rare events in molecular simulations often give rise to complicated nonlinear dynamics that are of broad interest [13]. Molecular dynamics simulations involve integrating a known equation of motion, such as Newton's equations, in the presence of many degrees of freedom. For temporal integration we choose a quadrature timestep that is usually smaller than the fastest modes of motion in the system, that of atomic translations and molecular rotations in the sub-picosecond scale. However, condensed phase environments stabilize multiple metastable states in molecules, and crucial dynamics like the folding of a protein from

its secondary to tertiary structure, or the nucleation of a stable phase from a metastable one, occur at timescales many orders of magnitude longer than molecular motion [14, 15]. This is not due to the progress along these reaction coordinates being inherently slow, but arises from the presence of kinetic bottlenecks with narrow funnels of entry that are only rarely encountered. As a result, these dynamics are only observed in molecular simulations with a low probability even though they drastically alter the properties of the system on experimental timescales. Developing special numerical techniques to sample such rare but interesting events with correct statistical likelihood in molecular dynamics simulations has occupied much of the e orts of statistical physicists working in physical chemistry problems. A majority of these techniques assume the system to be in thermal equilibrium, which simplies the uctuations of observables due to the con gurational distribution being Boltzmann. This however limits their applicability to driven systems that are ubiquitously encountered and exploited in functional chemistry and biology.

A system kept away from thermal equilibrium by hindering its relaxation or through a continuous supply of energy is subject to fewer physical constraints than one evolving within an equilibrium state. As a consequence, the application of external forces or the internal consumption of energy can produce structures and responses without equilibrium equivalent [16, 17, 18]. Driving a system dissipatively can decouple its dynamical constraints from its structure, and thus stabilize otherwise transient states, or generate dynamical uxes between otherwise non-reactive structures. Nonequilibrium driving can assist molecular motors to transduce energy and move ions against thermodynamic gradients [19], stabilize clusters of motile bacteria without attractive interactions [20], and overcome thermodynamic limits on the selectivity of assembly and control [21, 22]. A majority of manufacturing processes in industrial chemistry, synthesis protocols for nanomaterials, geophysics and geochemistry of climate uctuations, and biological reactions that govern living systems occur in nonequilibrium conditions [23, 24, 25, 26]. Advances in the theory and modeling of nonequilibrium steady-states, borrowing tools from probability theory and stochastic calculus [19, 27, 28], have resulted in an increased interest in trying to understand the behavior in systems out of equilibrium and leverage their versatility to design new functional materials [20, 29, 30, 31, 22, 32, 33]. Thermodynamic partition function formalism has been generalized to ensembles of trajectories via large deviation theory [34] and transport theories have been developed to quantify linear and higher-order response functions to nonequilibrium perturbations [35, 36]. The second law of thermodynamics has also been generalized to far-from-equilibrium dynamics through uctuation theorems that relate uctuations of nonequilibrium structure and dynamics to thermodynamic quantities like entropy or work [37, 38, 39], and thermodynamic uncertainty relations that always bound current uctuations [40, 41].

Despite progress in these theoretical directions, quantifying emergent nonequilibrium behavior with computer simulations is currently hampered by the lack of robust tools to sample the rare uctuations required to estimate response functions, overcome kinetic bottlenecks, and reach the timescales of experimental relevance. Commonly used numerical methods for sampling nonequilibrium dynamics rely on stratifying trajectory ensembles to build up rare events stepwise, but do not apply additional importance sampling forces that can gener-

ate inherently uncorrelated statistical realizations. In the present work we have developed variational techniques to sample and design rare dynamical uctuations for a generic class of stochastic systems that violate detailed balance. Our approach relies on optimizing an additional driving force that reproduces the unbiased rare uctuations of the unperturbed system. Our numerical methods have allowed us to learn this force in complex molecular and colloidal systems within arbitrary basis sets and e ciently extract the statistics and mechanism of nonequilibrium rare events. Chapter 1 brie y discusses the trajectory ensemble formalism to quantify rare events in transient and steady-state trajectories, and the backgrounds of large deviations and reaction rate theory that our formalism is based on. In Chapter 2 we develop a variational algorithm for the use of optimized control forces to sample rare current and activity uctuations of particles in a nonequilibrium steady-state on a one-dimensional model potential. Chapter 3 applies this algorithm towards the automated discovery of inverse design principles for DNA-labeled colloids in a shear ow. In Chapter 4 we then develop more e cient optimization algorithms based on reinforcement learning tools to solve nite and in nite duration variational problems. Finally, Chapter 5 applies this approach to develop a new algorithm to compute reaction rates in far-from-equilibrium systems like active matter, that we call Variational Path Sampling (VPS).

1.2 Langevin dynamics and trajectory ensembles

We aim to temporally integrate the equations of motion at molecular and colloidal scales to simulate rare dynamics of systems of interest. Conventional molecular dynamics in an isolated system is described by Newton's equations of motion that are integrated with symplectic schemes that conserve energy, like the velocity Verlet algorithm [42]. However, driving a system with external forces performs work on it, and in order to relax into a steady-state, the excess energy must be discarded by the system as dissipation into a bath. Driven systems are thus often modeled as open systems where energy is not conserved, and where there is on average a nite dissipative ow between the system and the bath, resulting in nonequilibrium. Moreover, dynamics in a condensed phase can often be simpli ed by interpreting the interesting part of the system to be open and in contact with the rest of the system as a thermal bath, which only a ects the interesting part through frictional forces and Brownian noise. This dynamics could be interpreted as a method to thermostat the system at a xed temperature, or as an approximation to the noisy di usive dynamics encountered by molecules and colloids suspended in liquid solvents or in contact with a large number of thermalized vibrational excitations. We will describe our nonequilibrium variational algorithms only in the context of this Langevin dynamics. Generalization to other kinds of stochastic dynamics is straightforward as long as every trajectory has a closed form nonzero probability.

We consider dynamics described by a Langevin equation of the form [11],

$$\dot{\underline{r}}_i(t) = v_i(t) \tag{1.1}$$

$$m_i \dot{\underline{v}}_i(t) = F_i[r^N(t)] \quad {}_i v_i(t) + \quad {}_i(t) \tag{1.2}$$

where $\dot{\underline{r}}_i$ and $\dot{\underline{v}}_i$ are the rate of change of the i-th particle's position and velocity, $_i$ is the corresponding friction coe cient, and $F_i[r^N(t)]$ is the sum of all conservative and noncon-servative forces exerted on the i-th particle that depends on the full con guration of the N-particle system, r^N. The nal term, $_i(t)$, represents a thermal bath-induced white-noise, as a Gaussian random number with zero mea $h_{,i}(t)i = 0$, and delta-correlated in time,

$$h_i(t)\ _j(t^0)i = 2\ _i k_B T\ _{ij} \quad (t \quad t^0) \tag{1.3}$$

for component (;), with $k_B T$ being Boltzmann's constant times the temperature. The relation between the variance of the Gaussian white noise with the friction coe cient of the bath derives from the second uctuation-dissipation theorem. Given the approximate diameter of a particle is and it is suspended in the bath with friction , the time taken to di use its own length is given by $t = {}^2/D$, where $D_i = k_B T/$ is the di usion constant. In most cases we will consider the high-friction limit of the Langevin equation, formally taken as $m/t \ ! \ 0$, where the particles lose their inertia and instead are governed by the overdamped Langevin equation

$$_i \dot{\underline{r}}_i(t) = F_i[r^N(t)] + \quad {}_i(t) \tag{1.4}$$

The overdamped equation can model the dynamics of a wide range of colloidal systems including nanocrystals, proteins and polymers in a condensed phase with a su ciently large bath. On the other hand, describing inertial dynamics like vibrations and hydrodynamic e ects require using the full Langevin equation in an underdamped regime. We note that the underdamped equation could be considered formally to be a generalized overdamped equation for all dynamical coordinates, in the joint phase space of position and velocity, with the latter being in contact with a bath at zero friction. Hence we will do most discussions of numerical techniques developed here in the context of the overdamped Langevin equations without loss of generality. Similarly, generalizations to systems with state-dependent di usion matrices are straightforward from our formalism.

The Langevin equation is simulated by discretizing time into integration timesteps. Given $\dot{\underline{r}}(t)dt$ is a stochastic di erential, care must be taken with respect to discretizing time-integrals and connecting them to physical observables. We will follow in most cases the Ito formalism, where the integration of the dynamics is interpreted as

$$_i r_i(t + \quad t) = \quad {}_i r_i(t) + F_i[r^N(t)] \quad t + \sqrt{2\ _i k_B T \quad t} N_i^t \tag{1.5}$$

where N_i^t is a vector of independent Gaussian random numbers with zero mean and unit variance. Here each step taken in is fully de ned by the forces evaluated at the previous

value of r. This should be contrasted with the Stratonovich convention,

$$ {}_i r_i(t + \quad t) = \quad {}_i r_i(t) + F_i \quad \frac{r^N(t) + r^N(t + \quad t)}{2} \quad t + \sqrt{2 \; {}_i k_B T \quad t} N_i^t; \qquad (1.6) $$

where the dynamical equation has to be solved self-consistently at every timestep. Choosing either of the two conventions results in di erent dynamics if the pre-factor of the Gaussian noise depends on the con gurations. The Stratonovich formalism makes di erential calculus easier as the chain rule of di erentiation holds unchanged, while the Ito formalism simpli es integral calculus as the change of $f(t)$ in every step is uncorrelated from the noise N_i^t. In our work we will use the Ito formalism for integrating the Langevin equation as well as integrating all stochastic integrals for physical observables. With this convention, the time evolution of the probability distribution of the con gurations $\{r^N\}$, denoted as $P(r^N; t)$, is described in forward direction by the Fokker-Planck equation [43]

$$ \frac{@P(r^N; t)}{@t} = \sum_i \quad r \; {}_i \quad F_i[r^N(t)]P(r^N; t) \; + D_i r \; {}_i^2 P(r^N; t) \qquad (1.7) $$

This equation can be derived from the Langevin dynamics by a Kramers-Moyal expansion, a moment expansion of the transition probabilities due to the stochastic noise [43]. The Fokker-Planck equation, also known as the forward Kolmogorov equation, along with boundary conditions describes the forward time evolution of the full probability distribution of the stochastic dynamics arising from Eq. 1.4. In principle, the Fokker-Planck equation formally predicts the probability of all rare dynamics that the Langevin dynamics can generate. In practice, it is often not possible to obtain analytical or even numerical solutions from the Fokker-Planck equation when the system size is large. In many-body systems with identical particles, the Fokker-Planck equation is often hierarchically coarse-grained to describe e ective evolution of one or two-body densities [44].

Dynamical events of interest simulated with the Langevin equation also manifest in the ensemble of trajectories. We will denote a trajectory as a stochastic object $X(t_f)$ de ned as a sequence of con gurations generated from a Langevin dynamics of duration t_f, $X(t_f)$ $\{r^N(0); r^N(\quad t); r^N(2 \quad t); : : : ; r^N(t_f)\}$. An ensemble generated by Langevin trajectories with all possible statistical realizations of the Gaussian noise will be referred to as the trajectory ensemble associated with the dynamics. For any particular trajectory with a pre-speci ed initial condition, every step has been taken with a Gaussian probability $\exp(\quad {}_i^2/2)/\sqrt{2}$. We write the total trajectory probability in the ne discretization limit as [45]

$$ P[X] / \quad p(r^N(0)) \exp \left(\int_0^{t_f} dt \sum_i \frac{(\; {}_i \dot{r}_i \quad F_i)^2}{4 \; {}_i k_B T} \right) \qquad (1.8) $$

where the normalization constant is formally in nite. The exponent is referred to as the stochastic action associated with each trajectory, in analogy with the path-integral description of Schrödinger's wave mechanics. All three levels of descriptions of the stochastic dynamics of our system of interest, through con gurations, densities and trajectories as in

Equations 1.4, 1.7 and 1.8, are fully equivalent, and we will often draw connections between the three to fully characterize the physical and chemical mechanisms of rare fluctuations. However, the fact that the full trajectory probability has a closed form expression in terms of the applied forces, even in cases where the forces are nonconservative and the dynamics is out of equilibrium, makes our approach of reweighting trajectories with an optimal control force viable. Hence we will be using trajectory ensembles constructed from molecular dynamics simulations to apply our variational algorithms.

1.3 Steady-state rare events: large deviation theory

The statistics of persistent rare events in nonequilibrium steady-states can be analyzed using large deviation theory. Large deviation theory is a branch of probability dealing with the exponential scaling of large fluctuations of the sample mean or empirical mean of a large number of stochastic observables. This can be thought of as a generalization of the central limit theorem in probability towards describing large fluctuations of sample means, or large deviations. The central components in large deviation theory are the rate function, which describes the log-likelihood of the sample mean, and the scaled cumulant generating function for its probability distribution [46]. These two are jointly referred to as large deviation functions and, in the case of the probability distribution being convex and differentiable, are connected through the Legendre-Fenchel transform, a generalization of the Legendre transform used for thermodynamic free energies. In recent years, the classical theory for constructing thermodynamic free energies from statistical mechanical partition functions has been reinterpreted with the language of large deviation theory, with the thermodynamic system size being the large deviation limit [34]. This has allowed interpreting steady-state trajectory ensembles to also be in the (temporal) large deviation limit, and has provided large deviation functions as dynamical free energies for specific observables of interest, encoding stability and response in arbitrarily far-from-equilibrium steady-states. In this section we will review the large deviation formalism for trajectory ensembles and its use to obtain statistics of rare events.

For ease of notation we will denote all coordinates of the system $\mathbf{r}^N(t)$, sometimes simply as $\mathbf{r}$. For a specific trajectory, $X(t_f) = \{\mathbf{r}(0), \ldots, \mathbf{r}(t_f)\}$ spanning an observation time t_f, we are interested in fluctuations of time-averaged dynamical observables A_{t_f} of the form

$$A_{t_f}[X(t_f)] = \frac{1}{t_f}\int_0^{t_f} dt\, f[\mathbf{r}(t)] + \frac{1}{t_f}\int_0^{t_f} dt\, \mathbf{g}[\mathbf{r}(t)] \cdot \dot{\mathbf{r}}(t) \tag{1.9}$$

where f is a scalar function and $\mathbf{g}$ is a vector function with components g_i, with the second term being evaluated in the Ito sense [47]. Many physically relevant observables like the particle density, particle current, and entropy production can all be expressed in this trajectory averaged form. We will be interested in the statistics of rare time-extensive fluctuations, the large deviations, of A_{t_f} in the long time limit, $t_f \to \infty$.

Figure 1.1: Large deviation principle for observable A_t arising from Brownian diffusion. a) Trajectories of the time-integrated observable tA_t starting from the same point diverge in time with a variance proportional to t over times longer than intrinsic correlation times of the system. c) Trajectories of the time-averaged analogue A_t of the same observable. The probabilities concentrate around the mean value with asymptotic scaling as exponential in time multiplied by the negative of the rate function $I(A_t)$.

We assume that in the long time limit, the probability distribution of A_{t_f} satisfies a large deviation principle, with a rate function, or log likelihood, $I(A)$, defined by [34]

$$I(A) = \lim_{t_f \to \infty} \frac{1}{t_f} \ln \langle \delta(A - A_{t_f}[X(t_f)]) \rangle \tag{1.10}$$

where the angular brackets denote a trajectory average

$$\langle \delta(A - A_{t_f}[X(t_f)]) \rangle = \int \mathcal{D}[X(t_f)] \, \delta(A - A_{t_f}[X(t_f)]) P[X(t_f)] \tag{1.11}$$

and $P[X(t_f)]$ denotes the path probability associated with trajectory $X(t_f)$. We will consider thermalized systems that have exponentially decaying correlation functions and thus are expected to obey the large deviation principle. The large deviation principle for time-integrated and time-averaged observables in the long time limit is illustrated in Fig. 1.1.

The long time behavior of A_{t_f} can also be characterized by its scaled cumulant generating function (SCGF), defined as

$$\lambda(s) = \lim_{t_f \to \infty} \frac{1}{t_f} \ln \left\langle e^{s t_f A_{t_f}} \right\rangle \tag{1.12}$$

where s is a counting parameter conjugate to A_{t_f}, and denotes the extent of biasing or tilting on the typical value of A_{t_f}. Larger positive or negative values of s probe rarer fluctuations.

This is clear by noting that the derivatives of (s) report on the cumulants of A_{t_f}, such as

$$\frac{d(s)}{ds} = \langle A_{t_f} \rangle; \text{ and } \quad \frac{d^2(s)}{ds^2} = t_f \langle (A_{t_f})^2 \rangle \tag{1.13}$$

where $A = A - \langle A \rangle$ are the fluctuations of A. We refer to the rate function, $I(A)$, and the SCGF, (s), collectively as the large deviation functions. When the rate function is convex, it can be obtained from the SCGF using a Legendre-Fenchel transform

$$I(A) = \inf_s [sA - (s)] \tag{1.14}$$

where inf refers to an infimum taken over all possible values of s. Together, the large deviation functions encode stability of dynamical phases in nonequilibrium systems [48, 49, 50, 51, 52], characterize complex dynamical behavior [53, 54, 55, 56, 57] and describe linear and nonlinear multivariate response of nanoscale systems to nonequilibrium perturbations [58, 36, 59, 60, 61]. The SCGF defined in the trajectory ensemble picture can also be derived from the eigenspectrum of a tilted backward Fokker-Planck propagator defined as

$$\begin{aligned} L_s = & -sf + \sum_i \frac{h_s}{2}(r_i - g_i)g_i + F_i \cdot (r_i - sg_i) \\ & + \gamma_i k_B T(r_i^2 - s(r_i \cdot g_i) - 2sg_i \cdot r_i + s^2 g_i^2) \end{aligned} \tag{1.15}$$

where we have suppressed the arguments F_i, g_i, and f for compactness. This operator can be derived similar to the Fokker-Planck equation from the Kramers-Moyal expansion for infinitesimal propagation of $\langle \exp(-st_f A_f) \rangle$ with increasing t_f [62]. We have adopted the form of the operator used in [63] towards the Ito definition of A_{t_f} rather than Stratonovich, by adding the first term in the square brackets in Eq. 1.15. This operator satisfies an eigenvalue equation

$$L_s \, {}_s(r^N) = (s) \, {}_s(r^N) \tag{1.16}$$

where (s) and ${}_s(r^N)$ are respectively the largest real eigenvalue and corresponding right eigenvector of L_s, which follows from the Perron-Frobenius theorem and the long time limit of the SCGF. Constructing and solving this eigenvalue equation in a chosen functional basis for ${}_s(r^N)$ is an option to obtain the SCGF (s), however, the dimension of the solution space scales exponentially with the system size, rendering this approach impractical in many-body systems.

1.4 Barrier crossing events: reaction rate theory

Interesting chemical events in liquid and solid phase involving transduction of energy from one class of dynamical modes in a molecule to another often happen when a trajectory goes over a free energy barrier that separates long-lived metastable states. Rate constants of such rare events in equilibrium depend on the free energy cost of being at the top of

the barrier as well as the dynamical flux contributing to such a barrier crossing. One of the fundamental approaches to quantify reaction rates in equilibrium systems is transition state theory that assumes configurations at the top of the barrier to be in quasi-equilibrium with the rest of the system, thus decoupling the dynamics from the Boltzmann energetics. In reality, this approximation leads to an overestimation of the rate, as many trajectories that cross the barrier still recross back to the reactant basin in a correlated fashion to the configurational statistics. Thus a full treatment for exact reaction rate computation in even equilibrium system requires a trajectory ensemble formalism [13, 64]. Additionally, in nonequilibrium reactions, such as the nucleation of a liquid phase in repulsive motile particles [65], in the shear induced unfolding of proteins in blood vessels [66] and in rare switching events in biochemical networks [67], the existence of a free energy as a state-function is not guaranteed, and dynamics is inherently coupled to structural fluctuations. In the limit of low temperature compared to the barrier-height, Freidlin-Wentzell theory approximates barrier-crossing trajectories to be comprised of a long sequence of infinitesimal noise fluctuations along a progress coordinate, thus mapping the fluctuation into a large deviation and yielding corresponding rate estimates [68, 69]. However, for systems with diffusive degrees of freedom at ambient temperature, the overlap between the problem of sampling short barrier-crossing trajectories of finite duration and that of sampling large deviations in steady-state trajectories of infinite duration remains unexplored and the tools used in either disciplines are mostly disparate.

In this section we will discuss how to compute reaction rates from ensembles of trajectories. When we want to sample reactive trajectories that cross a free energy barrier transiently in time but otherwise follow typical steady-state statistics, either in or out of equilibrium, we want to restrict our trajectory ensemble to have a finite duration t_f, but initialized from an ensemble of steady-state configurations $\{r^N(0)\}$. We will denote the initial and final basins in configuration space as A and B, and the rate constant of going from A to B as k. In case of the reaction being rare on simulation timescales t_f, usually is in the range $\tau < t_f \ll 1/k$, where τ is the timescale of relaxation within the initial metastable basin or that of the typical transition path duration, while $1/k$ is the typical waiting time for the reaction to happen spontaneously. This separation of timescales is usually associated with a large gap in the eigenspectrum of the Fokker-Planck operator governing the dynamics in the system. We have schematically illustrated this for a general order parameter q in Fig. 1.2(a). We also define indicator functions $h_A(t)$ and $h_B(t)$ at any given time t as

$$h_{A(B)}[r^N(t)] = \begin{cases} 1 & \text{if } r^N(t) \in A(B) \\ 0 & \text{else} \end{cases} \tag{1.17}$$

The rate constant k can be computed as the reactive flux of trajectories into B conditioned on starting from A,

$$k = \frac{d}{dt_f}\langle h_B(t_f)|h_A(0)\rangle = \frac{d}{dt_f}\frac{\langle h_A(0)h_B(t_f)\rangle}{\langle h_A\rangle} \tag{1.18}$$

Figure 1.2: Reaction rate computation from side-side correlation function. a) Separation of timescales between the relaxation time within each metastable well and the typical waiting time for barrier-crossing. Blue line is a schematic illustration of the proportion of time spent by a barrier-crossing trajectory in the wells and on the barrier(Inset) Time series of order parameter q. b) Linear regime of the side-side correlation function, with the slope given by the reaction rate.

where the averages are computed over trajectories $x(t_f)$ and the average in the denominator is time-independent in a steady-state. Eq. 1.18 for the reaction rate yields a time-independent constant for all times that fall between the separated molecular and reactive timescales, and so the side-side correlation function $\langle h_B(t) h_A(0)\rangle$ is linear in t in that regime. When the separation of timescales is large, the o set to the correlation function from the initial nonlinear portion is small enough, and the rate constant can be computed from a simple time-scaling [70], as

$$k \quad \frac{\langle h_B(t_f) h_A(0)\rangle}{t_f} = \frac{1}{t_f}\frac{Z_{AB}(t_f)}{Z_A} \tag{1.19}$$

where $Z_{AB}(t)$ and Z_A are de ned as trajectory partition functions $Z_{AB}(t) = \langle h_A(0) h_B(t)\rangle$ and $Z_A = \langle h_A\rangle$. This linear scaling regime is schematically demonstrated in Fig. 1.2 (b). Reaction rate constants can thus be computed by counting the number of reactive trajectories proportional to all trajectories.

The side-side correlation is the central observable of interest for all sampling techniques for reactive trajectories. Sometimes the computation of the side-side correlation is done in two steps,

$$k_{AB} = \langle h_B(t) h_A(0)\rangle = \frac{\langle h_{A+}(t) h_A(0)\rangle}{\langle h_A\rangle}\langle h_B(t) h_{A+}(t)\rangle \tag{1.20}$$

where the rst term in the product is the escape ux from the basin A towards B, with $\dot{h}(t)$ denoting the time-derivative of the indicator $h(t)$ when some in nitesimal time t has elapsed after last being in A, and the second term in the product is the probability of reaching B given a trajectory has escaped towards B but is still in the vicinity of A. The escape ux and the conditional reactive probabilties can be evaluated separately and then combined to produce the reactive ux, i.e., the reaction rate.

The growth statistics of the side-side correlation produces only the rate constant, but does not by itself contain information about the reaction mechanism. A microscopic quantity that encodes the reaction mechanism is the committor function, de ned as the probability of reaching target state B after some transient time $_z$ when started from system con gurations $\{r^N\}$, as

$$p(r^N; \ _z) = \frac{h(r^N(0) \quad r^N)h_B(\ _z)i}{h(r^N(0) \quad r^N)i} \tag{1.21}$$

where angular brackets again mean a trajectory average [70, 64]. We see that the committor function is just a generalization of a side-side correlation function for starting from all coordinates instead of from just the initial metastable basin. The transient time $_z$ is comparable to the typical time for natural transition paths, i.e., the time taken between the last escape from A and the rst subsequent entry into B. Since trajectories started from most regions in con guration space do not react within this short time, the committor will be mostly either 0 or 1, but for only a few special con gurations that connect A to B along the correct reaction coordinate, the committor will smoothly change in value, taking a value of 0.5 at the exact set of transition states. Thus the committor function encodes the intrinsic mechanism of the rare event by identifying the dynamical bottleneck. Regions of con guration space having di erent values of the committor function are said to be partitioned by a separatrix, another name for an iso-committor surface. The direction of the gradient of the committor function, comprising of normals to consecutive separatrix surfaces, can be identi ed as the reaction coordinate, which is the natural coordinate on which the reaction proceeds. Knowing the reaction coordinate will equip us to tune the exact reaction mechanism microscopically and thus a key goal of enhanced sampling algorithms is to learn a committor surface from harvested reactive trajectories.

1.5 Existing sampling techniques for rare nonequilibrium trajectories

Enhanced sampling methods within equilibrium ensembles are standard tools that enable the determination of phase diagrams and the calculation of rates of rare events, through the evaluation of equilibrium free energies and rare dynamics that obeys detailed balance. The methods typically rely on either enhancing the sampling of the tails of a Boltzmann distribution by adding a suitable umbrella potential [42, 71], or by resampling parts of con gurational

and trajectory ensembles to construct rare fluctuations [72], or both [73]. Nonequilibrium systems pose additional complications over equilibrium ones in terms of sampling both rare structures and dynamics. Nonequilibrium structural configurations are not Boltzmann distributed and the existence of free energies as a functional of only density-like variables is not guaranteed, so methods analogous to umbrella sampling in configuration space have not been possible. Though some recent approaches discuss candidate theories for quantifying the amplification of reaction rates with nonequilibrium driving, they feature bounds instead of exact reweighting relations, and involve dynamical observables like heat dissipation rather than thermodynamic state-functions like energy or entropy [74, 75]. Nonequilibrium dynamics on the other hand breaks detailed balance, and it is not possible to construct a rare barrier-crossing trajectory by propagating a transition state configuration forward and backward in time. Instead, most nonequilibrium importance sampling methods for barrier-crossing trajectories start with a distribution of initial configurations and push the trajectory ensemble up the barrier in a forward ratchet-like fashion, performing nonequilibrium metawork in the space of the trajectories [76]. We have discussed below some of these trajectory sampling methods for nonequilibrium large deviations and reactive events.

The dynamical analogues of equilibrium free energies for nonequilibrium steady-states are large deviation functions that describe the likelihood of fluctuations of time-averaged observables and encode the response and stability of the system. The aim of dynamical importance sampling methods for nonequilibrium steady-states is thus to estimate large deviation functions from a trajectory ensemble. As large deviation functions like the SCGF contain exponentially rare statistics, the tails of the probability distribution of the observable of interest must be sampled in order to have an accurate estimate. Efforts to compute large deviation functions in systems with many degrees of freedom have largely been restricted to Monte Carlo based approaches like cloning [77, 78] and list-based algorithms [79]. Most current algorithms scale exponentially in computational effort the further the rare fluctuation is from the mean behavior, as apart from stratification or population dynamics, most do not employ additional importance sampling [80, 81, 82, 83]. Recent work adding control forces to importance sample trajectory based Monte Carlo has demonstrated that even an approximate force can greatly improve the efficiency of Monte Carlo methods in estimating large deviation functions [84, 85, 86]. Consequently there has been much work to find approximate control forces analytically or through empirical arguments in both lattice-based and continuous systems [87, 88, 89] and several iterative effective force optimization techniques have been proposed with varying levels of generality or accuracy [90, 91, 92, 93]. The control forces in general can have many-body components in interacting particle systems [94, 89], can be long-ranged in systems with dynamical phase transitions [95], and can stabilize otherwise metastable states [96]. However, many-body control forces have been derived only from analytical approximations in simple limits, and efficient numerical techniques to learn many-body optimal forces in systems of chemical relevance is currently lacking.

In contrast, nonequilibrium rate calculation has been entirely reliant on trajectory stratification and Monte-Carlo resampling based approaches. Most such techniques build on the classic Weighted Ensemble method introduced by Huber and Kim [97] and subsequently re-

ned and applied to equilibrium biological problems like protein folding and protein-ligand binding by Zuckerman, Chong and others [98]. The original Weighted Ensemble algorithm is applicable for nonequilibrium dynamics but not very e cient for generating reactive trajectories. The Transition Path Sampling algorithm [70, 99], widely used to compute rates of molecular isomerization and dissociation in solvents at equilibrium, can be applied to nonequilibrium systems by either restricting the set of moves that require detailed balance [81], or by directly accounting for the change in nonequilibrium path action by noise sampling [100]. The most widely used method for computing reaction rates in stochastic nonequilibrium systems is Forward Flux Sampling, that starts with the forward ux of escaping trajectories from the reactant basin and generates reactive trajectories stepwise by dividing the progress coordinate into multiple bins [101, 102]. A related approach is nonequilibrium umbrella sampling, that keeps track of the entry and exit ux of bins in con guration space in order to output the nonequilibrium steady-state distribution [103, 104]. The application of control forces to enhance the sampling of reactive events has been extremely limited and been con ned only to equilibrium systems. Recent works by Rotsko and Vanden-Eijnden recast the rare event sampling as a variational learning of an equilibrium committor surface [105, 106], however, connections have not been made between the committor function and the optimal force. Other approaches that use an importance sampling potential have explicitly used Boltzmann statistics, and have only seen moderate improvements in sampling statistics on adding suboptimal control potentials [107, 108]. However, the exact optimal control force is formally known to generate uncorrelated statistical realizations of rare trajectories directly [109], thus making the rate computationdirect without need of additional importance sampling, in arbitrarily far-from-equilibrium systems. Hence there exists a dearth of e cient numerical methods to compute the optimal control force within conveniently chosen basis sets for evaluating nonequilibrium reactive events in complex molecular and colloidal systems. It is this gap that our variational techniques are able to address.

Chapter 2

Variational optimization of steady-state trajectories

In this chapter we will develop a variational algorithm to optimize a control force within a chosen basis to compute large deviation functions. For Markovian systems, there exists an optimal control force, which is the unique additional force having the smallest contribution to the path ensemble measure that can be added to the system to make a rare uctuation typical [63, 110]. This optimal control force satis es several variational identities [111]. By deriving such a variational principle and explicit forms for the gradients required to optimize it, we develop an algorithm that approximates the control force su ciently well so as to make quantitatively accurate estimates of the likelihood of rare events within nonequilibrium steady-states. In this way, we generalize previous work on variational control of single particle systems to interacting, continuous force systems, bypassing the need for exponentially scaling Monte Carlo sampling. Our algorithm is similar in strategy to the recent use of thermodynamic variational principles to compute equilibrium free energies [71], and to the Rayleigh-Ritz variational principle that others have used to nonperturbatively compute e ective forces far from equilibrium [112]. The variational principle that underlies our algorithm is related to minimum-entropy production principles [113, 114]. and the Donsker-Varadhan formula in Markov Stochastic processes [115]. While our variational estimate of the large deviation function is subject to errors associated with the representation of the control force, we derive exact corrections that can be evaluated straightforwardly. In the rst two systems studied, these corrections are easy to evaluate, as our control forces are su ciently close to the optimal control forces to make these corrections perturbatively small. However, in cases where the corrections are large, we show that using optimized control forces in conjunction with standard Monte Carlo algorithms can increase the statistical e ciency in the estimation of large deviation functions by orders of magnitude. In this way, our algorithm is similar to the use of variationally optimized wavefunctions for quantum Di usion Monte Carlo calculations [116][1].

[1] Most of the content of this chapter was originally part of the publication [117].

2.1 Variational principle from Girsanov reweighting

Computing either of the large deviation functions of A_{t_f} requires sampling exponentially rare fluctuations. These rare fluctuations can in principle be made to occur more frequently by introducing an additional control force into the system as a means of importance sampling. In the presence of the additional force, (r^N), over the original force, $F(r^N)$, the computation of the SCGF can done by changing the path ensemble measure,

$$\begin{aligned}(s) &= \lim_{t_f \to \infty} \frac{1}{t_f} \ln \int \mathcal{D}[X(t_f)] e^{-s t_f A_{t_f}} \frac{P[X(t_f)]}{P[X(t_f)]} P[X(t_f)] \\ &= \lim_{t_f \to \infty} \frac{1}{t_f} \ln \langle e^{\mathcal{O}} \rangle \end{aligned} \tag{2.1}$$

where $\langle \cdot \rangle$ denotes an average in the controlled path ensemble with path probabilities $P[X(t_f)]$, and $\mathcal{O}[X(t_f)]$ can be derived from the difference in Onsager-Machlup path-actions $\mathcal{U}[X(t_f)]$, [45]

$$\mathcal{O}[X(t_f)] = -s t_f A_{t_f} + \int_0^{t_f} dt \sum_i \frac{\gamma_i^2 - 2\gamma_i(\gamma_i \dot{r}_i - F_i)}{4\gamma_i k_B T} - s t_f A_{t_f} - \mathcal{U} \tag{2.2}$$

interpreted in the Ito sense. Changing the force for such a Gaussian process does not change the normalization constant associated with the path ensemble in the long time limit where boundary terms from the initial and final configurations can be ignored.

Expanding Eq. 2.1 in terms of its cumulants, and using Jensen's inequality, we find a variational expression for the SCGF,

$$(s) \geq \lim_{t_f \to \infty} \frac{1}{t_f} \langle \mathcal{O} \rangle \tag{2.3}$$

in terms of the mean of $\mathcal{O}[X(t_f)]$, averaged within the controlled path ensemble. This expression is identical to previous work by Chetrite and Touchette that was derived using the contraction principle [111]. To see this, here we note that we can write $\langle \mathcal{O} \rangle$ in two ways,

$$\langle \mathcal{O} \rangle = \left\langle -s t_f A_{t_f} + \int_0^{t_f} dt \sum_i \frac{\gamma_i^2 - 2\gamma_i(\gamma_i \dot{r}_i - F_i)}{4\gamma_i k_B T} \right\rangle \tag{2.4}$$

$$= \left\langle -s t_f A_{t_f} - \int_0^{t_f} dt \sum_i \frac{\gamma_i^2}{4\gamma_i k_B T} \right\rangle \tag{2.5}$$

where we derive the second expression from the first by using the driven Langevin equation $\gamma_i \dot{r}_i = F_i + \xi_i + \eta_i$ and that, for an Ito dynamics, $\langle \xi_i(t) \eta_i(t) \rangle = 0$ because of causality.

Among the forces that make the rare value of the observable statistically typical, the one closest to the original force is the optimal force that realizes the supremum of the inequality.

This many-body function can be approximated within a chosen ansatz with variationally optimizable parameters $\{ \}$. In the limit that $\{ \}$ represents all possible functional forms of the many-body force, this ansatz becomes exact [111], so that

$$(s) = \sup_{\{ _1; _2;:::\}} \lim_{t_f !1} \frac{1}{t_f} O_{[\{ \}]} \quad _{[\{ \}]} \tag{2.6}$$

where the optimal coe cients $\{ \}$ will in general depend on s.

The existence of a control force that saturates the supremum in Eq. (2.6) follows again from the largest real eigenvalue (s) and the corresponding right eigenvector $_s(r^N)$, of the generator of the SCGF in Eq. 1.15, that directly gives the value of the SCGF,. The optimal force $_s$ that solves Eq. 2.6 is related to $_s$ through a Hopf-Cole transform [118, 63, 111] de ned as

$$_{s;i} = F_i + 2 \ _i k_B T(\ sg + r \ _i \ln \ _s) \tag{2.7}$$

and the controlled dynamics associated with this optimal force can be obtained from a generalized Doob transform of L_s [63, 119]. That $_s$ saturates the variational inequality 2.3 can be seen by writing the eigenvalue equation as

$$(s) = \ _s{}^1 L_s \ _s; \tag{2.8}$$

and multiplying both sides by a steady-state con gurational probability distribution associated with a generic control force, (r^N). If we then integrate over all coordinates with appropriate boundary conditions, it can be shown that the variational surface is convex around the supremum [113],

$$(s) = \lim_{t_f !1} \frac{1}{t_f} \left\langle O + \int_0^{t_f} dt \sum_i \frac{(\quad)^2}{4 \ _i k_B T} \right\rangle^{+} \tag{2.9}$$

However, though the variationally surface is locally convex, it may not be globally convex. Following the same derivation as above but with starting from an eigenvalue equation for a general (non-dominant) eigenvalue, the corresponding eigenvector, and its Hopf-Cole transform, it can be shown that the SCGF estimator in Eq. 2.6 can give the exact value of the real part of any eigenvalue of the tilted generator, for a physically meaningful force. This means that at dynamical crossovers when two eigenvalues approach each other and cross as a function of s, the variational expression is almost (in nitesimally) satis ed by two distinct forces, making the variational surface either slowly varying or bimodal.

For an interacting many-body system, the dominant eigenvector is a many body state, and therefore the optimal control force is many-bodied. Generally, we will assume that the control force is well approximated by a low rank ansatz such as obtained from a low order many body expansion.

Obtaining the SCGF from directly diagonalizing the tilted generator in many-body systems is prohibitively expensive due to the size of the multi-dimensional state space over

which L_s is de ned. There have been recent advances to approximate this state space using Matrix Product States for lattice based models [120, 121]. However, for continuous space systems with many particles, it is expected that Eq. 2.6 will present a physically motivated way to formulate approximate solutions to the eigenvalue problem and to the computation of (s), and subsequently, $I(A)$. It is worth noting that the constrained optimization of a variational expression analogous to (2.6) can also be directly used to compute $I(A)$ [111], with a straightforward extension of the algorithm described below.

2.2 Optimization algorithm with explicit gradients

In order to optimize Eq. 2.6 by gradient descent, we need to calculate derivatives of $\langle O[X(t_f)]\rangle$ with respect to the variational parameters $\{ g$ in the limit of a large t_f. Using these explicitly calculated gradients in the optimization algorithm can reduce the noise and numerical instabilities associated with nite di erence schemes, that are generally used to empirically estimate the gradients from the optimization trajectory through the parameter space. The explicit gradients that we use have the form of expectation values in the controlled ensemble,

$$\lim_{t_f \to 1} \frac{1}{t_f} \frac{@}{@} \langle O \rangle$$
$$= \lim_{t_f \to 1} \frac{1}{t_f} \left\langle \frac{O[X]}{} \frac{@}{@} \right\rangle + \left\langle O[X] \frac{@\ln P[X]}{@} \right\rangle \tag{2.10}$$

where is any of the optimizable parameters specifying the control force. The rst term is a simple trajectory average and is straightforward to compute. In a steady-state, we replace the time-extensive integral divided by the trajectory time with only the instantaneous contribution,

$$\lim_{t_f \to 1} \frac{1}{t_f} \left\langle \frac{O[X]}{} \frac{@}{@} \right\rangle = \left\langle \frac{O}{} \frac{@}{@} \right\rangle \tag{2.11}$$

where Eqns. (1.9) and (2.2) have been used to write $O[X]$ as a time integral of $O(t)$.

The second term is a two-time correlation function that is expected to have a high variance if directly computed in this form in a steady-state. Explicit functional forms of $@\ln P[X]/@$ can be derived from the normalized path probabilities as,

$$\left\langle O[X] \frac{@\ln P[X]}{@} \right\rangle = \left\langle \int_0^{t_f} dt\, O(t) \int_0^{t_f} dt^0 \sum_i \frac{{}_i(t^0)}{2\,{}_i k_B T} \frac{@\,{}_i(t^0)}{@} \right\rangle$$
$$- \left\langle \int_0^{t_f} dt\, O(t) \right\rangle \left\langle \int_0^{t_f} dt^0 \sum_i \frac{{}_i(t^0)}{2\,{}_i k_B T} \frac{@\,{}_i(t^0)}{@} \right\rangle : \tag{2.12}$$

The second term comes from the derivative of the normalization constant for $P[X]$, and it is formally zero as reweighting with a changed force doesn't change the normalization constants for the Gaussian white noise. We see this explicitly as $\langle s_i(t^0)\ @_i(t^0)/@i\rangle = 0$ due to causality where the product is Ito. Nevertheless, including this term during empirical estimation of the gradient reduces its variance, working as a noisy baseline subtraction.

The averages in Eq. 2.12 can be computed by propagating additional coordinates $y(t)$ associated with each variational parameter as

$$y(0) = 0\ ; \quad \underline{y}(t) = \sum_i \frac{\ _i(t)}{2\ _i k_B T}\ \frac{@\ _i(t)}{@} \tag{2.13}$$

where the sum has been performed over all dynamical coordinates of the system, and its uctuation is de ned as $\underline{y}(t) = \underline{y}(t) - \langle \underline{y}(t)\rangle$. These ctitious coordinates are known in the literature as Malliavin weights [122] and have previously been used to calculate parameter sensitivity of steady-state distributions in Langevin systems [123]. Provided these averages are evaluated in the steady-state generated by the control force $\underline{r}_i = F_i + \ _i + \ _i$, we can invoke time-translational invariance and note that only past noise history correlates with the observable to simplify Eq. 2.10. We start by de ning the uctuations in $\underline{O}(t)$ as $\underline{O}(t) = \underline{O}(t) - \langle \underline{O}(t)\rangle$ and rewriting the right-hand side of Eq. 2.10 as

$$\begin{aligned}
&\int_0^{t_f} dt \int_0^{t_f} dt^0 \langle \underline{O}(t)\ \underline{y}(t^0)\rangle \\
&= \int_0^{t_f} d \int_0^{t_f} dt^0 \langle \underline{O}(\ + t^0)\ \underline{y}(t^0)\rangle + \int_0^{t_f} d \int_0^{t_f} dt^0 \langle \underline{O}(\ + t^0)\ \underline{y}(t^0)\rangle && (2.14)\\
&= \int_0^{t_f} d \int_{t_f}^{t_f} dt^0 \langle \underline{O}(\ )\ \underline{y}(0)\rangle + \int_0^{t_f} d \int_0^{t_f} dt^0 \langle \underline{O}(\ )\ \underline{y}(0)\rangle && (2.15)
\end{aligned}$$

where at rst we have replaced $= t - t^0$ and changed the limits of integration appropriately, and then used time-translation invariance in the nonequilibrium steady-state. We now note that $\langle \underline{O}(\ )\ \underline{y}(0)\rangle$ is zero for all < 0, as the Malliavin weight $\underline{y}(0)$ is proportional to $_i(0)$ which is causally independent from all past values of $\underline{O}$. Hence the rst term vanishes and the second term simpli es to,

$$\int_0^{t_f} dt \int_0^{t_f} dt^0 \langle \underline{O}(t)\ \underline{y}(t^0)\rangle = \int_0^{t_f} (t_f - \ )d\ \langle \underline{O}(\ )\ \underline{y}(0)\rangle \tag{2.16}$$

Dividing both sides by t_f, taking the limit of $t_f \to 1$ and assuming a rapidly decaying correlation function in the steady-state gives the second term in Eq. 2.10 as

$$\lim_{t_f \to 1} \frac{1}{t_f} \langle O[X] \frac{@\ln P[X]}{@}\rangle = \int_0^1 dt \left\langle \underline{y}(0)\ \underline{O}(t)\right\rangle \tag{2.17}$$

where the gradient is proportional to an integrated time-correlation function. This is an example of a generalized uctuation-dissipation formula [124].

Putting together the two contributions,

$$\lim_{t_f \to \infty} \frac{1}{t_f} \frac{\partial}{\partial} \langle \mathcal{O} \rangle = \left\langle \frac{\mathcal{O}[X]}{} \frac{\partial}{\partial} \right\rangle + \int_0^{\infty} dt \left\langle \underline{y}(0)\, \mathcal{O}(t) \right\rangle ; \tag{2.18}$$

we arrive at an explicit form for the gradient of our SCGF estimate with respect to the variational parameters that can be estimated as time-averages from a straightforward molecular dynamics trajectory with the control forces. In practice, we will take the integral over the time correlation function in Eq. (2.18) up to a time t. The exact form we use to evaluate this correlation function on-the-fly is obtained from Eq. 2.17 by translating the correlation time-lag $[0; t]$ to $[-t; 0]$, replacing the dummy variable t by $-t$ and carrying out the integration explicitly, as

$$\int_0^{\infty} dt \left\langle \underline{y}(0)\, \mathcal{O}(t) \right\rangle = \int_{-t}^{0} dt \left\langle \underline{y}(-t)\, \mathcal{O}(0) \right\rangle = \int_{0}^{t} dt \left\langle \underline{y}(t)\, \mathcal{O}(0) \right\rangle \tag{2.19}$$

$$\left\langle \underline{y}(0+)\, \mathcal{O}(0) \right\rangle - \left\langle \underline{y}(-t)\, \mathcal{O}(0) \right\rangle \tag{2.20}$$

Here the time 0+ denotes the next infinitesimal timestep after time 0, and t is large enough such that the estimator is accurate upto a desired tolerance. We propagate the dynamics of y and keep a history upto lag-time t, such that we can statically correlate with $\mathcal{O}$ at every timestep. The baseline subtraction at $-t$ and the subtraction of the averages for computing the correlations from joint expectations helps keep the variance of the gradient well-behaved, even in a steady-state where $y(t)$ is a random walk with a time-extensive variance [122].

We also note here that when $\langle \mathcal{O} \rangle$ is represented as Eq. 2.4, the first term in the gradient in Eq. 2.18 is zero. This is seen readily by inserting Eq. 2.4 into Eq. 2.18 to get the first term as

$$\begin{aligned} \left\langle \frac{\mathcal{O}[X]}{} \frac{\partial}{\partial} \right\rangle &= \left\langle \int_0^{t_f} dt \sum_i \frac{{}_i \quad {}_i \underline{r}_i + F_i}{2\, {}_i k_B T} \frac{\partial}{\partial} \right\rangle \\ &= \left\langle \int_0^{t_f} dt \sum_i \frac{{}_i(t)}{2\, {}_i k_B T} \frac{\partial\, (t)}{\partial} \right\rangle = 0 \end{aligned} \tag{2.21}$$

due to causality. It can be shown that this term is also proportional to the derivative of the normalization constant for the trajectory probabilities with respect to changing the force, and this is zero [125]. However, if one uses the form $\langle \mathcal{O} \rangle$ from Eq. 2.5, the first term in Eq. 2.18 is not zero and must be independently computed as a static average.

When the explicit gradients are averaged over a long steady-state trajectory at every optimization step, the noise in the gradients become small compared to their means and they can be directly used to perform variants of gradient descent. We perform an iterative optimization in the parameter space spanned by $\{ \}$ in order to estimate the SCGF. We use

an algorithm called Nesterov's Accelerated Gradient Descent [126, 127] which shows super-linear convergence. This algorithm updates a conjugate momentum p in parameter space to accelerate optimization with inertia, as well as prevents inertial oscillations with a predictive step. The conjugate momentum is a sum of the scaled gradient with scaling parameter, and the scaled momentum from the previous optimization step with scaling parameter. These parameters can be chosen at the start of the optimization or can be tuned with iterations to improve convergence. We stop either a fixed number of optimization steps or when the norm of the force gradients are less than a tolerance value. The optimization algorithm is given in pseudocode form in Algorithm 1.

This algorithm converges to a local optimum in the parameter space, which can be different from the global optimum when the variational surface or its projection on a truncated basis space is not convex. The convergence can also be significantly slow at values of far a crossover point or a phase transition. In the event that we converge to a local optimum, we incur a systematic error in the SCGF that can be corrected with a cumulant expansion.

2.3 Corrections of systematic errors

In general, the ansatz specified by the parameters will not form a complete basis for a many body system. This is because generically, the dominant eigenvector of Eq. 1.16 is a many-body state, containing exponentially many parameters, and not expected to be exactly expressible with a low rank form. Because of this, the variationally converged SCGF $\theta(s)$ obtained from Eq. 2.6 with an optimized force e will have a systematic error. This error, and errors associated with convergence to a local maximum, can both be corrected in principle by computing the remaining terms of the cumulant expansion

$$(s) = \theta(s) + \lim_{t_f \to 1} \frac{1}{t_f} \sum_{\ell=2}^{\infty} \frac{\ell}{\ell!} \tag{2.22}$$

where ℓ are the second and higher cumulants in the expansion of the $\exp(O_e)$ in e and the force e is the solution of the variational problem in the approximate and incomplete ansatz. If the ansatz used to express the control force e, is close enough to the optimal force obtained from the Doob transform, the correction terms are small in magnitude and the series will converge quickly. This will occur when the trajectory distribution generated by the controlled dynamics has significant overlap with the tilted distribution of the original dynamics.

In cases where the ansatz is poor and many cumulants are needed, brute force convergence of the correction will be difficult. In such cases, control forces can be used as guiding functions for estimating the SCGF through Monte Carlo based approaches like the cloning algorithm. In the cloning algorithm, an ensemble of N_w trajectories generated from the ordinary path probabilities $P[X(t_f)]$ are branched with corresponding weights of $\exp(st_f A_{t_f})$. However, under the controlled dynamics, following Eq. 2.1, the weighted path probabilities can be

Algorithm 1 Nesterov's Accelerated Gradient Descent with Malliavin weights

inputs Variational parameters for a general force $(r^N; t)$
parameters Force gradient and momentum learning rates , ; total optimization steps I; trajectory length t_f consisting of J timesteps of duration t each; number of trajectories N
initialize choose initial weights , define iteration variables i and j, force gradients and conjugate momenta and p, define functional form for stepwise differential increments (rewards) to the loss-functions $t_f A_{t_f} + U$
$i \quad 0$
$p \quad 0$
repeat
  Generate trajectory $[X(t)]$ with first-order Euler propagation in presence of the additional force $[f + p\ g]$ and wait till it relaxes into a nonequilibrium steady-state. Configurations, times, noises (with variance $2k_B T \ t$), Malliavin weights, and rewards are denoted by $r_j^N; t_j; \ _j; y(t_j)$ and $(t_j) = \ _j$ respectively.
  $j \quad 0$
  $\quad 0$
  $y(t_0) \quad 0$
  repeat
    $\underline{y}(t_j) \quad _j \ r \quad (r_j^N; t_j)/2k_B T \ t$
    $y(t_{j+1}) \quad y(t_j) + t\underline{y}(t_j)$
    $+ \ _j y(t_{j+1})$
    $j \quad j + 1$
  until $j = J$
  $= J$
  average over N trajectories to get $\bar{\ }$
  $p \quad p$
  $+ p$
  $i \quad i + 1$
until $i = I$

written as [84, 89]

$$P_s[X(t_f)] \,/\, e^{-s t_f A_{t_f}} P[X] = e^{-\mathcal{O}[X]} P[X] \tag{2.23}$$

where system evolution under an approximate controlled dynamics is nonconservative and must be accompanied by branching steps with weights given by $\exp(-\mathcal{O})$. An estimate of the SCGF is then obtained from the normalization constant of this weight, so that in the limit of large N_w,

$$\Lambda(s) = \frac{1}{t_f} \ln \frac{1}{N_w} \sum_{j=1}^{N_w} e^{-\mathcal{O}[X_j]} \tag{2.24}$$

where $\mathcal{O}[X_j]$ denotes the time-integrated observable for the walker labelled j.

When the variationally optimized θ is used to generate trajectories and to compute the branching probabilities, the efficiency of the cloning algorithm is improved as the control force samples the rare fluctuations in the observable. When θ is actually the optimal force θ_s derived from the Doob transform, all trajectories achieve the rare fluctuation as typical behavior, and the weight of each trajectory becomes a constant. In this situation no trajectories are killed in the branching step of the cloning algorithm, and the sampling is statistically optimal [77]. However, even with an approximate ansatz the variationally optimized force slows down the rate of death of uncorrelated trajectories with increasing t_f, as demonstrated in Section 2.5.

The variational algorithm along with the cumulant-correction has improved scaling properties compared to the cloning algorithm. By adopting an approximate ansatz for the many-body force containing, for example, one-body and two-body terms, for a system of identical particles we can exploit their permutation symmetry and optimize a single one-body and two-body force. Hence the variational algorithm scales linearly with the system size, the computational cost arising only from the propagation of trajectories of interacting particles. This is in contrast to the cloning algorithm, which has an exponential scaling for observables that are system size extensive [81]. Also, while the cloning algorithm scales exponentially with s, the variational algorithm depends on the bias only through the complexity of the optimal force and scales linearly with the number of variational parameters required to approximate the force. Hence in cases that the dominant part of the optimal force can be simply expressed within the choice of the ansatz, the computational cost for the algorithm to converge does not increase with s. This indicates a resummation of the exponential bias through the modification of the control force. Neither does the algorithm scale with increasing observation time t_f, as the $t_f \to \infty$ limit has already been incorporated in the algorithm. Lastly, this algorithm can be parallelized trivially by distributing the computation of the expectation values at each step of the iteration to independent trajectories on independent processors.

To study the accuracy and efficiency of our variational algorithm to compute the SCGF and the optimal force, in the next sections we apply it to three different continuous time and space systems. The first is a benchmark system where we can test our algorithm against a numerically exact result. This model consists of a driven underdamped particle in a

one-dimensional periodic potential, for which we have studied rare uctuations of the total current. The second system is comprised of multiple repulsive overdamped particles, where we have focused on the uctuations of the total activity, which measures how much the particles explore con guration space. In this system, we demonstrate the ability of our algorithm to compute the optimal control force even through singular changes in the SCGF across a dynamical phase transition. The third system consists of multiple repulsive particles in a periodic potential having a wavelength equal to the particle diameter and an external nonequilibrium force, such that it is a Brownian analogue of the Asymmetric Simple Exclusion Process (BASEP). We study its current SCGF and nd that many-body e ective forces are necessary to generate rare backward current uctuations at high packing fractions.

2.4 Current uctuations in non-interacting overdamped/underdamped system

An underdamped particle being driven on a periodic potential by a constant external force is a simple system with two dynamical coordinates, position and velocity, that can exhibit nontrivial nonequilibrium properties due to competing ballistic and di usive modes of transport [128, 129]. Large deviation functions for current uctuations in this model can be obtained by numerically exact diagonalizations of the tilted generator, and the controlled ensemble can show diverse behavior in di erent parameter regimes [130]. We consider this model to benchmark our variational optimization algorithm.

Speci cally, we consider an underdamped particle of mass m moving in a one-dimenional periodic box of length $L = 2$. The forces acting on the particle are derived from a cosine potential, $V(x) = V_0 \cos(x)$, where V_0 is the magnitude of the potential, and include a constant external driving force, F_{ext}. For the particle in contact with a bath of temperature, T, and friction coe cient, , the equations of motion for the position, x, and velocity, v, are

$$\begin{aligned} \underline{x} &= v \\ m\underline{v} &= F(x) \quad v + \end{aligned} \tag{2.25}$$

where $F(x) = V^{0}(x) + F_{ext}$ and (t) is a Gaussian white noise with

$$h(t)i = 0 \qquad h(t)\ (t^{0})i = 2\ k_B T\ (t \quad t^{0}) \tag{2.26}$$

where k_B is Boltzmann's constant. These equations of motion have the form of Eqs. (1.1) and (1.3) with two dynamical coordinates and a vanishing noise in position [130].

We investigate the statistics of the time-averaged current owing through the system,

$$J_{t_f} = \frac{1}{t_f} \int_0^{t_f} dt\, v(t) \tag{2.27}$$

which measures the total displacement of the particle. The SCGF for current is given by

$$(\quad s) = \sup_{u(x;v)} \frac{1}{t_f} \int_0^{t_f} dt \quad sv + \frac{u^2 \quad F^2 \quad 2v(u \quad F)}{4k_B T} \quad \frac{m\underline{v}(u \quad F)}{2k_B T} \quad {}_u \tag{2.28}$$

where the path average is obtained from the controlled dynamics

$$m\underline{v} = u(x;v) \quad v + \tag{2.29}$$

and the optimal force is in general a function of both position and velocity. We expand the additional force $(x;v) = u(x;v) \quad F(x)$ in an ansatz

$$u(x;v) = F(x) + \sum_{p=\ M_1}^{M_1} \sum_{q=0}^{M_2} c_{p;q} e^{ipx} v^q \tag{2.30}$$

where c_{pq} are parameters that can be optimized variationally subject to $c_{\ p;q} = c_{p;q}$, and the number of position and velocity basis functions are $(2M_1 + 1)$ and $(M_2 + 1)$ respectively. The basis is complete in the limit of large M_1 and M_2. Note that this force incorporates the periodicity of x and also allows the external nonequilibrium driving, which is the $p = q = 0$ term, to be optimized. In the high friction limit, the dynamics becomes overdamped and in that limit the optimal force becomes a function of just the particle position. For small friction, inertia is important and the general form of the optimal force must be considered. We note that this velocity-dependent drift function is a force only in a generalized sense.

The SCGF and the optimized control force obtained from the variational algorithm can be compared to numerically exact results obtained by solving the eigenvalue equation for the tilted generator given by [130]

$$L_s = v\frac{@}{@x} \quad \frac{1}{m}[v \quad F(x)]\frac{@}{@v} + \frac{k_B T}{m^2}\frac{@^2}{@v^2} \quad sv \tag{2.31}$$

as in Eq. (1.16). The exact control force is obtained using the right eigenvector $_s(x;v)$ corresponding to the largest real eigenvalue, as

$$u(x;v) = F(x) + \frac{2k_B T}{m}\frac{@\ln \ _s(x;v)}{@v} \tag{2.32}$$

where numerical diagonalization of L_s can be performed by expressing the right and left eigenvectors over a position-velocity grid and representing the di erential operators in L_s using a second order nite di erence scheme. The boundary conditions are periodic in the position grid and re ective in the velocity grid, so that only forward (backward) di erence at the minimum (maximum) velocity grid point is used to represent the di erential operator.

We have computed the cumulant-corrected large deviation functions in this system and have compared them to the numerically exact results. We have worked with $k_B T = 1$ and $= 1$. These parameters along with the length of the box $L = 2$ let us de ne our natural

Figure 2.1: a) Convergence of $_n(\ t)$ for $s = \ 0{:}5$, for di erent n. Shaded region represents optimal choice of $\ t$ for gradient descent. b) Simultaneous convergence of SCGF and biased density for $s = \ 0{:}5$.

time unit as $t\ = 4\ ^2 k_B T = L^2$. All observables have been reported in dimensionless units following these de nitions. We have done our computations at two values of mass, $m = t\ = 1$ and $m = t\ !\ 0$. We have also chosen $V_0 = 2$ and $F_{ext} = 1$. The numerically exact result was obtained with a grid of $140\ \ 50$ points in the position-velocity space. The position points span all of the box and the velocity points are centered at $F(_{ext}\ \ 2sk_B T) = \ $ corresponding to the mean velocity in the $V_0\ !\ 0$ limit. For all the simulations, the timestep was chosen to be $0{:}001$ natural time units. For $m = t\ !\ 0$, an Euler scheme was used to integrate the overdamped equation of motion, while for $m = t\ = 1$, a velocity Verlet scheme was used [42].

For each iterative step during the optimization, a trajectory of duration 10^4 units was simulated. During the rst half of each trajectory, the system was allowed to come to a steady-state, and the time-averaged gradients were computed only with the second half of the trajectories. For implementing Eq. 2.18, we integrated the correlation function up to $t = 100$. The size of the basis was $M_1 = 3; M_2 = 1$ for $m = t\ = 1$ and $M_1 = 3; M_2 = 0$ for $m = t\ !\ 0$, the overdamped limit. The optimization parameters used for the gradient descent were $\ = 0{:}5; \ = 0{:}2$. Near $s = 0$, all c_{pq} were initialized at zero, and subsequent optimizations with increasing magnitude of s were initialized from a previously optimized set of c_{pq} taken from the nearest value of s. In the overdamped limit, an accurate estimate of the SCGF could be obtained with just the variational optimization, with the cumulant correction merely a con rmation of the optimal control forces being correct. However for $m = t\ = 1$, the variational SCGF had to be corrected with cumulants computed with an observation time $t_f = 100$ and a total trajectory length 10^5 units.

The choice of a nite integration limit $\ t$ to compute the integral in Eq. 2.17 depends on both the intrinsic correlation times of the system and the timescale of the variance of the

Figure 2.2: Large deviation functions for current uctuations in a driven underdamped system in a periodic potential. a) SCGF for $m = t = 1$ with $M_1 = 3; M_2 = 1$ and for $m = t \to 0$ with $M_1 = 3; M_2 = 0$. b) Rate functions obtained by a numerical Legendre-Fenchel transform of the SCGFs. The legend is the same as that used in a) Inset: Schematic diagram of the simulated system.

integrated correlation function to diverge. To illustrate this, we plot

$$_n(t) = \int_0^t dt \left\langle y_n(0)\, Q_-(t) \right\rangle \tag{2.33}$$

for this system in the $m = t \to 0$ limit. The ansatz can be written in this limit as

$$u(x) = F(x) + c_0 + \sum_{n=1}^{3} [c_n \cos x + c_n \sin x] \tag{2.34}$$

and for Fig. 2.1, we have chosen c_n parameter values randomly between 1 and 1, with $s = 0.5$. We see that even though the correlation function converges for large t, the error in the computed gradient increases steadily. For all the results in this paper, t was chosen to balance between these two e ects so that the computed gradients su ers from no systematic error and minimum statistical error.

Using these numerically estimated gradients, the accelerated gradient descent algorithm converges superlinearly, and in Fig. 2.1 we have plotted the decrease of the systematic error (s) in the current SCGF estimate with optimization steps in the limit $m = t \to 0$. We also show the simultaneous convergence of the controlled ensemble steady-state density $_u^{ss}(x)$ to the true biased steady-state density $_s(x) / \ _s(x)\ _s(x)$ where $_s$ and $_s$ are the dominant left and right eigenvectors of the tilted generator (2.31). We demonstrate this by plotting the relative entropy of the two,

$$D(\ _u^{ss} \| \ _s) = \int dx\ \ _u^{ss}(x) \log \frac{_u^{ss}(x)}{_s(x)} \tag{2.35}$$

Figure 2.3: Overdamped limit, $m = t \to 0$, of the driven particle on a periodic potential. a) Optimized control forces (dashed lines) overlaid on the exact control force (solid lines). The thick curve is for $s = 0$ and the curves above (below) are for s in intervals of $+0.5$ (-0.5). b) Basis size errors in the variational estimate of (s), where the deviation $(s) = \tilde{(}s) - (s)$ is the difference between the finite basis result $\tilde{(}s)$ from the exact SCGF.

which shows that even as only the current is being optimized to have a nontypical value, the entire trajectory ensemble simultaneously converges to the exact biased ensemble.

Following this procedure, we obtain estimates of SCGFs that are in quantitative agreement with the numerically exact results throughout the range of s considered, as shown in Fig. 2.2(a). We have also calculated the rate functions for the current, Fig. 2.2(b), in these two parameter regimes by a numerical Legendre-Fenchel transform of the SCGFs.

The SCGFs in Fig. 2.2(a) both have a *locked* region where the current changes slowly with s, and an *unlocked* region for larger magnitudes of s. Due to the time-reversal properties of L_s, the SCGF shows a Gallavotti-Cohen symmetry [131]

$$(\ ) = (\ F_{ext}/k_B T\ ) \tag{2.36}$$

which is clear through the reflection symmetry about $s = 0.5$ of the SCGF in Fig. 2.2(a). Analogously, the rate function obeys a fluctuation theorem symmetry

$$I(J) = I(-J) + F_{ext} J/k_B T \tag{2.37}$$

indicating the exponentially rare probability of a current in the direction opposite to the applied force.

Figure 2.3(a) shows the position-dependent optimal forces obtained in the overdamped limit, $u^{(1)}(x)$, overlaid on the numerically exact answers obtained from diagonalization [132], for multiple values of s. In the limit of $|s| \to 1$, the optimal forces approach the free-diffusion limit, where the majority of contribution comes from a constant nonequilibrium driving.

Figure 2.4: Underdamped system, $m = t = 1$, of the driven particle on a periodic potential. a) (L-R) Exact and optimized control forces, $u(x; v)$, for $s = 1.5$, with the solid contour at $u(x; v) = 2$, and the dashed (dotted) contours being at differences of +1 (1). Exact and optimized control forces, $u(x; v)$, for $s = 2$ with the solid contour at $u(x; v) = 5$ and the dashed (dotted) contours being at differences of +1 (1). b) Convergence of the cumulant expansion for representative values of s.

When $|s|$ is of the order $|F_{ext}|/k_B T$, the forces have a non-trivial position dependence. This is manifested in the size of the basis-set, M_1, required to obtain the optimal control force accurately. Figure 2.3(b) shows the effect of finite basis size on the error made in estimating (s). Increasing M_1 reduces the error and ultimately the ansatz becomes exact when M_1 is large. The error decreases when going to larger $|s|$ as the forces are easier to represent using the first few basis functions. The error bars were computed from 5 independent estimates of the SCGF using independent trajectories.

For the $m = t = 1$ system, inertial effects are important and the optimal force depends on both position and velocity, and the optimal force has a complicated functional dependency that is difficult to represent using a small number of basis functions. Using a truncated basis to represent the control force leads to a systematic error in the SCGF estimate obtained using Eq. 2.28 that can be corrected using the cumulant expansion in Eq. (2.22). Figure 2.4(a) shows the approximate forces obtained from the variational optimization compared to the numerically exact results. When s is near the Gallavotti-Cohen symmetry point, the average current is small and the optimal control force is a complicated function of both v and x. Within our ansatz, the optimized $u(x; v)$ does not reproduce the exact form of the optimal control force. Nevertheless, these approximate forces recover the majority of the SCGF, so that the cumulant expansion converges for all tested s points. Figure 2.4(a) also contains the optimal force at a larger positive s, where the forces lose their velocity dependence and simplify towards the free-diffusion limit. In this limit, position based forces are sufficient to

recover the SCGF quantitatively.

Figure 2.4(b) shows the convergence of the consecutive terms of the cumulant expansion in Eq. 2.22 for di erent values of s. $_1$, the rst cumulant, is identical to $\theta(s)$, the variational estimate. Error bars were calculated using 5 independent trajectories for the estimation of the cumulants. Even though our basis is small and approximate, the cumulants computed from a single trajectory have decreasing amplitudes for various values of s, showing that the variational force is accurate enough to approach the force derived from the Doob transform. We note that the sign of the cumulants need not be positive, and therefore the variational structure in the estimate of (s) holds only for the rst cumulant. Further, the magnitude of the terms in the cumulant expansion need not be strictly decreasing. Figure 2.4(b) includes an example of a nonmonotonic convergence for $s = 1.1$. Moreover, the sign of the error of the approximate SCGF at a given truncation of the cumulant expansion can change resulting in the cancellation of errors of two oppositely signed cumulant corrections and an accidental near agreement of the exact SCGF. We have found that by considering the convergence of the consecutive terms of the cumulant expansion we can reliably determine the accuracy of the approximate SCGF.

2.5 Activity uctuations in interacting overdamped system

To study how this algorithm performs in an interacting system, we consider the uctuations of the activity in a system of overdamped repulsive particles on a line. In both lattice and continuum models of volume excluding particles in one dimension, it has been reported that there are two characteristically distinct types of activity uctuations, with a dynamical phase transition separating them [133]. For rare large negative values of the activity, such systems spontaneously phase separate into macroscopically sized clusters, whereas for rare small values of the activity, they form a hyperuniform phase in which long-wavelength density uctuations are suppressed. This behavior emerges as a singularity in the SCGF and a closing of the gap in the eigenspectrum of the tilted operator, which in the hydrodynamic scaling limit is predicted to occur with a critical point at $s_c \to 0^+$ [133, 89]. This system is thus suitable to test the e ectiveness of the variational algorithm in computing rare uctuations that are collective in origin.

Speci cally, we study the uctuations of dynamical activity in a system of N overdamped repulsive particles in a one-dimensional periodic box of length L. The equation of motion is

$$\dot{x}_i = F_i(x) + \eta_i \tag{2.38}$$

where $F_i(x)$ is the total force felt by the i-th particle,

$$F_i(x) = -\frac{\partial}{\partial x_i} \sum_{j \neq i} V_{WCA}(x_{ij}) \tag{2.39}$$

where $x_{ij} = x_i \quad x_j$ and the force is derived from a WCA pair potential

$$V_{WCA}(r) = \quad 4 \quad \frac{\quad^{12}}{r^{12}} \quad \frac{\quad^{6}}{r^{6}} + \quad ; \quad r < 2^{1=6} \tag{2.40}$$
$$= 0 ; \quad r \quad 2^{1=6}$$

with characteristic energy, , and length scale, . The Gaussian white noise, $_i$, is speci ed by

$$h_i(t)i = 0 ; \qquad h_i(t) \; _j(t^0)i = 2 \; k_B T \; _{ij} \; (t \quad t^0) \tag{2.41}$$

We work with $k_B T = 0:5$, $= 1$ and $= 1$. As before, we de ne our unit of time for this system as $2k_B T = \quad ^2$ and we have reported all observables in dimensionless units. Additionally, we set $= 1$ and consider a density of $= N =L = 0:5$, so that the box is half- lled.

We study a measure of activity derived from the probability that the particles stay in the same state in a short time interval [134]. This form of the activity,

$$K_{t_f} = \frac{1}{t_f} \int_0 dt \sum_i \quad \frac{F_i^2}{4 \; k_B T} + \frac{1}{2} \frac{@F_i}{@x_i} \tag{2.42}$$

is also a part of the time-symmetric component of the path-action [135], and its long time statistics are similar to other commonly used metrics that count the total number of hops for particles on a lattice [136, 137]. Using Ito's Lemma to simplify the last term in Eq. 2.2, the variational expression for the SCGF becomes

$$(\; s) = \sup_{u(x_1;x_2;::::;x_N)} \frac{1}{t_f} \int_0^{t_f} dt \sum_i \quad s \quad \frac{F_i^2}{4 \; k_B T} + \frac{1}{2} \frac{@F_i}{@x_i}$$
$$+ \frac{u_i^2 \quad F_i^2}{4 \; k_B T} + \frac{1}{2} \frac{@(u_i \quad F_i)}{@x_i} \Big_u \tag{2.43}$$

where in addition to the force, we require the gradient of both the original and the control force.

For this system, the optimal control force $u(x)$ is in general long-range and many-bodied. Previous work on related one-dimensional systems have shown long-range repulsive interactions stabilizing the hyperuniform state for values of activity small in magnitude [95], and long-range attractive forces acting on the surface of particle clusters that emerge in rare large negative uctuations of the activity [89]. For our variational ansatz, we have approximated the many-body force as a sum of long-range pairwise interactions. Pair forces are the lowest rank approximation to this system due to its translational invariance. From the Hopf-Cole transform, optimization of a pair force is analogous to optimization of a two-body Jastrow function as used in variational quantum Monte Carlo [138].

To represent the control force, we expand it in a basis of Laguerre polynomials L_p with coe cients c_p as

$$u_i = \sum_{j \in i} \frac{@}{@x_i} V_{WCA}(x_{ij}) + \sum_{p=1}^{M_3} c_p L_p(x_{ij}) e^{\; x_{ij} =2} \frac{x_{ij}}{jx_{ij} j} \tag{2.44}$$

Figure 2.5: Size-scaling of activity uctuations of repulsive particles on a line. a) $O(N^2)$ scaling of (s) in the phase-separated state. Inset) O(N) scaling in the hyperuniform state. b) Change in mean activity across the dynamical phase transition. Inset) Schematic representation of the phase-separated (left) and hyperuniform (right) states.

where x_{ij} = $|x_{ij}|$ is a linear transformation on the distance between particles i and j . The parameters and can be adjusted to set a scale and a cuto for where the force smoothly decays to zero, and M_3 determines the size of the basis. The basis is complete for all possible two-body forces in the limit of large M_3. The exponential factor makes the basis functions orthogonal and aids in the convergence of the optimization. We have used M_3 = 10 for all of our results. We have xed = 2=L, and optimized $\{c_p\}$ and with starting values of 0 and L=2 respectively. In each iteration of the optimization, a trajectory of length $2 \cdot 10^4$ time units is simulated, the rst half again reserved for equilibration and the second half being used to compute the gradients. For computing the integrated correlation function in Eq. 2.18, we have used t = 200 units. After obtaining the optimized control force in this ansatz, we use it to compute the unbiased SCGF using a cumulant expansion as before, with an observation time t_f = 10 and a total trajectory length of $5 \cdot 10^4$ units. Across the range of s considered, we nd convergence using the rst three cumulants to correct the variational result. The SCGF obtained from this cumulant expansion is identical to results obtained using a guided cloning algorithm that has been described later in this section.

In Figure 2.5(a) we have plotted the size scaled SCGF, and the mean activity, for positive and negative values of s. For s < 0, we nd the system in a hyperuniform state, where all particles are pushed apart from each other and long-range density uctuations are suppressed [133]. The SCGF is size-extensive in this range of s. For s 0 the particles phase separate, forming a single cluster. In the region where s is positive but small, there is a phase transition to this clustered state accompanied by an in ection point in the mean activity, shown in Fig. 2.5(b), obtained from taking the numerical derivative of the SCGF, $\langle K \rangle_s = {}^0(s)$. The extensive scaling regime has been explored systematically in a related model and found to

Figure 2.6: Optimal pair-potential for positive and negative s for $N = 40$.

agree well with predictions from macroscopical uctuation theory [89]. In our studies, we nd it limited to $0 < s < 0.02$. For large positive values of s, the cluster is a highly compressed solid with system-spanning correlations that result in the SCGF scaling super-extensively. In this regime of the SCGF, the typical force is on the order of $\sqrt{N}$, and can continue to increase with increasing s because of the soft repulsion of the WCA potential. Inspection of the distribution of mean squared forces reveals that the cluster is not homogeneous, but most compressed in its interior with lower density near the edges, with a system size independent pro le. The phase transition from a disordered state to a clustered state is in accord with previous observations in related systems, and result in diverging correlation times rendering the precise study of the critical point di cult [89, 133]. We therefore focus our attention on the two phases on either side of that transition. Error bars were obtained from independent statistics from 3 distinct trajectories.

Figure 2.6 shows the e ective pair-potential, $V^{(2)}(r)$, derived from the optimal control force at di erent values of s, for $N = 40$, obtained by the numerical integration of the control force. The potential is long-ranged and repulsive in the hyperuniform phase, and long-ranged and attractive in the clustered phase. The long-range potential leads to the observed size scaling in Fig. 2.5, because it imposes in nite range correlations. We also observe that the depth of the attractive potential for increasingly positive values of s tends to saturate, while the magnitude of the repulsive potential for increasingly negative s does not. This di erence arises from the steeply rising WCA forces that can achieve more negative values of K_s with just a slight decrease in the nearest neighbor distance in the controlled system. In the hyperuniform phase, achieving the rarer values of activity implies an exponentially small number of collisions between the particles, which leads to an increasing repulsive control force. These optimal control forces derived from the variational ansatz do not contain many-body components unlike analytically derived approximate forces [89], yet they achieve the same phenomenology of phase separation and hyperuniformity described previously.

Figure 2.7: Characterization of the two dynamical phases for $N = 80$. a) Pair distribution functions within the phase separated, $s = 0.1$, and (inset) hyperuniform, $s = -0.1$, states. b) Structure factor for various system sizes in the hyperuniform state, $s = -1$.

Figure 2.7(a) characterizes the steady-state radial distribution function $g(r)$,

$$g(r) = N \langle \delta(r - |x_{12}|) \rangle_u \tag{2.45}$$

obtained in these phases, for a system size of $N = 80$, where x_{12} denotes the interparticle distance between each distinct pair of particles. In the phase-separated state, the particles form a solid cluster that has sharp peaks in $g(r)$ at intervals of . In the hyperuniform phase, the particles are repelled away from each other and $g(r)$ has little structure aside from the volume-exclusion. We also characterize the structure of the hyperuniform state through the structure factor, $S(q)$, as a function of the wavenumber q, obtained from

$$S(q) = \frac{1}{L} \left\langle \left| \sum_{j=1}^{N} e^{iqx_j} \right|^2 \right\rangle_u \tag{2.46}$$

where the averages are computed in the ensemble with the control force. A linear increase of $S(q)$ from zero at small q is a signature of the suppression of long-wavelength density uctuations in the hyperuniform phase, which we con rm in Fig. 2.7(b). The spike at $q = 2 = 2^{1/6}$ results from $2^{1/6}$ being the distance of closest approach of the repulsive particles without experiencing a force.

Under large positive activity bias, we nd that the overdamped repulsive particles form a highly compressed cluster. This cluster is described by system-spanning correlations. Shown in Fig. 2.8 is the size-scaled pro le for the rst term of the collective activity (2.42), $\langle F_i^2 \rangle_s / 4 k_B T$, with respect to a size-scaled particle index $N_i = i - (N + 1)/2$. The particles are indexed from one end of the cluster to the other, such that the center of the cluster is indexed at $N_i = 0$. The compressed cluster does not break apart during the duration of the

Figure 2.8: Size-scaling of the mean squared force pro le within the cluster for $s = 0:1$.

trajectories observed, so that large $|N_i|$ unambiguously refers to particles close to the surface of the cluster. The total mean activity $\langle K \rangle_s$ is proportional to the total mean squared force appearing in the rst term, such that the pro le of the second term in the de nition looks analogous only with an opposite sign [137]. The $\Theta(N^2)$ scaling of the mean squared force and its size-invariant parabolic pro le explains the super-extensive SCGF scaling and the system spanning correlations in this s regime.

While we have not investigated the phase transition directly, the disparate behavior of either side of the dynamical phase transition provides a useful test of our ability to obtain control forces, as the structure and dynamics of the system in the phase separated and hyperuniform states are very di erent. Despite their di erences in both regimes, we are able to obtain control forces that are near enough to the optimal force to converge the large deviation functions using a brute force evaluation of the remaining cumulant expansion. Nevertheless, we expect this strategy may fail in general, in which case a more robust means of estimating the remaining contribution must be employed. To explore such alternatives, we apply these control forces as guiding functions within the cloning algorithm [84]. To quantify the statistical bene t from the control forces, we start with a trajectory ensemble of $N_w = 32000$ walkers and monitor the decay rate in the number of uncorrelated walkers, N_c, with and without the control forces. The number of uncorrelated walkers is de ned as those with a distinct history, having not been previously merged into an existing walker. Figures 2.9(a) and (b) show the statistics of the walkers with respect to observation time, with and without the control forces, in a system with 20 particles, and branching steps taken every 0.5 time units. We have plotted $f_s^c(t) = N_c(t)/N_w(t)$, where t is the observation time, to represent the growth of correlation in the trajectory ensemble.

In the clustered state, incorporating the control forces improves the number of uncorrelated walkers by multiple orders of magnitude. For larger positive s, an unbiased estimate of the SCGF can be obtained only when the variational control forces are used. The improve-

Figure 2.9: Improvement of walker statistics of the cloning algorithm using approximate control forces as guiding functions in an N=20 system, represented by $f_s^c(t) = N_c(t)/N_w(t)$, after an observation time t. Blue circles are without a guiding force and green squares are with the variationally optimized guiding force. Decay of the fraction of uncorrelated walkers with increasing observation time in a) the phase-separated state $s = 0.04$) and b) the hyperuniform state ($s = 0.2$). (Insets) Decay of the fraction of uncorrelated walkers after $t = 20$ as a function of s in a) the phase-separated state ($s = 0.04$) and b) the hyperuniform state ($s = 0.2$).

ment in the statistics of the walkers increases for more positive s because the magnitude of the SCGF grows rapidly, and therefore the weight carried by the branching step increases. We see this effect in the inset, where we show the fraction of uncorrelated walkers left after an observation time and how it varies with s [84].

The decay of the walkers depends on the overlap between the tilted trajectory ensemble and that generated from the controlled dynamics. Slower decay will result when the control dynamics generates a trajectory ensemble that is close, in this sense, to the tilted trajectory ensemble. This behavior is analogous to other approximate guiding function based importance sampling, such as that arrived by iterative feedback [86] or analytical approximation [89]. These effects are seen in the hyperuniform phase as well, albeit the decay of walkers in the ordinary cloning algorithm is less drastic, and so is the improvement by incorporating the guiding forces. The improvement in statistical efficiency upon including the optimized forces is not restricted to the cloning algorithm, and could be analogously adopted within transition path sampling [89] or forward flux sampling [102].

2.6 Current uctuations in interacting overdamped system

The Asymmetric Simple Exclusion Process (ASEP) has been extensive studied as a model for many-body systems in nonequilibrium statistical physics [139]. ASEP is de ned on a one-dimensional lattice where each site can accommodate at most one particle, and the particles hop forward or backward to neighboring sites with asymmetric rates without overtaking each other. Boundary-driven nonequilibrium steady-states in ASEP have been well-characterized to contain several dynamical phases that arise from both local density and collective current order-parameters [140, 141]. Conditioning trajectory ensembles on large deviations of currents have been previously shown to induce long-range e ective forces that decay logarithmically [95, 142]. Specialized tools using matrix product states and other tensor network techniques have also recently been developed to sample rare current uctuations and characterize the dynamical phase behavior in ASEP and its generalization to two dimensions [121, 143].

Despite being de ned only on a discrete lattice in one spatial dimension, ASEP has been widely used to study the phenomenology of the nonequilibrium stochastic transport of biomolecular cargo inside the cell [144]. The movement of ribosomes on RNA during protein synthesis [145], the active tra c of load-carrying molecular motors on microtubules [146] and the single- le transport of ions and small molecules through ion-channels [147] have been mapped into ASEP to characterize their dynamical regimes and their energetic e ciency. The mapping between ASEP and these systems comprising of continuous space hopping between a sequence of metastable wells becomes accurate in the limit that barrier heights are large compared to thermal energy. However, rare current uctuations are often collective and involve correlated hops involving neighboring multiple particles, for which the introduction of in nitesimal uctuations in inter-particle distances may change the phase behavior. Indeed, dynamical phases featuring current reversals have been previously observed in lattice models but not found in its continuous space analogue in a model for stochastic pumps [148, 149]. To better understand the continuous space generalization of ASEP and to characterize its driven phase behavior, Brownian ASEP (BASEP) has been recently introduced as a model for single- le transport [150, 151]. BASEP consists of hard volume-excluding spheres di using on a one-dimensional periodic potential driven by a constant external force. This model features rich structural phenomenogy depending on the ux of incoming and outgoing particles, average density and the commensurability of particle diameters with the wavelength of the periodic potential. But the rare current uctuations, response and the dynamical phase behavior of BASEP trajectory ensembles have not been studied and compared against those of ASEP to determine the latter's applicability in studying real space dynamics.

In this section we have used our variational algorithm to construct the large deviation functions for currents in a soft-sphere analogue of the BASEP. We have generalized our model we used in Section 2.4 for describing multiple interacting particles in a periodic potential.

We consider N di usive particles governed by the overdamped Langevin equation,

$$\dot{x}_i = F_i(x^N) + \quad _i \tag{2.47}$$

in a one-dimensional box of length $L = \quad$, where $\quad$ is the packing fraction of the system, F_i is the total force on the i-th particle, and $\quad _i$ is a Gaussian white noise with zero mean and variance $2k_BT$. Each particle experiences a periodic potential $V(x_i) = V_0 \cos(x_i)$, a constant external driving force F_{ext}, and a pairwise WCA repulsion $V_{WCA}(x_i \quad x_j)$ from all the other particles, similar to the previous section in Eq. 2.40. We consider only the case $= 2 \quad$, such that each potential well can accommodate at most one particle without high energetic cost. We have de ned a collective current J over a trajectory $X(t_f)$ of duration t_f as

$$J = \frac{1}{t_f} \int_0^{t_f} dt \sum_i \dot{x}_i dt \tag{2.48}$$

The collective current can thus be also interpreted as the drift in the center of mass of the particles. A peculiarity of working with a continuous description of space is that if $V_0 = 0$, momentum is well-de ned and conserved modulo the e ect of the thermal noise. The current in the $V_0 = 0$ limit is exactly as if each particle is being driven independently by the external force F_{ext}. This means that the e ect of having a high density of particles only becomes manifest in the statistics of average collective current through a non-zero external potential energy function. When V_0 is non-zero and comparable to k_BT, at high $\quad$ non-zero current is generated from sequential correlated hops from collections of adjacent particles. Such current-carrying mechanism is shown in Fig. 2.10(a) via the collection of particle coordinates $\{x_i\}$ as a function of time t at parameters $= 0.5$, $k_BT = 0.5$, $V_0 = 1$ and $F_{ext} = 1.5$.

For a system of $N = 100$ particles, we compute the SCGF (s) associated with biasing this time-averaged current, by applying our variational algorithm to compute an optimal force. As basis functions we choose a combination of one-body periodic functions and two-body pairwise functions, $u_i = F_i(x) + u_i^{(1)}(x) + \sum_{j \neq i} u_i^{(2)}(x_i \quad x_j)$, that are described by Equations 2.34 and 2.44. We study the convergence of the variational estimate for the SCGF at three values of packing fraction, a dilute limit, an intermediate regime and a dense limit where every well is occupied. We have learnt this force with $M_1 = 5$ and $M_3 = 5$ basis functions for biass in the range of $[\quad 2; 5]$. We have worked with units $k_BT = 0.5$, $= 1$ and our natural time unit as $t \quad = 2 \quad ^2 k_BT = N \quad ^2$ which evaluates to the same timescale de ned in Section 2.4. All observables have been reported in dimensionless units following these de nitions. We have also chosen $V_0 = 1$ and the external driving force $F_{ext} = 1$. For all the simulations, the timestep was chosen to be 0.001 natural time units and an Euler scheme was used to integrate the overdamped equation of motion. For each iterative step during the optimization, a trajectory of duration $5 \quad 10^7$ units was simulated. Similar to our computations with the single overdamped particle, the system was allowed to come to a steady-state for the rst half of the trajectory, and the time-averaged gradients were

Figure 2.10: a) Snapshot of the many-particle trajectory ensemble at half- lling, demonstrating correlated sequential hops that lead to non-zero current. b) Variational SCGF estimators for $= 0;\ 0.5$ and 1.0. Dotted vertical line represents Gallavotti-Cohen symmetry point at $s = 1.5$. Green dashed line represents the segment of SCGF for $= 1,\ s < 1.5$ re ected to the region $s > 1.5$.

computed only with the second half of the trajectories. For implementing Eq. 2.18, we integrated the correlation function up to $t = 100$. We chose optimization learning rates $= 0.1;\ = 0.0$.

In Fig. 2.10(b) we have shown the variational SCGF estimates obtained for our system at three di erent densities. We nd that the SCGF follows the similar locked and unlocked regimes we had seen for the single particle case, which is identical to the dilute limit. The SCGF does not change much from the dilute limit to $= 0.5$, and only changes appreciably near full packing At a high density, the biased current value in the locked part of the SCGF reduces from the low-density limit. This is manifested as a sharp change at $s = 0$, which is arising from a diverging second cumulant of current, the latter being proportional to the second derivative of the SCGF. For any density, the SCGF should respect a Gallavotti-Cohen re ection symmetry about the line $s = 1.5$. This is respected for small and intermediate densities, but when the box is fully packed, we see this symmetry not being bveyed by the variational estimator, which implies that the optimal force cannot be adequately represented within our basis set ansatz. Nevertheless, given the variational principle and that the symmetry must hold for the correct (s), we get an estimate of the error being made by re ecting one-half of the SCGF on the other half. This construction is shown as the green dashed line

in Fig. 2.10(b).

Despite the percentage error being small, we could not correct the variational estimate with a converging cumulant expansion or a guided cloning procedure with statistics from upto 10^9 timesteps. This is due to the uctuations in the system being extremely rare, and also hints at the converged forces being quite far from the actual optimal force in the system. Looking closer into the one-body and two-body components of our optimized force, we nd that for all values of s the two-body component optimizes to a negligible value compared to the one-body parts. This suggests that a pairwise two-body force ansatz is a poor approximation for the many-body nature of the true optimal force. Improved representations of many-body e ective forces through neural networks [152] might be promising as better basis sets for optimal forces generating rare current uctuations in this system.

2.7 Conclusion

We have developed a variational algorithm to compute optimal control forces for Langevin models driven into nonequilibrium steady-states. We have used the control forces to sample rare uctuations in time integrated dynamical observables like current and activity, in order to compute large deviation functions, and shown that they can be used to improve the e ciency of the cloning algorithm. Our variational algorithm, along with the correction of the systematic error with the cumulant expansion, has improved scaling properties compared to trajectory ensemble methods, and can be useful in dealing with many-particle chemical or biological systems.

Though we worked with Langevin models of structureless particles, the algorithm is straightforward to generalize to higher dimensions, where optimal control forces might have signi cant rotational components. It can also be extended to lattice models, where the rate matrix has to be expressed in a variational ansatz. A system modeled by a di erent stochastic equation of motion, like that employing an Andersen thermostat [42] or quantum trajectory-based approaches [153, 154], can also be treated through this algorithm by changing only the functional forms of the path-actions provided a Doob transformation exists.

The versatility of the variational algorithm allows for its use with di erent force ansatzes. In the activity-biased system, using a low-rank approximation for a many-body optimal control force was su ciently accurate. However in cases where the control force is not expressible in a simple functional form or even as a many-body expansion, machine learning using arti cial neural networks could be used to approximate it. The variational algorithm relies on evaluating functional derivatives of the force with respect to the parameters, which can be automated with autodi erentiation algorithms [155], as has already been demonstrated in equilibrium free energy calculations [156]. The use of techniques developed in this paper can aid the formulation of such optimization algorithms in the future. Additionally, this algorithm can be used for model reduction in high-dimensional systems [157], and hence to extend Variational Force-Matching and Ultra Coarse Graining algorithms [158, 159, 160] out

of equilibrium, so that biomolecular and other soft matter systems can be simulated over large length and time scales with e ective forces in nonequilibrium steady-states.

Lastly, this framework of solving the optimal forces can tackle inverse-design problems out of equilibrium. Various inverse-design algorithms have been proposed that can obtain optimal forces to rationalize materials design with targeted properties and to guide directed self-assembly of smaller objects [161, 162]. Our variational algorithm can be used to obtain optimal forces suitable for targeted assembly or tailored particle distributions when nonequilibrium driving forces are present, and hence can be used to characterize and predict dynamical phases in new functional materials.

Chapter 3

Inverse design of nonequilibrium colloidal self-assembly

The self-assembly of soft and biological matter out of equilibrium can result in novel structures and dynamical responses not constrained by thermodynamic considerations [163, 164, 165, 166, 167, 33]. The microscopic violation of detailed balance in such systems can be used to design a wide range of functional materials with enhanced thermomechanical, optoelectronic or drug-delivery properties [168, 169, 170]. Predictive inverse design to drive the assembly of target dissipative structures requires a dynamical description of the system [171, 172, 173, 48, 174]. In this chapter we apply our variational algorithm to automate the discovery of inverse design principles for colloidal self assembly in a nonequilibrium steady-state in molecular dynamics simulations. The algorithm uses a variational principle arising from rare dynamical uctuations of the system in a trajectory ensemble, and optimizes the yield of target clusters or their rate of formation with statistically estimated explicit gradients in the design parameter space. We demonstrate the performance of this algorithm by obtaining optimal design principles for the self-assembly of DNA-labeled colloids [175] driven out of equilibrium by a shear ow. We expect that the ability to uncover general optimal inverse design principles away from equilibrium will enable bottom-up synthesis of new materials and elucidate the processes encoding structure in biological contexts [176, 177, 18].

Self-assembly of nanoscale building blocks is increasingly used to engineer functional materials with novel properties arising from their complex nanostructures. Colloidal systems o er a versatile paradigm for inverse design towards a desired target structure due to the independent tunability of shape, valency and assembly environment [178]. In thermodynamic equilibrium, stabilizing a target structure amounts to lowering its free energy, which is typically achieved by increasing both the interaction strength and speci city. This principle has been exploited to achieve the self-assembly of a variety of clusters and superlattices from colloids and nanocrystals with crystal facets decorated with organic ligands or DNA [179, 180, 8]. However, employing these principles in practice requires mitigating dynamical e ects like slow coarsening and kinetic trapping [173, 172]. Optimal forces for self-assembly must achieve a trade-o between slow relaxation at high interaction strengths, and slow

growth at high interaction speci city [181, 182]. Self-assembly of colloids and biomolecules in nonequilibrium steady-states provide a route to decouple kinetics from stability and mitigate this tradeo . Directive self-assembly has been achieved by driving the system with a constant supply of chemical fuel, or by applying external elds [183, 184, 185, 186]. However, the design of such systems must confront the continuous supply of energy necessary to prevent the system from relaxing to equilibrium. Existing computational methods to discover inverse design principles for nonequilibrium self assembly are limited due to the con gurational probability not following the Boltzmann distribution and the corresponding variational structure a orded by the free energy no longer being valid under such dissipative conditions.

Recent advances in the theoretical treatment of the stochastic thermodynamics of nonequilibrium steady-states have made possible a trajectory ensemble description of self-assembly, treating structure and dynamics on an equal statistical footing [27, 19]. This has enabled basic principles governing assembly away from equilibrium to be formulated [22, 74]. In this work we develop a perspective and accompanying numerical technique based on these insights. Rather than considering the probability of observing a state and tuning its associated free energy, we consider the likelihood that a trajectory forms a speci c structure as quanti- ed by a stochastic action, and how that action is changed by modifying the intermolecular and applied forces. We show that uctuations around a nonequilibrium steady-state encode the susceptibility of a system to assemble, in a manner analogous to a uctuation-dissipation relationship. Further, optimal forces that assemble a target structure or maximize reactive ux while minimizing the change to the stochastic action satisfy a variational principle [63, 111]. We extend and apply our optimization algorithm to solve this variational expression and compute the optimal control force to sample rare dynamical phases [117]. We show that this algorithm can be used to solve the inverse design problem, deciphering how rare uctuations encode stability away from equilibrium.

We outline below an inverse design algorithm for the self-assembly of sheared DNA-coated colloids into di erent target nanoclusters. The algorithm is based on a variational principle relating rare uctuations in an ensemble of trajectories conditioned on evolving a target structure, to e ective forces achieving the target as the typical dynamical state. Working with a trajectory ensemble, where the probability distribution is known, circumvents the di culty of not knowing the distribution of con gurations within a nonequilibrium steady-state. To solve the variational problem, we have used generalized response relations for the gradients of the steady-state trajectory probability to a change in the inter-particle and externally driven forces.[1]

[1] Most of the content of this chapter was originally part of the publication [187].

3.1 Coarse-grained model for DNA-labeled colloids

For concreteness, we consider a model N of colloidal particles in a cubic box of length L, evolving with an overdamped Langevin equation of the form,

$$\dot{\underline{r}}_i = u_i + \xi_i \tag{3.1}$$

where $\dot{\underline{r}}_i$ are time derivatives of the coordinates of the i-th particle and u_i are the forces acting on it. The friction coefficient of the colloids with the thermal bath is denoted γ and ξ_i are Gaussian white noise that satisfy

$$\langle \xi_i(t)\rangle = 0\,; \quad \langle \xi_i(t)\,\xi_j(t')\rangle = 2\,\gamma\, k_B T I_{3\times 3}\,\delta_{ij}\,\delta(t-t') \tag{3.2}$$

where $I_{3\times 3}$ is the 3×3 identity matrix and $k_B T$ is Boltzmann's constant times the temperature. The angular brackets denote an averaging operation over the random noise distribution. As we consider dynamics in the presence of a shear flow, we use Lees Edwards boundary conditions [188].

We use an ansatz of DNA-labeled spherical isotropic colloids as programmable building blocks for self-assembly. The interaction between these colloids, mediated by the DNA molecules attached to their surface, consists of a volume-exclusion repulsion and a short-range attraction [189]. The effective interaction strengths and the pairwise specificity can be independently tuned by varying the sequences of the grafted DNA molecules. During self-assembly, the short-range forces generate a competition between local and global order that leads to frustration and unique phase behavior and dynamical effects [190, 191]. This system has been computationally and experimentally demonstrated to form finite nanoclusters with specific target structures [192, 193]. The high-dimensional design space has the possibility to offer multiple pathways to stabilize any cluster out of the many nearly degenerate states formed without the specificity of the DNA-mediated attraction. To illustrate the performance of the variational algorithm, we consider the nonequilibrium self-assembly of 21 such rigid and nonrigid clusters, some examples of which are demonstrated in Figure 3.1a.

We examine the self-assembly of these colloidal particles under a constant linear shear flow. Shear flows are known phenomenologically to alter the stability of compact and extended colloidal structures [194, 195]. A recent paradigm of colloidal assembly being increasingly explored is that in a microfluidic device, where the confining walls generate a strong shear on the assembling clusters [196, 197]. This system offers a canonical nonequilibrium setting to explore inverse design principles. Taken together, the forces acting on the i-th colloid are

$$u_i = f_i^S(r_i) - \nabla_{r_i} \sum_{j\neq i} V(r_{ij}) \tag{3.3}$$

where $V(r) = V_{WCA}(r) + V_{Morse}(r)$ and $V_{WCA}(r_{ij})$ is a WCA pair potential representing the volume exclusion interactions, and $V_{Morse}(r_{ij})$ denotes the DNA-mediated short-range pairwise attraction. The force due to a shear flow, $f_i^S(r_i)$, has the form

$$f_i^S(r_i) = f\, z_i\, \hat{x} \tag{3.4}$$

Figure 3.1: Model details for self-assembly of DNA-labeled colloids. a) Examples of nite rigid and nonrigid nanoclusters for which we have studied design principles, along with the corresponding point groups for molecular symmetries denoted underneath. b) Graphical forms of the potential energy functions, the WCA potential (blue) and the WCA and Morse potential combined (red) for $D_{ij} = 10k_BT$. The orange dashed line denotes the bond cuto $r_b = 1:35$ and the black dashed line denotes the potential cuto $r_c = 2:12$.

which has magnitude f and generates a constant gradient of the x component of the velocity along the z direction. The WCA pair potential has the functional form

$$V_{WCA}(r_{ij}) = 4 \left[\frac{}{r_{ij}}^{12} \quad \frac{}{r_{ij}}^{6} \right] + \quad ; \quad r_{ij} < 2^{1=6}$$

$$= 0 \; ; \quad r_{ij} \quad 2^{1=6} \tag{3.5}$$

with particle diameter and energy scale. The attractive Morse potential has the functional form

$$V_{Morse}(r_{ij}) = D_{ij} \quad e^{2\,(r_{ij} \quad 2^{1=6})} \quad 2e^{(r_{ij} \quad 2^{1=6})} \tag{3.6}$$

where D_{ij} is the magnitude of the bond energy and determines its width.

We work in units of $k_BT = 1$, $= 1$ and $= 1$. The natural time scale with these units is $t_0 = \quad ^2=k_BT$ and times are expressed in these units throughout. We set $= 10k_BT$

and $^{1} = = 10$. The attractive energy scale D_{ij} and the shear ow rate f are tuned as variational parameters to induce self-assembly. They have been restricted to vary within the range $0 \quad D_{ij} \quad 10 k_B T$ and $0 \quad f \quad 50 k_B T = {}^{2}$ to avoid large relaxation times and to stay within the overdamped regime. Figure 3.1b shows the potentials for the inter-particle interactions. The Morse potential and its force have both been truncated and shifted, using the Shifted Forces approximation [198], to decay smoothly to zero at $r_c = 2.12$.

In order to avoid nite size e ects in the formation of small clusters, we study a low packing fraction of $= 0.01$. We use a rst order Euler discretization for the equation of motion in Eq. 3.1. Since the potentials in Eq. 3.3 are narrow and short-range, we have to use a small timestep of $5 \quad 10^{5} t_0$ in order to sample the potentials accurately. We have used trajectories of length ranging from $= t_0 = 2.5 \quad 10^{3}$ to 10^{4}.

3.2 Automated inverse design algorithm

Variational principle

In order to uncover design principles for self-assembly, we consider the task of nding the set of forces that ful ll the condition of assembling a target structure as the typical state of the system in the long time limit. Such tasks in stochastic dynamics are generalizations of Brownian bridges and known to have unique solutions [111]. They have played an important role recently in the application of large deviation theory to physical systems driven far from equilibrium [86, 81, 93, 199, 117].

We start by de ning an observable A as a time averaged indicator function for a target cluster,

$$A \; [r^N(t)] = \frac{1}{} \int_0 \; 1[r^N(t)] \, dt \tag{3.7}$$

where $1[r^N(t)] = 1$ for a con guration satisfying a geometric criterion consistent with a target cluster and $1[r^N(t)] = 0$ otherwise, for each time t along a trajectory $r^N(t)$ of total duration . The average value of the observable quanti es the yield of the target cluster. For all colloidal clusters considered, 1 is computed by constructing a bond-connectivity matrix. A cuto of $r_b = 1.35$ has been used to de ne a bond between two particles. Indicator functions for rigid target clusters are then uniquely determined by permutation-invariant measures of this connectivity matrix [200]. For nonrigid target clusters, along with the bond-connectivity matrix, we additionally consider measures of the geometry of the cluster for de ning the indicator function.

Rather than considering trajectories conditioned on a particular value of A directly, which is numerically cumbersome, we work within an ensemble equivalent representation [118]. Using a counting parameter , we can statistically bias a system towards a particular value of A within a nonequilibrium steady state. The cumulant generating function (s) is the partition function associated with the trajectory ensemble under the statistical action

of s,

$$(s) = \lim_{!1} \frac{1}{} \ln \; e^{ s A} \;_0 \tag{3.8}$$

where the angular brackets denote a path average over trajectory probability $P_0[r^N(t)]$, as

$$e^{ s A} \;_0 = \int D[r^N(t)] \exp \; s A \; [r^N(t)] \; P_0[r^N(t)] \tag{3.9}$$

The subscript 0 refers to the average being computed in a reference ensemble where the particles do not typically show the desired self-assembly behavior. For this reference system we have chosen an equilibrium ensemble of colloids interacting only with the WCA repulsive forces,i.e., $D_{ij} = f = 0$, such that $u_i = r r_i \sum_{j \neq i} V_{WCA}(r_{ij})$ which is denoted as $F_i^{WCA}(r^N)$.

When the optimizable parameters are tuned to vary u, the trajectory probability $P_0[r^N(t)]$ changes to $P_u[r^N(t)]$. The cumulant generating function can be estimated in the modi ed ensemble as,

$$\begin{aligned} (s) &= \lim_{!1} \frac{1}{} \ln \int D[r^N(t)] e^{ s A} \frac{P_0[r^N(t)]}{P_u[r^N(t)]} P_u[r^N(t)] \\ &= \lim_{!1} \frac{1}{} \ln \; e^{ s A \; + \; S[u]} \;_u \end{aligned} \tag{3.10}$$

where the functional form of the relative action $S[u]$ can be derived from Onsager-Machlup theory [45],

$$\begin{aligned} S[u] &= S[u] \quad S[F^{WCA}] \\ &= \int_0 \sum_{i=1}^N \frac{(\underline{r}_i \quad u_i)^2 \quad (\underline{r}_i \quad F_i^{WCA})^2}{4 k_B T} dt \end{aligned} \tag{3.11}$$

with the integral being computed in the Ito sense. This change of measure analogous to a Girsanov transform [63] relates the original likelihood of self-assembly in the reference ensemble to the ensemble under the control force.

Since the exponential is a convex function, we apply Jensen's inequality to Eq. 3.10

$$(s) \quad \lim_{!1} \frac{1}{} h \; s A \; + \; S[u] i_u \tag{3.12}$$

to obtain a variational expression for the cumulant generating function. In the long time limit for $! 1$, we can replace trajectory averages with static averages and simplify the relative action using the equation of motion for $\underline{r}_i$. Hence we arrive at our nal variational expression,

$$(s) \quad h \; [u] i_u = \left\langle s 1(r^N) \sum_{i=1}^N \frac{(u_i \quad F_i^{WCA})^2}{4 k_B T} \right\rangle_u \tag{3.13}$$

where $h[u] i_u$ is the target function to optimize.

For a bounded observable like the indicator function, a large negative value of s enforces the desired conditioning. The problem of saturating the variational inequality is known to have a unique solution when $u(r^N)$ can take all possible functional forms, the optimal force being a generalization of Doob's h-transform [111]. Optimizing $\langle g[u]\rangle_u$ will lead to a set of many-body forces for assembling target clusters in high yield. In practice the use of only one-body and two-body forces in Eqs. 3.3-3.6 need not saturate the inequality. The second term in the variational expression is associated with the Kullback-Leibler divergence between the reference and conditioned trajectory ensembles. This term enforces the smallest excess force out of all possible control forces, and thus acts as a regularizer in the optimization process. While the solution to Eq. 3.13 uniquely selects the force that shows fluctuations closest to rare fluctuations in the original ensemble, it is not a unique inverse design criterion and alternatives can in principle be constructed [201]. However, the optimization scheme that we construct in the next section can be generally extended to other functional forms of regularizers.

Stochastic gradient descent

To numerically optimize Eq. 3.13, we derive explicit gradients of the variational estimator $\langle h[u]\rangle_u$, using an algorithm that we have previously employed to estimate large deviation functions in nonequilibrium steady-states [117]. The general form of the gradient with respect to any variational parameter $c \in \{D_{ij}; f\}$ is

$$\frac{\partial \langle h[u]\rangle_u}{\partial c} = \frac{\partial \langle \dot{S}[u]\rangle_u}{\partial c} \int_0^{\infty} \langle \dot{S}(t)\, \dot{S}(0)\rangle_u \, dt \qquad (3.14)$$

where $\dot{S}[u]$ is the time derivative of the action $S[u]$. Equation 3.14 is a generalized fluctuation-dissipation relation for a nonequilibrium response in the design parameter space. For computational purposes, we approximate the gradient expression by integrating the correlation function in the second term up to a fixed large time interval $t = 5t_0$. Due to the small density, we have to use a low variance estimate for the explicit gradient in Eq. 3.14 for the specific case of optimizing the shear flow rate f as described later.

The corresponding variational algorithm consists of a stochastic gradient descent optimization [202] for a large negative value of s in Eq. 3.13. Starting from an initial point in parameter space $\{D_{ij}; f\}$, we simulate the dynamics of the system using Eq. 3.1, and after relaxation into a steady state, statistically estimate the explicit gradient of the variational estimator, Eq. 3.14. We perform stochastic gradient descent updating all variational parameters c at every step of the optimization, with the update rule at the n-th step being

$$c_{n+1} = c_n + \eta_n \frac{\partial \langle h[u]\rangle_u}{\partial c}\bigg|_{c_n} \qquad (3.15)$$

where the stochastic gradients are evaluated within a steady-state with the current value of the parameters, and η_n is the learning rate for any of the c parameters in the n-th optimiza-

Figure 3.2: Optimization procedure for an O_h cluster. a) The convergence of the yield with increasing number of optimization steps. Di erent colors represent varying learning rates. b) Convergence of D_{ij} for the green yield curve in (a). (Inset) Bond structure and the corresponding MAD D_{ij} matrix for the O_h cluster. Blue and white elements in the matrix denote bonds with $D_{ij} = 10 k_B T$ and $D_{ij} = 0$, respectively.

tion step. The level of noisy uctuations in each parameter during the optimization process can be tuned independently through the corresponding learning rates. If the variational surface changes sharply, the learning rate has to be decreased with increasing anneal to the optimal solution basin. The learning rates have also been chosen individually for each example such that in each optimization step, the rate of change of $D_{ij} = k_B T$ is in the range $[0.1; 0.5]$ and that of $f^2 = k_B T$ is in the range $[1; 5]$.

Convergence and choice of s

To illustrate the performance of the optimization algorithm, we study the assembly of 6 particles into an octahedral (O_h) target cluster. An octahedron is not the highest yield cluster formed in a system of 6 hard sphere colloids with in nitely short-range attractions [192], and is formed in only 6% yield with strong, nonspeci c interactions. Figure 3.2a shows the yield as a function of optimization steps with di erent learning rates and di erent trajectory noise histories. For all these examples, the yield is optimized over multiple orders of magnitude with the nal converged value being close to 100%. This change of the order parameter over several orders of magnitude arises from the observable being de ned as the probability of forming the target cluster, and in cases where the change is more drastic, would necessitate the use of two di erent learning rates in Eq. 3.15. At a constant learning rate, the learning curves show two distinct regions, such that a gradual rise in yield is followed by

Figure 3.3: Stability of the high-yield solution with s. a) Convergence of the yield starting from the $10k_BT$ MA solution. Di erent colors represent varying values of st_0 in the range [4 10^4; 10^3], with the converged value of the yield increasing monotonically with $|s|$. b) Blue crosses with errorbars are the converged bond energy of the MA solution at varying values of s. Red line is a linear t.

a rapid convergence to the saturation value.

Figure 3.2b shows the convergence of D_{ij} for one of the optimization runs. For 6 particles there are 15 distinct interactions, all of which are optimized. The starting point is a non-speci c attraction $D_{ij} = 4k_BT$ for all ij pairs. The optimization curve shows two regions, an initial spreading of the D_{ij} values followed by a rapid permutation symmetry breaking and a clear segregation of the 15 interactions into 12 attractive and 3 repulsive parameters. The 12 attractive interactions are all statistically equal, as are the 3 repulsive interactions. The attractive interactions correspond to the 12 bonds in the connectivity matrix for the octahedron. The symmetry breaking is spontaneous and is aided by the initial noisy uctuations during optimization. Di erent noise histories in the trajectory lead to a symmetry breaking for which di erent sets of D_{ij} parameters become attractive or repulsive. For the nite clusters considered, this symmetry breaking is general. We refer to the specific D_{ij} solutions for the optimal yield of a target cluster as an *alphabet*, and the particular D_{ij} in which there is a pairwise attractive interaction for every contact in the target structure as a Maximal Alphabet (MA). This strategy has been previously shown to be e ective in the self-assembly of short-range interacting colloids into small clusters [175, 193].

For the octahedral cluster, we have studied the stability of the MA solution for varying values of s. We nd a bimodal structure of the variational surface, with the algorithm converging to either a MA solution or a trivial solution $u = F^{WCA}$, depending on the value of s. Figure 3.3a demonstrates the convergence of the octahedral yield with di erent values of st_0 varying in the range [4 10^4; 10^3], starting from an MA solution with $10k_BT$ bond

Figure 3.4: Design principles for rigid clusters. a) Bond structure of clusters (i-ix), along with their corresponding point groups, optimal yields and their converged alphabets on a color scheme indicated by the colorbar at the top. b) Yield as a function of a nonspecific attraction $D_{ij} = D$ for fixed shear $f = 0$. c) Yield as a function of the shear rate f for fixed optimal alphabets. The colors of the points in b) and c) correspond to the colors of the clusters in a).

energies and a yield of 100%. For moderate but decreasing values of s, the algorithm remains stable in the MA solution, but with monotonically decreasing yields and bond energies. For less negative values of s than a critical value of $s_c = 5 \quad 10^3/t_0$, the MA solution becomes unstable and the algorithm finds the $u = F^{WCA}$ solution. Rather than optimizing the yield in Eq. 3.13, at small values of s the second term is optimized. For some moderate s values, the MA basin is only a local optimum and the crossover behavior shows s-dependent hysteresis. We find this bistablity of the variational surface to be generic.

Asymptotic dependence of solution on s

We analytically solve for the asymptotic dependence of the bond energy D as a function of the bias s, for the formation of one bond, independent from the dynamics of the other bonds. We can represent this simplified system as an equilibrium, two state Markov model with the rates of transition between the bonded and unbonded states k_b and k_u, respectively. These rates are determined by their mean $= (k_u + k_b)/2$ and their ratio $k_u/k_b = \exp(\quad F/k_B T)$,

where F is the free energy di erence of the two states.

The cumulant generating function associated with uctuations in the bonded state sati es a eigenvalue equation of the form,

$$W_s r_s = (\quad s) r_s \tag{3.16}$$

where r_s is a right eigenvecture, and operator W_s is given by

$$W_s = \begin{matrix} k_u & s & k_u \\ & k_b & k_b \end{matrix} \tag{3.17}$$

which is equal to the adjoint of the transition rate matrix when $s = 0$. For the optimal rates that generate the statistics equivalent to this rate matrix, we need the Doob's transform of the matrix [63]. For this purpose we diagonalize the matrix to nd the eigenvector corresponding to the dominant eigenvalue as $r(_1; 1)$ where

$$r_1 = \frac{e^{\frac{F}{k_B T}} \quad (\quad s + 2 \quad) + e^{\frac{F}{k_B T}} (\quad s \quad 2 \quad)}{q \; e^{2\frac{F}{k_B T}} (\quad s \quad 2 \quad)^2 + (\quad s + 2 \quad)^2 + 2 e^{\frac{F}{k_B T}} (s^2 + 4 \quad ^2)} \tag{3.18}$$

and k_u and k_b have been rewritten with F and .

The modi ed rates that generate the optimal dynamics are given by $\tilde{k}_u = k_u = r_1 t_0$ and $\tilde{k}_b = k_b r_1 t_0$. The modi ed free energy di erence corresponding to these rates is $\tilde{F} = k_B T \ln(\tilde{k}_u = \tilde{k}_b)$. Using Eq. (3.18), in the $s \; ! \; 1$ limit the optimal free energy goes as

$$\tilde{F} \quad F \quad 2k_B T \ln(jst_0 j) \tag{3.19}$$

In this limit, the free energy is dominated by the negative of the bond-formation energy D, and hence the latter is asymptotically $D \quad 2k_B T \ln(jst_0 j)$.

Figure 3.3b shows the dependence of the average optimized bond energies of the converged MA solutions from the previous gure, as a function of varying s. This illustrates the depth of the MA basin in the parameter space. In the limit that each bond is formed independently, D_s is expected to asymptotically vary as $2k_B T \ln(jst_0 j)$. Over the range of s considered, we nd a logarithmic dependence but with a di erent coe cient, $1{:}3 \ln(jst_0 j)$. This suggests that the free energy is approximately pairwise additive.

In the limit of large negative s, which in practice is chosen such that the estimate of the rst term in Eq. 3.13 is at least an order of magnitude larger than the negative second term, the variational algorithm can be used to automate the discovery of optimal forces for the self-assembly of clusters of arbitrary shapes and sizes, in a nonequilibrium steady-state. The optimal forces stabilize the target clusters in an arbitrary ensemble without accounting for the dynamics of transient relaxation towards its steady-state. For xed number of tunable parameters, the computational cost scales linearly with system size since the only bottleneck

Figure 3.5: Design principles for nonrigid clusters. a) Bond structure of clusters (x-xv), along with their corresponding point groups, optimal yields and their converged alphabets on the same color scheme that was used in Fig. 3.4. b) Yield as a function of the shear rate for D_{ij} xed at the optimal alphabet. The colors of the points correspond to the clusters in a).

is propagating a steady-state trajectory long enough to compute statistically converged gradients. The algorithm also scales linearly with the number of variational parameters, but with a small proportionality constant as all the gradients are estimated from the same trajectory. The use of the statistically estimated gradients signi cantly lowers the computational cost in contrast to numerically estimating the gradients from nite di erence techniques by propagating multiple trajectories at di erent points in the parameter space. We next use our variational algorithm to study and rationalize the optimal design principles for a collection of rigid and nonrigid clusters.

3.3 Design principles for yields of rigid and nonrigid clusters

We have investigated the formation of small low-energy rigid and nonrigid clusters with 6,7 or 8 particles. We discover distinct design principles of these clusters and rationalize our ndings by analyzing the response function of yield to the shear ow rate. We also demonstrate that the variational algorithm can obtain high yield optimal solutions even with constraints imposed on the total number of experimentally realizable design parameters. The design principles we obtain are expected to be general for the nonequilibrium self-assembly of short-range interacting colloids in a sheared steady-state.

Rigid clusters

We study the formation of a family of rigid clusters that are all known to be the lowest energy structures for systems of hard sphere colloids with identical in nitely short-range attractions. These clusters have previously been systematically enumerated and tabulated [203, 204], and their free-energy landscapes have been theoretically and experimentally studied [192]. These nite clusters are the colloidal analogs to small molecules and have been shown to be involved in the controlled seeding and growth of polycrystalline phases and kinetically arrested gels [205, 206].

For each of these clusters, we optimize $\{D_{ij}; f\}$ to extremize the yield within this force ansatz. In the limit that the attractive interactions between the particles were in nitely short-range, there would be no internal low-energy distortion modes and the bond-connectivity matrix would correspond to a unique geometry. For our optimization, the indicator function refers to the corresponding bond-connectivity matrix conditions being satis ed. Figure. 3.4a summarizes the design principles discovered for these clusters. The point groups for the symmetries of each of these clusters have been indicated along with the highest yields obtained. For chiral C_n clusters, the yields are the racemic yield.

For each of these clusters, the corresponding optimal alphabet discovered by the variational algorithm has also been indicated. We nd that for clusters (i-iv), the optimal solution for D_{ij} is the MA. For the chiral C_2 cluster (v), the optimal alphabet is closely related to the MA but has a higher symmetry and is equivalent to a smaller 3 3 alphabet, while having the same yield. For clusters (vi-ix), all of which contain a radial 5-fold motif, the optimal yields are much less than 100%, and the optimal alphabets are not MA. The reason is the competition with structures with higher number of bonds. These lower energy structures would be geometrically unfeasible for in nitely short-range attractions, however in our model the short-range bonds have nonzero vibrations, which is su cient to lead to the formation of the extra bonds. Unlike MA, the optimal alphabet discovered by the variational algorithm penalizes the formation of these lower energy competing structures.

Figures 3.4b and 3.4c show two slices through the optimization landscape in the parameter space of $\{D_{ij}; f\}$. Figure 3.4b is a diagonal slice through D_{ij}, such that all $D_{ij} = D$, while xing $f = 0$. We nd that for the two 6 particle clusters (i) and (ii), there is a monotonic increase in yield with increasing D. This suggests that even when the attractive forces are not in nitely short-range, both of these clusters are energetically the most stable con gurations, and the only competing structures are higher in energy. When $D = 10k_BT$, the C_{2v} cluster (i) is formed with a yield of 93% compared to the 6% yield of the O_h cluster (ii), which is consistent with the stabilization due to the rotational entropy in the former [192]. All the other clusters (iii-ix) in Fig. 3.4b show a turnover in yield with increasing D. This is due to competing low energy structures that are formed at large enough nonspec D. We expect this design principle of a turnover in yield with increasing nonspeci c attractions to be general for larger clusters, since most larger clusters built from short-range interacting particles will contain the radial 5-fold motif [204]. The value of the attraction D at the yield turnover is determined by a competition between the energetic stabilization of the lower

energy structure and the destabilization of structures with missing bonds.

Figure 3.4c shows a slice through the optimization landscape with varying the shear rate $\dot{\gamma}$, while fixing D_{ij} at their optimal value found by the variational algorithm. We see that for this class of rigid clusters, the yield monotonically decreases with increasing shear rate. We expect this feature to be general for rigid clusters, since rigid clusters need no additional geometric stabilization that can be provided by shear, which only energetically destabilizes the bonds in the cluster. This perspective is confirmed in Sec. 3.3 using a linear response theory.

Nonrigid clusters

For our ansatz of short-range interacting colloids, clusters with N particles but fewer than $3N - 6$ bonds in total, and fewer than 3 bonds for every particle, are not minimally rigid in that they have zero energy deformation modes [203, 204]. These clusters are not formed in high yield as stable ground state structures in equilibrium. We have used the variational algorithm to uncover optimal nonequilibrium design principles for a family of such nonrigid clusters. The clusters we have considered belong to a family of planar two-dimensional structures known as polyiamonds. They have been shown to self-assemble from colloids in the presence of a spatial heterogeneity, like in hydrodynamically driven assembly of sedimenting colloids in the presence of a substrate [207, 208, 209]. Within our control force ansatz, we investigate whether the shear flow planes are sufficient to stabilize these clusters.

For the optimization process, the indicator function for the cluster yield has been defined using both the bond connectivity matrix and the flatness of the clusters, as discussed in the next subsection. Figure 3.5a shows the optimal design principles obtained for clusters (x-xiv). The variational algorithm converges on the MA solution for the D_{ij} parameters for all of these clusters. The yields, however, are not 100% due to contribution from competing buckled configurations where the polyiamonds fold over the triangular faces to form tetrahedral motifs. Moreover, with the MA fixed, the optimal yield is at a non-zero shear flow. Figure 3.5b shows the yield as a function of the shear rate for fixed optimal alphabets. The location of the turnover in yield depends on the competition between geometric stabilization of the planar structure from the shear flow lines and energetic destabilization of the bonds. This design principle of planar two-dimensional clusters being stabilized by a shear flow appears to be general, and stands in contrast to the rigid clusters which are strictly destabilized by shear.

Effect of shear on geometry of nonrigid clusters

Shear flow enhances the yield of nonrigid clusters (x-xv) by stabilizing a planar geometry and suppressing buckling modes. Here we have computed the probability distribution $P(\overline{\cos\ })$ of the average flatness, defined as $\overline{\cos\ } = [\cos(\ _{ABC}) + \cos(\ _{DEF})]/2$, conditioned on the correct bond connectivity matrix for the cluster being satisfied. The angles, $_{ABC}$ and $_{DEF}$ are defined in Fig. 3.6. The flatness is 1 for a perfectly planar geometry, but increases due

Figure 3.6: Flatness distribution for the geometry of the C_{2h} cluster (x), keeping the D_{ij} xed at the optimal solution and changing the shear rate f. (Inset) The atness is measured through the angles $_{ABC}$ and $_{DEF}$.

to buckling and bending of the nonrigid cluster. For de ning the indicator function for the nonrigid cluster, we used a atness cuto of $\overline{\cos}$ 0:8. We have looked at a population where the bond-connectivity matrix condition is satis ed but the atness is unconditioned. This is illustrated in Fig. 3.6 for the C_{2h} cluster (x). We have kept the D_{ij} forces xed at the optimized alphabet and plotted the distribution of atness at two values of shear, at equilibrium with $f = 0$ and also at $f = 20k_BT/^2$ which is close to the optimal value for highest yield. We nd that the planar geometry is a transient state at equilibrium, with the most probable states corresponding to the buckling of one or both of the angles $_{ABC}$ and $_{DEF}$. The shear ow destabilizes the buckled conformations and stabilizes the planar state instead, so that at $f = 20k_BT/^2$ the most probable conformation is the correct planar geometry.

Estimating explicit gradients in the low-density limit

Using Eq. 3.14, the second term in the explicit gradient with respect to the shear ow rate f is

$$\left\langle \left(\frac{\partial S[u]}{\partial u} \frac{\partial u}{\partial f} \right)(t) \right\rangle_{u(0)} = \frac{1}{2k_BT} \left\langle \left(\sum_{i=1}^{N} x_i z_i \right)(t) \right\rangle_{u(0)} \tag{3.20}$$

Figure 3.7: Response of yield to shear for a rigid and a nonrigid cluster, with D_{ij} xed at the corresponding optimal alphabets, and shear xed at $f = 5k_BT/ ^2$. a) Total correlation function (black squares) for an Q_n cluster (ii), and its torque (orange circles) and virial stress (red triangles) parts. b) The same correlation functions for the C_{2n} cluster (x).

where we have simpli ed the stochastic action using the equation of motion. Since z_n appears in the expression independently for each particle, and the density of the particles is vanishingly small, the z-di usion timescale of the cluster diverges, and the correlation function takes a long time to converge. Thus any gradient estimate we obtain by integrating the correlation function to a nite time t will contain a systematic error. In order to obtain an unbiased gradient, we recognize that in the large limit we are working in, the major part of (t) is from 1(t), which by our de nition depends only on the internal coordinates of the particle, and due to the spatial translation symmetry in our system, is decoupled from the center-of-mass di usion. This decoupling is directly expressed by a regrouping of terms in

the sum over particles,

$$
\begin{aligned}
\left\langle s^{*}\, 1(t) \sum_{i=1}^{N} \left(\dot{x}_i z_i \right)(0) \right\rangle_u &= \frac{s^{*}}{N} \left\langle 1(t) \left(\sum_{i=1}^{N} \dot{x}_i \right) \left(\sum_{i=1}^{N} z_i \right)(0) \right\rangle_u \\
&+ \frac{s^{*}}{N} \left\langle 1(t) \left(\sum_{i,j=1,\, i>j}^{N} (\dot{x}_i - \dot{x}_j)(z_i - z_j) \right)(0) \right\rangle_u
\end{aligned}
\tag{3.21}
$$

where in the first term the z coordinate of the center of mass has been explicitly factored out. We identify that the indicator function does not correlate with the center-of-mass motion and so the first term is 0. We use only the second term to approximately evaluate the gradients of Eq. 3.13 with respect to the shear flow rate $\dot{\gamma}$.

Response of structure to shear

The origin of the response of yield to shear flow is related to the relaxation dynamics of order parameter fluctuations in the unperturbed system. The response coefficients for rigid and nonrigid clusters can be understood using a generalized linear response theory [210, 36]. Keeping the D_{ij} parameters fixed, the linear response of the yield to a change in shear flow rate can be decomposed into two terms,

$$
\begin{aligned}
\frac{\partial \langle 1 \rangle_u}{\partial \dot{\gamma}} &= \frac{1}{k_B T} \int_0^{\infty} dt\, \langle h\, \dot{\gamma}(0)\, 1(t) \rangle_u \\
&+ \frac{1}{k_B T} \int_0^{\infty} dt\, \langle h\, M(0)\, 1(t) \rangle_u
\end{aligned}
\tag{3.22}
$$

where we have used the low variance estimator for the correlation function in 3.3. Here, is related to the dynamical torque acting on the cluster,

$$
= \frac{1}{2N} \sum_{i>j} [(\dot{\underline{x}}_i - \dot{\gamma} z_i) - (\dot{\underline{x}}_j - \dot{\gamma} z_j)](z_i - z_j) \tag{3.23}
$$

and M refers to the virial stress on the cluster due to internal forces,

$$
M = \frac{1}{2N} \sum_{i>j} [F_i^x - F_j^x](z_i - z_j) \tag{3.24}
$$

where $F_i = \nabla_{\underline{r}_i} \sum_{j \neq i} [V_{WCA}(r_{ij}) + V_{Morse}(r_{ij})]$ is the conservative force acting on the i-th particle. and M are time-reversal asymmetric and symmetric parts respectively [211] of

the full stochastic action gradient, $R = \quad + \quad M$. Decomposing in this form is necessary to preserve the deconvolution of the center-of-mass motion from the internal coordinates of the cluster [212].

This linear response function is 0 at equilibrium due to the spatial parity symmetry of the system. Hence we have characterized the di erent components of this correlation function at a small value of shear $f = 5k_BT= \quad ^2$, for the rigid cluster (ii) and the nonrigid cluster (x), xing D_{ij} to their corresponding MA interactions. The results are shown in Fig. 3.7. We nd that the component coming from the virial stress has opposite signs at small times for the rigid and the nonrigid cluster, which accounts for the opposite signs of the gradient of the yield. The tumbling motion of the clusters in a shear ow couples positively with the internal order parameter in the case of a nonrigid cluster and leads to an increase in yield with increasing shear ow rate at small values of shear. The shear ow planes function as a spatial heterogeneity that is generally a precondition for stabilizing these planar clusters during self-assembly. At large shear, the yields of both rigid and nonrigid clusters are decreased with increasing shear due to the larger anti-correlation between the torque and the indicator function.

These design principles in and out-of-equilibrium are general in their scope of applicability for small clusters formed by DNA-coated colloids. Nevertheless, a key limitation of this approach is the linear system size scaling of the number of di erent kinds of DNA-labeled colloids required in order to assemble a cluster, evident in the corresponding quadratic scaling of the number of variational parameters. We have addressed this limitation in the next section.

Smaller alphabets

Engineering a system with an extensive number of speci c interactions is di cult, even with DNA-coated colloids. It is advantageous in this regard to uncover *alphabets* that code for the minimal su cient interactions to stabilize a target structure, in such a way that does not increase with increasing system size. For example, polymers and crystals are macroscopic structures that can be assembled with a nite number of speci c interactions, as both only require a repeating microscopic number of components to be stabilized, either a sequence of monomers or a unit cell. For clusters that do not have a clear repeating unit, discovering optimal design principles with smaller *alphabets* is nontrivial.

We have studied this problem by considering the 6 di erent low-symmetry 8-particle clusters (xvi-xxi) shown in Fig. 3.8. For each of these clusters, there is no direct way to partition the interactions into 2 or 4 classes based on their bonding environment or symmetry. We have used the variational algorithm to optimize the yield of each of these clusters, with a 2 particle, 4 particle and 8 particle *alphabet*, in which D_{ij} has 3, 10 and 28 independent variational parameters respectively. Clusters (xvi-xix) have near 100% yields for the full sized *alphabet*. Clusters (xx) and (xxi) compete with higher bonded clusters containing the radial 5-fold motif, and so have an optimal yield of lower than 100% even with an 8 particle *alphabet*. Nevertheless, for all the clusters, a 4 particle *alphabet* can give quite large

Figure 3.8: Design principles for low symmetry clusters (xvi-xxi) with smaller alphabets. Each column corresponds to the point group of each cluster, its optimal yields and forces with 2 2, 4 4, and 8 8 sized alphabets in the three rows. The color scheme for the bonded structure refers to the optimal partitioning of 8 particles within 2, 4 or 8 labels, and the visualization of the D_{ij} matrices follow the same color scheme as in Fig. 3.4.

yields in comparison to the maximum possible yield. The variational algorithm identi es the optimal way to partition the groups of interactions of these clusters despite the lack of clear symmetry.

For the C_1 cluster (xix), even a 2 particle alphabet has a high yield, despite not having any exact two-fold symmetry. In this case, the algorithm has recognized a near-symmetry in the cluster and has partitioned it into 2 groups. The symmetry of these letters is close to the symmetrical alphabet that was discovered by the algorithm in a related C_2 cluster (v) in Fig. 3.4. We expect this potential to discover optimal design principles for large clusters with a small number of groups, to be promising towards the self-assembly of experimentally realizable systems with practical constraints on the limits of bottom-up design.

3.4 Design principles for yields of microphase separation

In the previous sections we discussed optimization of the yield of colloidal nanoclusters under a shear ow. The design principles we found however are only valid in systems consisting of particles of only one set of the corresponding alphabet. In cases where there are many copies of the alphabet, it is an entirely new paradigm to search for a set of design principles that produce many copies of identically sized nanoclusters. Formation of such a phase in a self-assembled system, identi ed by the spontaneous emergence of a microscopic cluster size scale even in the presence of a thermodynamic number of particles, is referred to as

Figure 3.9: Optimized design principles for microphase separation into 6-particle clusters. (a) and (b) are optimization curves for shear driving and active driving respectively. Dashed black lines denote the optimization step where the learning rates are lowered in order to anneal into the nal solution well. (c) and (d) are optimized D_{ij} and v_i alphabets for (a) and (b) respectively. The shear in (a) optimizes to $50 k_B T = {}^2$, the maximum allowed value. (e) and (f) denote cluster size distributions for the optimized system with $N = 48$ and a test system with the converged forces but with $N = 96$.

microphase separation or self-limited assembly. Such a phase will compete with a typical gas phase or bulk condensed phase, where the cluster size distribution either has a most probable value at 1 or $O(N)$, with N being the thermodynamic system size. A system with self-limited assembly will show a peak at a microscopic size in the cluster size distribution, with microscopic being de ned as the peak position being $O(1)$, invariant to changing the system size in the thermodynamic limit. Apart from the cluster size distribution, peaks in structure factor corresponding to a recurring microscopic lengthscale can also be used as a suitable order parameter to detect a microphase [213].

Biological systems widely use self-limited assembly to regulate volumes of cells, vesicles, and viral capsids [214, 215], e ciently package genetic material in nuclei [216], and tune mechanoelastic properties of cytoskeletal bundles [217]. The mechanism for stabilizing these structures are current areas of research. Stabilizing a microphase against bulk condensation requires either conservative forces or dissipative driving to compete against surface tension. In equilibrium systems, multiple mechanisms of self-limited assembly via short-range at-

tractive and long-range repulsive interactions, through geometric frustration, and through kinetic control, have been well studied [213, 218, 219]. Inverse design algorithms for designing microphase separation by optimizing a Boltzmann distribution have also been previously studied [220]. But the role of nonequilibrium mechanical or chemical control of microphase separation, such as is ubiquitous in active processes in cells, is not well understood.

In this section we apply the variational algorithm to design DNA-labeled colloidal particles for microphase separation under nonequilibrium driving. Our observable to optimize is a time averaged indicator function, as de ned in Eq. 3.7, for the formation of 6-particle clusters of any geometry and connectivity. As basis set we choose a 6 alphabet, where now the nonzero diagonal elements denote pairwise interaction strength between multiple copies of identical particles. We simulated with overdamped Langevin dynamics a periodic box of 8 copies of the alphabet, with a total of $N = 48$ particles at a packing fraction of 0.02. The particles feel pairwise WCA repulsion given by Eq. 3.5, and a Morse potential attraction with tunable amplitudes given by Eq. 3.6, representing a coarse-grained DNA-mediated e ective force. We optimize the yield of 6-particle clusters under two di erent dissipative driving conditions: a shear ow with a tunable rate given by Eq. 3.4, and alternately by making the colloidal particles active with a self-propulsion force given by

$$f_i^A = v_i \hat{p}_i \tag{3.25}$$

where v_i is an optimizable magnitude of self-propulsion, and $\hat{p}_i$ is a three-dimensional polarization vector that di uses as [221]

$$\hat{p}_i = \sqrt{2D_r}(\hat{p}_i \times \xi_i) \tag{3.26}$$

Here $D_r = 3k_BT/\gamma^2$ is the rotational di usivity and ξ_i are unit variance Gaussian white noise that satisfy

$$\langle \xi_i(t)\rangle = 0 \qquad \langle \xi_i(t)\, \xi_j(t')\rangle = 2I_{3\times 3}\,\delta_{ij}\,\delta(t - t') \tag{3.27}$$

We optimize self-propulsion v_i also within a 6-particle alphabet. We choose the biasing parameters in Eq. 3.13 to be $\lambda = 10^{10}$ as the second term in the cost-function scales with the system size and is larger than in previous sections.

Figure 3.9 shows the results of the optimization for microphase separation with shear driving and active driving. (a) and (b) are learning curves showing the increasing yield of 6-particle clusters with optimization steps. Similar to previous sections, we have had to lower the learning rate in order to anneal smoothly into the solution well. The optimized values of the yield are 5 7%, compared with an initial value of 1 2%. Sub gures (c) and (d) show the optimized alphabets in the respective systems. In the shear-driven system, the shear ow rate optimizes to a value of $5k_BT/\gamma^2$ which is the maximum we allow during optimization. The optimized D_{ij} alphabet is relatively sparse, coding for a few strong attractions. This implies same particles from multiple alphabets bonding together to increase probability of a 6-particle cluster. This is con rmed by a size-scaling analysis of the cluster-size distribution

pro le. For N = 48, the distribution is peaked at 1 and shows a second peak at a cluster size of 6, similar to the expected distribution in a microphase separated system, however, on doubling the system size, the second peak moves to approximately double the desired cluster size. This con rms that the alphabet codes for multiple copies of the same kind of particle assembling into 6-particle cluster, and that this is not true microphase separation. Attempts at eliminating these nite size e ects by optimizing larger systems directly did not result in a second peak in the cluster-size distribution pro le.

The optimization run for active driving converges smoother than the previous case, however, the algorithm nds a large degenerate class of solutions where either all particles experience strong attraction and strong activity, or they are both weak, with many solutions in between. The high attraction and activity solution has been shown with colormaps in Fig. 3.9(d). This solution does not result in a microphase, as has been shown in the cluster-size distribution pro le in sub gure (f). The distribution is peaked at 1 and it is invariant to system size, thus being typical behavior for a gaseous phase. For optimal forces with weaker attractions, the optimized activities required to melt the clusters into a gas are also weak. The variational algorithm nds this same class of gas-phase solutions regardless of choice of starting points in parameter space.

Hence we see that shear driving or active driving are unable to generate microphase separation in coarse-grained models of DNA-labeled colloids. This results from the inability of these mechanically driven dissipative modes to couple to speci c probability uxes that increase the probability of a microphase separated state. In the next section we extend our algorithm to directly optimize reactive uxes and currents for single nanoclusters to isomerize between di erent shapes.

3.5 Design principles for reactive ux in steady-states

Self-assembly is a dynamic process that involves stabilizing a target structure as well as increasing the kinetic ux into the target [172]. Fluxionality of a target structure is important in order to correct structural errors during its assembly. Further, any functional structure like an enzyme or a molecular motor must be not only stable but also able to disassemle and reassemble in a dynamic nonequilibrium steady-state. Functional molecules are able to couple their internal reactive modes to external mechanical or chemical dissipative forces, thus breaking detailed balance in speci c reaction coordinates that are coupled to their output work. Thus, optimizing self-assembly for optimizing the probability ux in and out of a target structure is the path to designing functional materials. Previous work on optimizing the reactivity of self-assembled nanoclusters have used equilibrium approximations to look at single transient realizations of reactive barrier crossing [222, 162], however, the design of long-lived probability uxes and currents in and out of functional structures in nonequilibrium steady-states remain unexplored.

In this section, we use our variational algorithm to optimize reactive uxes and currents owing between 6-particle clusters of C_{2v} and O_h geometries in a sheared steady-state of

Figure 3.10: Optimal design principles for high reactive uxes and currents. (a) shows the three major states the sheared colloids go through to manifest high uxes and currents. (b) and (c) are optimized D_{ij} alphabets for high reactive uxes between C_{2v} and O_h, and for high reactive current from C_{2v} to O_h respectively. d) Integrated reactive displacement corresponding to the high-current solution, at zero and nonzero shear ow rate. (e) and (f) are the steady-state averaged time-dependent reactive uxes in both directions for the high ux and high current solutions respectively, at zero (dashed blue and red lines) and nonzero (blue and red symbols) shear ow rates. The vertical dashed black lines denote the delay used in the optimization.

N = 6 particles. Our ux observable is de ned as

$$Q_t^{C_{2v}!\ O_h} = \frac{1}{\ }\int_0 1_{C_{2v}}(t)1_{O_h}(t+\ t)dt \tag{3.28}$$

for a xed delay time t. Similarly, we de ne the reverse ux $Q_t^{O_h!\ C_{2v}}$, and currents $J_t^{C_{2v}!\ O_h}$

and $J_t^{O_h \to C_{2v}}$ as the differences of the fluxes,

$$J_t^{C_{2v} \to O_h} = Q_t^{C_{2v} \to O_h} - Q_t^{O_h \to C_{2v}} \tag{3.29}$$

The delay-time is chosen to be comparable to the typical transition path time needed for individual reactive events to happen, as observed from pre-optimization simulations. This definition of probability current is associated with a nonzero entropy production and is nonzero on average only in a nonequilibrium steady-state. Directly optimizing such probability currents will design the system to couple the specified reaction coordinate maximally to the ambient mechanical dissipatively driven mode.

We choose the same parameters for overdamped Langevin dynamics simulations and optimization as we had chosen in Section 3.2. We have chosen the time delay to be $\tau = 0.025\tau_0$, comparable to observed transition path timescales from a nonspecific alphabet simulation of $D_{ij} = 5k_BT$ for all pairwise attractions, at equilibrium. We find that both the unidirectional fluxes and the currents are maximized when the shear flow rate is the largest we allow, $f = 50k_BT/\sigma^2$. Designing the system for high fluxes and currents lead to the prevalence of three major states as shown in Fig. 3.10(a). The high flux optimization runs on either direction yield the same optimal alphabet, Fig, 3.10(b), that is an intermediate between the respective maximal alphabet solutions for C_{2v} and O_h. However, the optimal alphabet for high currents from C_{2v} to O_h resemble neither of those two, but codes for the octahedral cluster to sometimes unfold into the nonrigid C_{2h} cluster shown in sub figure (a), and, keeping the correct order of bonds, refold into a C_{2v} cluster. These irreversible hops are what causes the symmetry of the bidirectional reactive fluxes to break and contributes to positive current round the loop. We have shown in Fig. 3.10(d) the total displacement due to this current,

$$R[X(t)] = \int_0^t J_t^{C_{2v} \to O_h}(t)dt \tag{3.30}$$

that increases to nonzero values with time at the driven nonequilibrium steady-state, as opposed to an equilibrium system having the same conservative forces as the optimized alphabet. These nonzero displacements are used in enzymes and molecular motors to denote the amount of irreversible work done and the emergence of these currents denotes that the molecule has been successfully designed to harvest mechanical nonequilibrium driving to produce usable work [223]. In contrast, optimization runs for the reverse current $J_t^{O_h \to C_{2v}}$ were unsuccessful to find any set of forces that led to any positive value. The optimal value of the reverse current was found to be zero arising from either maintenance of the bidirectional reactive flux symmetry or from the nonequilibrium driving being too small to generate any fluxes at all. This selectivity in the ability of only specific reaction coordinates to positively couple to the shear driving is remarkable and could arise from the comparative mechanical deformation modes and radii of gyration of the C_{2v} and O_h clusters, as that is how they couple to a shear flow.

We have explicitly shown the steady-state averaged time-dependent reactive fluxes in the optimized solutions for high flux and high current in Figures 3.10(e) and (f) respectively The

vertical black dashed line denotes the choice of f, and we see that the side-side correlation functions there have gone into the linear regimes but haven't started saturating yet. The correlation functions in both cases have periodic oscillations corresponding to the fast tumbling motion of the cluster in the sheared steady-state. At the long time limit $t \to \infty$, the reactive fluxes in both directions become identical and time-independent due to the loss of correlation. In subfigure (e), we see that the reactive flux has been maximized in both directions while approximately preserving their symmetry. The shear flow has an active role in this high flux, as clear from the same correlation functions with the same optimized alphabet but at $f = 0$. Subfigure (f) similarly shows the reactive fluxes for the optimized high-current forces. Here we see that the algorithm has found a way to maximally break the symmetry between the two reactive fluxes at the given value of t, by delaying the reverse reactive flux compared to the forward. This breakdown of detailed balance in this prespecified reaction coordinate is evidently achieved by the only nonconservative force in the system, the shear flow, as demonstrated by the restoration of the symmetry by putting the shear flow rate to 0 even with the optimized pairwise forces.

We have thus shown that the variational algorithm can be used to tune reactive fluxes and currents in a nonequilibrium steady-state. In a sheared system this amounts to maximizing the coupling between mechanical driving and chemical degrees of freedom for the colloids, thus harvesting the dissipative driving to break detailed balance in select modes. The optimizability of the currents itself depends on the ability of the clusters in the relevant direction to be mechanically affected by a shear flow. This approach opens a promising avenue for application of the variational algorithm to directly design functional materials which may convert dissipation into usable work. In the next section we look at such a molecular machine and directly optimize observables related to its performance.

3.6 Optimal design principles for molecular machines

Molecular machines are functional molecules that consume chemical energy at a nonequilibrium steady-state to perform useful mechanical or chemical work. The operation of molecular machines far from thermal equilibrium enables them to overcome thermodynamic bounds on performance as well as produce a finite power. The operation of molecular machines to produce desired levels of power at high efficiency has been extensively studied with the tools of stochastic thermodynamics and fluctuation theorems[19], with geometric near-equilibrium response theory [201, 224] and with large deviation theory for efficiency fluctuations [225]. The nonequilibrium operation of molecular machines have been analyzed in the framework of free energy and information transduction between multiple nonequilibrium reservoirs and the role of tight internal coupling for efficient operation of molecular machines has been recognized [226]. Nevertheless, most design approaches for molecular dynamics models of machines use phenomenological surveys to find high power and high efficiency solutions, instead of an automated method that can discover novel solutions. In this section we extend our variational algorithm for the discovery of high power and high efficiency regimes

Figure 3.11: Optimal design principles for a model for ATP synthase. a) A schematic representation of the two-particle system, where the F_0 particle (blue) climbs down a potential ramp driven by $_{H+}$, and the coupled F_1 particle is pulled up a ramp doing work against $_{ATP}$. b) Representative statistical realization of the two-dimensional trajectories for a pre-optimization starting point (red), at a set of parameters with maximum e ciency but not high power (black), and a set of parameters with both maximum power and e ciency (orange). (c) and (d) are the optimization landscapes for power and e ciency respectively. Black lines are the trajectories taken by the variational algorithm through the parameter space in either of the two cases. The square, triangle and circle denote the set of parameters corresponding to (b).

that have been previously phenomenologically studied, in a coarde-grained model for the molecular motor ATP synthase.

The F_0 F_1 ATP synthase is vital towards most of the ATP production from aerobic respiration in mitochondria. This motor resides across membranes in mitochondria and transforms cross-membrane proton motive force into energetic phosphate bonds to make ATP. The motor consists of two subunits F_0 and F_1 each having a three-fold symmetry. The

F_0 subunit is rotated by protons moving across the membrane due to proton motive force $_{H+}$, which in turn drives the F_1 subunit to rotate and produce $_{ATP}$, a chemical potential leading to ATP synthesis from ADP and inorganic phosphate. We model the rotation of the F_0 and F_1 subunits as overdamped Langevin dynamics of two coupled particles in one-dimensional periodic cosine potentials [227]. This model is schematically illustrated in Fig. 3.11(a). The coordinates of the particles, $_0$ and $_1$, di use in the potential energy landscape

$$V(\ _0;\ _1) = \frac{1}{2}E_0 \cos n_0(\ _0 \quad) \quad \frac{1}{2}E_1 \cos n_1\ _1 \quad \frac{1}{2}E_{couple} \cos(\ _0 \quad _1) \tag{3.31}$$

where E_0 and E_1 are barrier heights in the periodic potential, is a phase o set between $_0$ and $_1$ and E_{couple} is the coupling energy of the two particles. We allow $_0$ and $_1$ to be in the range [0; 2] with periodic boundaries. Aside from the conservative forces, the F_0 particle is driven by a constant external force $_{H+}$ and the F_1 particle by a constant force $_{ATP}$. Choice of these two driving forces as $_{H+} > 0$, $_{ATP} < 0$ and $j\ _{H+}j > j\ _{ATP}j$ pushes the system out of equilibrium and allows the machine to do work against $_{ATP}$. Similar to Ref. [227], we choose $n_0 = n_1 = 3$ and $E_0 = E_1$. With this model, the input and output powers and the e ciency associated with each trajectory are de ned by

$$P_{H+}[X(t)] = \frac{2\ _{H+}}{\ _0}\int_0^{\ } \dot{\ _0}\,dt \tag{3.32}$$

$$P_{ATP}[X(t)] = \frac{2\ _{ATP}}{\ _0}\int_0^{\ } \dot{\ _1}\,dt \tag{3.33}$$

$$= \frac{P_{ATP}}{P_{H+}} \tag{3.34}$$

with the e ciency being bounded by $_{ATP} = \ _{H+}$. The parameter regimes for optimal performance in terms of these metrics in this model has been studied in detail in Ref. [227].

We use our variational algorithm to nd optimal choice of the parameters used in Eq. 3.31 for independently maximizing power and e ciency. For a cost-function, we retain only the rst term in Eq. 3.13 with the power or e ciency as observables, the choice of thus being rendered to that of the learning rate. Optimization of a trajectory-averaged output power is exactly similar to the approach used for yields and reactive uxes, and for gradients of trajectory averaged e ciency with respect to variational parameters $\{c\}$, we use the chain rule

$$\frac{@}{@c} = \frac{P_{H+}\,@P_{ATP} \quad P\ _{ATP}\,@P_{H+}}{P_{H+}^2} \tag{3.35}$$

We hold constant $E_0 = E_1 = 2k_BT$, $_{H+} = 4k_BT$ and $_{ATP} = 2k_BT$ to avoid trivial optimization of power with increased chemical potentials and decreasing barrier heights and that of e ciency with changing ratio of chemical potentials. Our variational parameters and E_{couple} are allowed to vary in [;] and $[5k_BT; 20k_BT]$ respectively. For simulating the dynamics at each optimization step and estimating gradients, we use a timestep of 10^{-5}, an

integration time t for Malliavin weights as 10, and a trajectory length of 500, all in units of the natural di usive timescale of the system with friction coe cients and di usion constant both being unity.

We initialize the power and e ciency optimization runs from $= 0; E_{couple} = 10k_BT$ and had to lower learning rates in two steps to anneal to the solution well. The results of the optimization runs through the two-dimensional parameter landscape for power and e ciency are shown in Fig. 3.11(c) and (d) respectively, with single statistical realizations of equal durations of unoptimized, high-e ciency and high-power trajectories shown in Fig. 3.11(b). We see that unoptimized trajectories have generally correlated motion between $_0$ and $_1$ albeit with sudden failure modes that lead to loss of power and e ciency. In contrast, high power is uniquely achieved at a high value of E_{couple} around a phase shift of $2:2$, which is similar to results in Ref. [227]. The high power trajectories also have high e ciency, as shown by the highly correlated motion of $_1$ versus $_0$ with a high slope. There is however a large near-degenerate regime of parameters, at high E_{couple}, where near-maximum e ciency close to the theoretical bound is achieved irrespective of the phase shift. This is the regime of tight coupling between input and output coordinates where any energy leakage is minimal. The algorithm nds this degenerate solution basin and di uses out in it, as the noise in the stochastically evaluated gradients becomes comparable in magnitude to the gradient of the e ciency landscape. There is therefore a class of solutions at high coupling strength and positive phase shift that have high e ciency but low output power. One such statistical realization is shown in Fig. 3.11(b), where the two coordinates have tightly e cient correlations but the total amount of displacement for the same amount of elapsed time is less than the high-power solution. Further analysis in terms of the exact mechanism of successful and failed energy transduction segments in these trajectories is required in order to generalize these design principles to machines with more degrees of freedom that can dissipate and reroute energy.

3.7 Conclusion

We have developed an inverse design algorithm for the self-assembly of colloidal clusters in a nonequilibrium steady-state. The formalism exploits a variational structure originating from large deviation techniques for importance sampling in trajectory ensembles. The algorithm optimizes the yield of clusters of arbitrary shapes, sizes and geometry by tuning control forces within an arbitrarily chosen ansatz, with statistically estimated explicit gradients. We have demonstrated the performance of the algorithm using an ansatz of DNA-labeled colloidal clusters self-assembled in a shear ow, and have obtained design rules for di erent families of rigid and nonrigid clusters. This algorithm scales linearly both in system size and in the number of optimizable parameters in the force ansatz, but its performance is independent of the speci c order parameter chosen for the self-assembly process. For example, the choice of a locally de ned structural order parameter such as the density or the degree of crystallinity as the optimized observable can produce design principles for the assembly of extended

dispersed or periodic structures out of equilibrium. Similarly, dynamical order parameters like the instantaneous ux between two stable states and performance metrics of molecular machines like output power and e ciency can also be optimized using the same variational procedure in a suitable trajectory ensemble. Hence this algorithm can be used to tune both structural and dynamical properties of clusters in a nonequilibrium steady-state to produce dynamical phases having no equilibrium analogs.

This variational algorithm di ers from other available inverse design algorithms for soft matter in and out of equilibrium. The equivalent variational structure in con guration space for systems in thermal equilibrium, where the potential energy function is optimizable and explicit gradients can be statistically estimated by autodi erentiation, has been used extensively as the basis of both importance sampling and inverse design algorithms [71, 156, 228]. This con guration space approach with explicit gradients is not feasible in nonequilibrium systems due to the probability measure being non-Boltzmann. Out of equilibrium, there have been theoretical approaches to rationalizing design principles in one or two component systems [22, 229, 21].

In the absence of a closed form expression for the con guration space measure, machine learning algorithms have been previously used to identity optimal design principles by tracking an order parameter during or at the end of nite-duration trajectories [230, 231, 232, 233]. Machine learning or neuroevolution based approaches are equivalent to estimating numerical gradients in the design space using nite-di erence methods, and have similar convergence properties as explicit gradient based methods in the limit of small optimization steps [234]. However, since our multidimensional statistical gradient estimates are obtained using information from the same trajectory, our explicit gradient based method is expected to be advantageous in a high-dimensional design space as typically encountered in the bottom-up design of soft materials.

Finally, a class of design algorithms employ a trajectory ensemble based approach to statistically estimate explicit gradients of the dynamical response in colloidal systems [162, 222] and are formally closest to our approach. These algorithms probe the transient dynamical response of self-assembly trajectories, which however do not predict the long-time properties of the self-assembled cluster in a nonequilibrium steady-state. In our algorithm, we have explicitly evaluated the long-time steady-state limit and arrived at the novel uctuation-dissipation relation in design space in Eq. 3.14. This will enable our explicit-gradient based method to be directly used to optimize both structural and dynamic properties of driven phases of soft matter, and automate the discovery of new functional materials.

Data Availability

The data that support the ndings of this study are openly available in Zenodo at https://doi.org/10.5281/zenodo.4289235 [235].

Chapter 4

Reinforcement Learning of rare di usive dynamics

We have so far developed an optimization technique based on stochastic gradient descent to optimize control forces to sample and design rare but important trajectories. The performance of our variational algorithm relies on getting a low-variance estimate of the explicit gradients of the cost function in parameter space. In the steady-state problems we have studied, we have seen a large improvement in statistical e ciency of sampling compared to direct population dynamics approaches, allowing us to probe novel regimes of rare uctuations. However, the accelerated gradient descent method only incorporates memory through an inertia vector, while discarding prior information about the parts of the cost-function landscape that it visited before. Additionally, generalization of the optimization algorithm to nite duration trajectories ampli es the problem of noisy gradients, as in nite duration the gradients have to explicitly computed by a temporal double integral of a correlation function, instead of a single steady-state integral. This necessitates development of more advanced learning algorithms where auxiliary functions can reduce the variance of gradients and store memory of the cost-function landscape. In this chapter, borrowing notions from reinforcement learning [236], we have developed a method to generate rare dynamical trajectories directly through the optimization of an auxiliary dynamics that generates an ensemble of trajectories with the correct relative statistical weights. Within this ensemble of trajectories, a variational estimate of the likelihood of the rare event is obtainable from a simple expectation value. We have built on the work in Refs. [117, 125] and past literature on reinforcement learning for continuous time processes [237, 238, 239, 240, 241, 242, 243, 244].

The techniques of reinforcement learning aim to learn the best decisions to make in each state in order to achieve some goal. Algorithms developed in this context have led to many signi cant advancements in recent years across tasks requiring an intelligent agent to interact with an environment, such as in gameplay [245, 246, 247] and robotics [248, 249, 250], with a variety of recent applications in physics [251, 252, 253, 254, 255, 256, 257, 258]. However, many of these situations are framed as discrete time problems, with relatively little work

done in stochastic continuous time control [237, 238]. For di usive processes and importance sampling molecular dynamics, we formulate a reinforcement learning procedure to learn the correct force to in uence the probability of choosing each next state. From this perspective, we take a policy gradient based approach [259, 260, 238, 248, 249], learning a generative model for the evolution of the state. The optimized force found is such that rare events are made typical while staying close to the original force, providing a dynamics that can aid in e ciently sampling the targeted trajectory ensemble.

A key advantage of the reinforcement learning techniques we develop is the use of an additional learning process for a function which guides the optimization of the dynamics, a so-called value function [261], which describes how relevant each state is to the rare events of interest. This value function substantially reduces the variance in estimates of the gradient of the parameters specifying a force, allowing for the use of less data in each optimization step and subsequently more complex approximations to the auxiliary dynamics. We show how this approach can be successfully applied to both nite time problems in which the dynamics is constrained to guarantee the occurrence of some rare transition like a barrier crossing, and to time-homogeneous problems where we are interested in the statistics of time-integrated observables in the long time limit as characterized by its large deviation function.[1]

4.1 Trajectory ensemble Formalism

We consider systems evolving with a di usive dynamics over time t of a con guration x. These con gurations evolve according to a force vector $F(x;t)$ and noise vector of equal dimension W with associated constant noise matrix G invertible within the stochastically evolving subspace, represented by the stochastic di erential equation (SDE)

$$dx = F(x;t)dt + G \quad dW; \tag{4.1}$$

where the noise W follows a Wiener process, with increments dW drawn from a Gaussian with zero mean and dt variance. Throughout we will work in dimensionless variables that imply unit energy scales and mobilities. The requirement of G being invertible within the stochastic subspace may in principle be relaxed, however in that case there may be multiple noise vectors corresponding to the same change of state, making the evaluation of transition probabilities necessary for our optimization approach di cult. We will follow the Ito convention for ease of notation and implementation with standard numerical integrators. Throughout, we do not assume in Eq. 4.1 that the force is gradient or that the noise obeys a detailed balance, and thus our approach is generally applicable to equilibrium as well as nonequilibrium dynamics.

We aim to probe rare uctuations in trajectory observables. Here we consider trajectories, $X_{0;T}$, de ned as the sequence of con gurations over an observation time T, though generalizations of uctuating observation times are possible [263] . Generally, we will consider

[1] Most of the content of this chapter was originally part of the publication [262].

observables that are functions of time-integrated variables over the trajectory,

$$O[X_{0;T}] = \int_0^T dt\, A[x_t;t] + B[x_t;t]\ \underline{x}(t); \tag{4.2}$$

where the rst term is a state dependent observable, while the second term depends on a stochastic increment, with both $A[x_t;t]$ and $B[x_t;t]$ being state dependent. However, we will also consider cases in which $A[x_t;t]$ is a function of a single time in order to impose end point conditioning. Expectations of functions of such observables are de ned through path integrals of the form

$$\langle f\ (O[X_{t;t^0}])\rangle_p = \int DX_{t;t^0} dx_t\, P[X_{t;t^0}]\, f\ (O[X_{t;t^0}]); \tag{4.3}$$

where $P[X_{t;t^0}]$ is the total probability of a trajectory decomposable into $P[X_{t;t^0}] = p[X_{t;t^0}|x_t]$ (x_t) where $p[X_{t;t^0}|x_t]$ is the transition probability conditioned on starting in con guration x_t with initial probability (x_t).

Probabilities for trajectories between times t and t^0 starting at x_t are de ned by

$$p[X_{t;t^0}|x_t]\ \ /\ \ \exp\left\{ \tfrac{1}{2}\int_t^{t^0} dt^{00}\, \|G^{-1}\ (\underline{x}\ \ F)\|^2 \right\} \tag{4.4}$$

where we suppressed the arguments of x_t and $F[x_t;t]$ for shorthand. This is the standard Onsager-Machlup form for the di usive dynamics considered here [264]. The measure over paths between times t and t^0 starting from position x_t is de ned such that

$$\int DX_{t;t^0} p[X_{t;t^0}|x_t] = 1 \tag{4.5}$$

where the transition probability is normalized when integrated over all trajectories. These path probabilities satisfy

$$p[X_{t;t^{00}}|x_t] = p[X_{t^0;t^{00}}|x_{t^0}]\, p[X_{t;t^0}|x_t] \tag{4.6}$$

and

$$DX_{t;t^{00}} = DX_{t^0;t^{00}} DX_{t;t^0} \tag{4.7}$$

due to the Markovian noise in Eq. 4.1.

Trajectories sampled with $P[X_{0;T}]$ will be dominated by the most typical values of $O[X_{0;T}]$. We will encode the rare trajectories with atypical values of $O[X_{0;T}]$ by reweighting the original trajectory ensemble de ned by Eq. 4.4, multiplying each trajectory by an observable dependent factor. Such reweightings occur naturally in statistical studies of rare events and are isomorphic to extended ensemble approaches in equilibrium con gurational problems. The ensemble of events we are interested in is constructed by weighting the probability of trajectories in the original dynamics by an exponentially positive number,

$$P_s[X_{0;T}] = e^{\ sO[X_{0;T}]\ \ (s;T)} P[X_{0;T}]; \tag{4.8}$$

where $P_s[X_{0;T}]$ is denoted as a tilted path ensemble, biased by a statistical field s in such a way to promote rare fluctuations in $O[X_{0;T}]$. The quantity $\Psi(s;T)$ normalizes the tilted distribution, and is identifiable as a cumulant generating function (CGF)

$$\Psi(s;T) = \ln Z(s;T) = \ln \left\langle e^{sO[X_{0;T}]} \right\rangle_p ; \tag{4.9}$$

and equal to the logarithm of the tilted path partition function $Z(s;T)$. The reweighted path ensemble generally defines a new transition probability $p_s[X_{t;t'}|x_t]$ and initial condition. The evaluation of $\Psi(s;T)$ is a common objective in studies of diffusive systems as it describes the statistics of $O[X_{0;T}]$. Contributions to $\Psi(s;T)$ or $P_s[X_{0;T}]$ are dominated by trajectories with large or small values of $O[X_{0;T}]$, depending on the sign of s. The exponential bias, $\exp(-sO[X_{0;T}])$, can also be constructed to function as a filter based on fulfilling specific criteria. In such cases $P_s[X_{0;T}]$ is identified as the probability that a trajectory fulfills a specific conditioning, and its ensemble a corresponding conditioned path ensemble. Common examples are Brownian bridges [109, 265, 266], where trajectories are conditioned to end at $x_T = x^0$, in which $O[X_{0;T}]$ is 1 if $x_T = x^0$ and is 0 otherwise, and s is taken sufficiently negative that only those trajectories for which the constraint is satisfied have significant weight.

4.2 Gradient optimization for finite time constrained dynamics

Our aim is to find a dynamics which generates trajectories with probability as close to the reweighted trajectories ensemble as possible. For the diffusive dynamics considered here, this is exactly achievable in principle through a so-called generalized Doob transformation [267, 142, 51, 63, 110, 268]. The generalized Doob transformation defines a modified dynamics with an added drift force that is generally time dependent but with an identical noise as in the original SDE. However, constructing this transformation is often not possible in practice, as it requires diagonalizing a modified Fokker-Planck operator which in interacting systems is exponentially complex [269]. Here we aim to parametrize a drift force with tunable parameters to approximate the generalized Doob transform. With the modified force $F(x;t)$, we have a modified SDE

$$dx = F(x;t)dt + G\,dW ; \tag{4.10}$$

with corresponding trajectory probabilities

$$p[X_{t;t'}|x_t] \propto \exp\left\{ \frac{1}{2}\int_t^{t'} dt'' \left| G^{-1}(\dot{x} - F) \right|^2 \right\} \tag{4.11}$$

which still satisfy the Markovian properties of the original dynamics and the same normalization constant. See Ref. [125] for a discussion of problems in which the optimal dynamics is required to be non-Markovian, in the context of discrete time Markov processes.

We seek to learn a set of parameters to minimize the Kullback-Leibler (KL) divergence between the modified dynamics and the reweighted trajectory ensemble defined by Eq. 4.8. The KL divergence is defined as

$$D_{KL}(p \,\|\, p_s) = \ln \frac{p_{\theta}[X_{0;T}|x_0]\,(x_0)}{p_s[X_{0;T}|x_0]\,(x_0)} \Bigg\rangle_p ; \tag{4.12}$$

where the expectation is taken with respect to the parametrized dynamics. This quantity is a measure of the similarity between the modified and reweighted trajectory ensembles. Achieving a zero value when p_{θ} is given by the generalized Doob transform, the KL divergence has a unique minimum when this Doob transformed dynamics is contained within the space of parametrized dynamics, providing a variational estimate of the CGF. We note that this definition of the KL divergence differs from much of the literature considering optimization of a parametrized diffusive dynamics [270, 271, 272, 273], where the parametrized dynamics p_{θ} and target dynamics p_s appear in an opposite way. In principle the initial distribution should also be parametrized, as it will be modified by the reweighting, however depending on the space of distributions chosen these can be hard to sample. We drop this modification for simplicity.

Low variance gradient estimation

In order to optimize the force, F_{θ}, we follow techniques introduced in the reinforcement learning literature [236, 274, 275, 276, 248, 277]. Substituting the parametrized and reweighted trajectory probabilities into the KL divergence, we may rewrite it as an average over a parameter dependent time-integrated observable

$$D_{KL}(p \,\|\, p_s) = \langle R[X_{0;T}] \rangle_p + \Phi(s;T) \tag{4.13}$$

where in the language of reinforcement learning we define a return R, $R[X_{0;T}]$, as

$$R[X_{0;T}] = s\mathcal{O}[X_{0;T}] - \ln \frac{p_{\theta}[X_{0;T}|x_0]}{p[X_{0;T}|x_0]} \tag{4.14}$$

with the negative of the average of the second term measuring the KL divergence, $D_{KL}(p_{\theta} \,\|\, p)$, between the parametrized dynamics and the original dynamics. This return is analogous to a regularized form of reinforcement learning [275, 277] similar to that considered in maximum-entropy reinforcement learning [248, 249, 276]. When evaluated at the generalized Doob transform the KL divergence vanishes and the return evaluates to the CGF. Away from the Doob transform, the positivity of the KL divergence results in the return variationally bounding the CGF from below [111].

We aim to minimize the KL divergence through stochastic gradient descent in the parameter space. For this we need the gradient of $D_{KL}(p_{\theta} \,\|\, p_s)$ with respect to θ,

$$\nabla_{\theta} D_{KL}(p_{\theta} \,\|\, p_s) = \langle R[X_{0;T}] \nabla_{\theta} \ln p_{\theta}[X_{0;T}|x_0] \rangle_p \tag{4.15}$$

where we note

$$\langle \nabla \; R[X_{0;T}] \rangle_p = 0 \tag{4.16}$$

due to conservation of probability [125]. The factor multiplying the return is commonly referred to as the Malliavin weight in the stochastic analysis literature [122], and corresponds to a particular case of the eligibility traces found in reinforcement learning [261, 236, 278, 279, 280], which we denote as $y_s(T) = \nabla \; \ln p \; [X_{0;T} | x_0]$. It can be rewritten by substituting the path probability,

$$y(t'') \quad y(t') = \int_{t'}^{t''} dt \, \underline{y}(t); \tag{4.17}$$

where

$$\underline{y}(t) = \quad G^{-1} \; (\underline{x}(t) \quad F(t)) \qquad G^{-1} \; \nabla \; F(t) \tag{4.18}$$

is the integrand of the Malliavin weight.

Were we to stop at Eq. 4.15, we would proceed to optimize a generative model (the diffusive dynamics with our parameterized force) of the trajectories using a score-function based approach, similar to standard unsupervised learning. However, following the methods of reinforcement learning, we can use a combination of the Markovianity of the generative model and other variance reduction techniques to produce a gradient estimator which is much more efficient to estimate. To begin with, we can simplify Eq. 4.15 by noting that due to Markovianity, the Malliavin weight only correlates with the return in the future, and we can rewrite the gradient as

$$\begin{aligned} \nabla \; D_{KL}(p \,|\, p_s) &= \int_0^T dt \, \langle R[X_{t \;;T}] \, \underline{y}(t) \rangle_p \\ &= \quad {}_{MCR}(\;;T); \end{aligned} \tag{4.19}$$

where we used t as a shorthand for t for some small positive . We refer to the optimization of the modified dynamics using this formulation of the gradient as ${}_{MCR}$, as it is analogous to the Monte-Carlo returns (MCR), or REINFORCE [281, 282] policy gradient algorithm in reinforcement learning. In the long observation time limit, employing this gradient in stochastic optimization reduces to previous variational Monte Carlo procedures [117].

This estimator of the gradient is non-optimal for two reasons. First, it requires evaluation of a two time correlation function. In steady state, stationarity can be invoked to eliminate one of those integrals, however under finite time conditioning this simplification is not possible. Second, it has a high variance and requires significant averaging to converge accurate gradients. This is because both the Malliavin weight and the return undergo a random walk with linearly increasing variance [122]. Building on the analogies with the reinforcement learning formalism we define a value function as a path average of the return,

$$V(x;t) = \langle R[X_{t;T}] \rangle_{p \;;x} : \tag{4.20}$$

conditioned on starting at the position and time, $x_t = x$. Introduced into the gradients of $D_{KL}(p \| p_s)$ in distinct ways, the value functions can be used to tame both problems of the naive MCR gradient estimate.

First, we introduce a value function as a baseline that depends only on the state at the time t in order to reduce the variance of the gradient. We note that $y(t)$ is linear in the noise and thus averages to zero when multiplied by a function of the state at or before t. Defining a temporal difference error

$$[X_{t:T}; t] = R[X_{t:T}] - V(x_t; t); \tag{4.21}$$

we write the dynamical gradient as

$$\begin{aligned} \nabla_\theta D_{KL}(p \| p_s) &= \left\langle \int_0^T dt\, [X_{t:T}; t]\, \underline{y}(t) \right\rangle_p \\ &= \mathcal{G}_{MCVB}(\theta; T) \end{aligned} \tag{4.22}$$

where we have formally subtracted zero. We refer to this gradient estimator as $\mathcal{G}_{MCVB}$, for Monte Carlo Value Baseline (MCVB) [236]. The subtraction of the state point dependent value function reduces the variance of the gradient by accounting for the mean uncorrelated part of each return between t and T with $\underline{y}(t)$, focusing on how this return differs from the average behaviour encoded by the value function.

Second, we introduce a value function that encodes an estimate of the return in the future in order to further reduce the variance and also the complications associated with estimating the two-time correlation function. We can replace part of the return by a value function that is conditioned at some τ, such that $t < \tau < T$,

$$\begin{aligned} \langle R[X_{t:T}]\, \underline{y}(t) \rangle &= \langle V(x_{t+\tau}; t+\tau)\, \underline{y}(t) \rangle \\ &\quad + \langle R[X_{t:t+\tau}]\, \underline{y}(t) \rangle \end{aligned} \tag{4.23}$$

where we set the value function to zero for $V(x; t)$ with $t > T$. Combining this value function form of the kernel of the gradient with the value baseline, we define another temporal difference error

$$\begin{aligned} &\delta'[X_{t:t+\tau}; t] \\ &= V(x_{t+\tau}; t+\tau) + R[X_{t:t+\tau}] - V(x_t; t); \end{aligned} \tag{4.24}$$

and we arrive at a distinct formulation of the gradient

$$\begin{aligned} \nabla_\theta D_{KL}(p \| p_s) &= \left\langle \int_0^T dt\, \delta'[X_{t:t+\tau}; t]\, \underline{y}(t) \right\rangle_p \\ &= \mathcal{G}_{AC}(\theta; T) \end{aligned} \tag{4.25}$$

which we denote $\mathcal{G}_{AC}(\theta; T)$ for actor-critic gradient (AC) estimator, for the analogous algorithm in reinforcement learning [236, 248]. Here the value function is seen as criticizing the

transitions generated by the dynamics, i.e. the actor. Variance reduction of gradient estimates is therefore achieved by replacing potentially noisy return samples with the average behaviour expected in the future of the $x_{t+\ }$ state. In Sec. 4.4, we will compare the accuracy and statistical e ciency of these three gradient estimators: MCR, MCVB, and AC. Before that we discuss how the value functions are simultaneously parametrized and learnt along side the modi ed force.

Parametrizing value functions

While the gradient expressions are exact and the use of value functions expected to facilitate their convergence, using them requires knowledge of the exact value function for the modi ed dynamics, a formidable task in complex problems. In order to make their use tractable, we optimize a representation of the value function in addition to the modi ed force. Speci - cally, we introduce a parametrization of the value function denoted V . To optimize this approximation we note that the value functions satisfy a self-consistency equation called the Bellman equation [283]

$$V(x;t) = \langle V(x_{t+\ }; t+\ ) + R[X_{t;t+\ }]\rangle_{p\ ;x}; \tag{4.26}$$

which has a unique solution for a given dynamics and return (as de ned by the tilting observable and the dynamics via Eq. 4.14). We aim to minimize the error in this equation, thus optimizing our parametrized value towards this unique solution. Our approach is to minimize the squared di erence between the two sides of Eq. 4.26 with the true value function replaced by the parametrized value function, and apply gradient descent to it. Such an approach is the subject of gradient temporal di erence methods [284, 285, 286], but produces a gradient estimate which is di cult to evaluate, containing products of expectations which require independent samples. A part of the resultant gradient is however simpler to compute. We derive it by substituting only the right hand side of Eq. 4.26 with our parametrized value function to provide a xed target for the left and de ning a corresponding error function based on the squared di erence. To construct a loss, we integrate these errors along each trajectory, and average them over the trajectory ensemble. This results in a loss function $L(\ ;\ _i)$, that we take as a function of two weights, and $_i$,

$$L(\ ;\ _i) = \frac{1}{2}\left\langle \int_0^T dt\ \left\langle V_i(x_{t+\ }; t+\ ) + R[X_{t;t+\ }]\right\rangle_{p\ ;x} + V\ (x_t;t)^2 \right\rangle_p; \tag{4.27}$$

where the weight $_i$ is the weights after update i, used to provide the xed target estimate towards which we want to move the functional of . The derivative is then taken with respect

to , before setting $= {}_i$ to nd the gradient of this loss for the current parameters. Such an approach is referred to as semi-gradient in the reinforcement learning literature,[236] used to achieve the majority of state-of-the-art reinforcement learning results, and proves stable provided the data used to estimate the gradient is sampled using a dynamics which is close to p as we intend to do. As mentioned above, alternative methods which additionally consider the variation of the target with can be found in the RL literature, allowing for the use of data sampled from an alternative dynamics, utilized via importance sampling [284, 285, 286].

Writing an approximate temporal di erence for the value function parametrization, within MCVB

$$[X_{t\ ;T}; t] = R[X_{t\ ;T}] \quad V\ (x_t; t); \tag{4.28}$$

or for AC

$$\begin{aligned} &{}^0[X_{t\ ;t+\ }; t] \\ &= V\ (x_{t+\ }; t+\ ) + R[X_{t\ ;t+\ }] \quad V\ (x_t; t); \end{aligned} \tag{4.29}$$

we have gradients of the form

$$\begin{aligned} &r\ \ L_*(\ ;\ {}_i)j_{=\ {}_i} \\ &= \int_0^T dt\ \ {}_i[X_{t\ ;T}; t]\, r\ \ V\ (x_t; t)j_{=\ {}_i} + {}_p; \end{aligned} \tag{4.30}$$

for the loss function from the value function parametrization, where for the AC algorithm ${}_i$ is replaced with ${}^0_i$. Given this value function approximation, we can approximate the gradient of the KL divergence by replacing the exact temporal di erence with these approximate temporal di erences. We then use the same trajectories to estimate the force and value function gradients and simultaneously learn both. For the MCVB algorithm, an approximate value function does not bias the gradients as the future return that correlates with the Malliavin weight stays intact and the expectation of the Malliavin weight is identically 0. However, for the AC algorithm, an approximate value function can introduce a bias into gradients as it replaces the average of the future return, which it may not accurately represent.

Employing gradients with or without value functions, we can construct a stochastic descent algorithm to optimize the modi ed forces which can be used to estimate the likelihoods of rare events and the trajectories by which they emerge. The algorithms require the evaluation of the forces, value function, their parametric gradients and noises over the course of simulating trajectories. Ensembles of trajectories can then be used to construct an empirical estimate of the gradient via computing the Malliavin weights, returns, and the temporal di erence. These empirical estimates then iterate the two weights with respective learning rates and for the force and value function respectively. The resultant algorithm is

outlined in pseudocode below in Alg. 2. Detailed versions of the individual algorithms with computationally e cient on-the- y implementations for simulating trajectories with discrete timesteps are presented in the next section.

Algorithm 2 Gradient optimization using nite time trajectories

```
inputs  dynamical approximation F (x; t), value approximation V (x; t)
parameters  learning rates  ,  ; total optimization steps I ; trajectory length T consisting of J timesteps of duration t each; number of trajectories N
initialize  choose initial weights  and  , de ne iteration variables i and j , force and value function gradients P, V, temporal di erence  (can be R[X t ;T] or  [X t ;T; t] or 0 [X t ;t+ ; t] for MCR/MCVB/AC)
i  0
repeat
    Using chosen method to generate trajectories X 0;T with con gurations, times and temporal di erences denoted by x j; t j and j respectively.
    j  0
    P  0
    V  0
    repeat
        P  P + j y(t j) t
        V  V + j r V (x j; t j) t
        j  j + 1
    until j = J
    average P, V over N trajectories to get P, V
        +  P
        +  V
    i  i + 1
until i = I
```

4.3 Discrete timestep implementations of nite time algorithms

We now describe how the time-continuous equations of the reinforcement learning algorithm are e ciently implemented in simulations with a xed discrete timestep t, though variable timesteps may be easily used. We use an Euler propagator to integrate the SDE in Equation (4.10) as

$$x_{t+\ t} = x_t + \ tF(x_t; t) + G \ W_t \tag{4.31}$$

where W is a Gaussian random variable with mean 0 and variance t. The trajectory probability from Eq. (4.11) is now given by products of stepwise probabilities

$$p\,[X_{t;t+n\,t}|x_t] = \frac{\exp\left[-\frac{1}{2\,t}|G^{-1}(x_{t+\,t} - x_t - tF(x_t;t))|^2\right]}{2\,t\,\det(G)} \tag{4.32}$$

Next we discretize the gradient of the logarithm of trajectory probabilities using the Ito convention. We propagate the Malliavin weights from Eq. (4.18) as

$$y(t+\,t) = y(t) + G^{-1}(x_{t+\,t} - x_t - tF(x_t;t)) - G^{-1}r\,F(t) \tag{4.33}$$

We also write the full return (4.14) through a sum of stepwise rewards as

$$R[x_{t\,;t+\,}] = \sum_{j:j\,t<} r(x_{j+1};x_j;t+j\,t) \tag{4.34}$$

where the timestep index j starts from -1 in this sum, with the notation t accounting for the timestep before the current one, and the subscript j refers to the time $t+j\,t$. The reward at each step is de ned as

$$\begin{aligned} r(x_{j+1};x_j;t+j\,t) &= s(A_j\,t + B_j\,(x_{j+1} - x_j) + A(x_{j+1})\,{}_{jn}) \\ &+ \frac{[G^{-1}(x_{j+1} - x_j - tF(x_j;t_j))]^2}{2} \\ &- \frac{[G^{-1}(x_{j+1} - x_j - tF(x_j;t_j))]^2}{2}; \end{aligned} \tag{4.35}$$

using the de nition of the observable from Eq. (4.2) and accounting for an additional singular reward at the end of the trajectory after the last timestep n. Here the rst three terms come from the observable and the last two terms represent the KL divergence between the original and optimized dynamics.

Now we combine the rewards, Malliavin weights and value functions in multiple ways to produce the gradients in the di erent algorithms. The pseudocodes of e cient implementations of these are presented below.

Monte-Carlo returns

The gradient in the Monte Carlo returns algorithm can be rewritten from Equation (4.19) as

$$
\begin{aligned}
{}_{MCR}(\ ;T) &= \int_0^T dt\, R[X_{t\ ;T}]\, \underline{y}(t) \\
&= \int_0^T dt\, \underline{y}(t) \int_t^T dt'\, \underline{R}(t') \\
&= \int_0^T dt\, \underline{R}(t) \int_{t^+}^T dt'\, \underline{y}(t') \\
&= \int_0^T dt\, \underline{R}(t)\, \underline{y}(t^+)
\end{aligned}
\tag{4.36}
$$

where the return has been written as a time integral of its differential changes, and t^+ is shorthand for $t+$ for some small positive. This has converted the double time integral into a single time integral, which is then evaluated on-the-fly while propagating the trajectory. An implementation of this algorithm with a fixed timestep Δt is described in the pseudocode in Alg. 3.

Algorithm 3 Finite time MCR

inputs dynamical approximation $F(x;t)$
parameters learning rate ; total optimization steps I; trajectory length T consisting of J timesteps of duration t each; number of trajectories N
initialize choose initial weights , define iteration variables i and j, force gradient $_P$, stepwise rewards r representing the increments in return
$i \quad 0$
repeat
  Using chosen method to generate trajectories $X_{0;T}$ with configurations, times, noises, Malliavin weights and rewards denoted by $x_j; t_j; \quad W_j; y(t_j)$ and $r(x_{j+1}; x_j; t_j) = r_j$ respectively
  $j \quad 0$
  $_P \quad 0$
  $y(t_0) \quad 0$
  repeat
    $y(t_{j+1}) \quad y(t_j) + \quad W_j \quad [G^{-1} r \quad F(x_j; t_j)]$
    $_P \quad _P + r_j y(t_{j+1})$
    $j \quad j + 1$
  until $j = J$
  average $_P$ over N trajectories to get $\bar{}_P$
    $+ \quad \bar{}_P$
  $i \quad i + 1$
until $i = I$

Monte-Carlo returns with a value baseline

We use a similar technique to rewrite the double time integral for the gradient in the Monte Carlo value baseline algorithm, Equation (4.22), using a single time integral as

$$
\begin{aligned}
&_{\text{MCVB}}(\ ; T) \\
&= \int_0^T dt \ \{ R[X_{t\ ;T}] \quad V(x_t; t) \} \underline{y}(t)\Big|_{p\ ;\ =\ i} \\
&= \int_0^T dt \ \left\{ R(t) y(t^+) \quad V(x_t; t) \underline{y}(t) \right\}\Big|_{p\ ;\ =\ i} :
\end{aligned}
\tag{4.37}
$$

We rewrite the gradient of the value error in Eq. (4.30) similarly as

$$
\begin{aligned}
&\nabla L(\theta; \theta_i) \\
&= \left[\int_0^T dt \; R(t) \int_0^{t^+} dt' \, \nabla V(t') + V(t) \nabla V(t) \right]_{p;\theta=\theta_i} \\
&= \left[\int_0^T dt \; R(t) z(t^+) - V(t) \underline{z}(t) \right]_{p;\theta=\theta_i} ; \qquad (4.38)
\end{aligned}
$$

where the arguments of the value function $V(x_t; t)$ have been suppressed as $V(t)$ and the integral of the gradient of the value function upto and including current time has been denoted as $z(t^+)$. We explicitly set the $V(x_t; t)$ to 0 for any $t \geq T$, i.e., after the last timestep, in these expressions. The single time integral is then evaluated on-the-fly as the trajectory is propagated. If the force and the value function approximations use the same set of basis functions as we do with a fixed grid of Gaussians, the MCVB algorithm incurs no additional computational cost over the MCR algorithm. An implementation of this algorithm with a fixed timestep Δt is described in the pseudocode in Alg. 4.

Actor-critic

We rewrite the gradient in the Actor-critic algorithm from Equation (4.25) using a shift in time origin as

$$
\begin{aligned}
&\nabla_{AC}(\theta; T) \\
&= \int_0^T dt \; \Delta^0[X_{t+\Delta t}; t] \, \underline{y}(t) \Big|_{p;\theta=\theta_i} \\
&= \int_{T+\Delta t}^0 dt \; \Delta^0[X_{t-\Delta t}; t-\Delta t] \, \underline{y}(t-\Delta t) \Big|_{p;\theta=\theta_i} \qquad (4.39)
\end{aligned}
$$

where the change in return and the value function for $t \geq T$ is explicitly set to 0. We similarly write the gradient of the value error from Eq. (4.30) as

$$
\begin{aligned}
\nabla L(\theta; \theta_i) \Big|_{\theta=\theta_i} &= \int_{T+\Delta t}^0 dt \; \Delta^0[X_{t-\Delta t}; t-\Delta t] \\
&\quad + \nabla V(x_{t-\Delta t}; t-\Delta t) \Big|_{p;\theta=\theta_i} \qquad (4.40)
\end{aligned}
$$

Algorithm 4 Finite time MCVB

inputs dynamical approximation $F(x;t)$, value approximation $V(x;t)$
parameters learning rates , ; total optimization steps I; trajectory length T consisting of J timesteps of duration t each; number of trajectories N
initialize choose initial weights and , define iteration variables i and j, force and value function gradients $_P$, $_V$, stepwise rewards r representing the increments in return
i 0
repeat
  Using chosen method to generate trajectories $X_{0;T}$ with configurations, times, noises, Malliavin weights, integral of value function gradients, and rewards denoted by $x_j; t_j;$ $W_j; y(t_j); z(t_j)$ and $r(x_{j+1}; x_j; t_j) = r_j$ respectively
  j 0
  $_P$ 0
  $_V$ 0
  $y(t_0)$ 0
  $z(t_0$ 0)
  repeat
    $\underline{y}(t_j)$ W_j $[G^{1}r$ $F(x_j; t_j)]=$ t
    $y(t_{j+1})$ $y(t_j) +$ $t\underline{y}(t_j)$
    $\underline{z}(t_j)$ r $V(x_j; t_j)$
    $z(t_{j+1})$ $z(t_j) +$ $t\underline{z}(t_j)$
    $_P$ $_P + r_j y(t_{j+1})$ $V(x_j; t_j)\underline{y}(t_j))$
    $_V$ $_V + r_j z(t_{j+1})$ $V(x_j; t_j)\underline{z}(t_j)$
    j $j + 1$
  until $j = J$
  average $_P$, $_V$ over N trajectories to get $\bar{}_P$, $\bar{}_V$
    $+$ $\bar{}_P$
    $+$ $\bar{}_V$
  i $i + 1$
until $i = I$

These integrals are then evaluated on-the- y along with trajectory propagation. Since the gradients involve correlations of the di erential return r with the di erential Malliavin weight $\underline{y}$ and the value function gradient $z = r \quad V$ from time in the past, this makes it necessary to store and use this history, along with the reward and the value function, for the past $= t$ timesteps. Aside from this additional memory requirement, given a delay time which is much smaller than the trajectory duration, the Actor-critic algorithm has similar computational cost comparable to the MCR and MCVB algorithms. This implementation of the algorithm is described in the pseudocode in Alg. 5.

Algorithm 5 Finite time AC

inputs dynamical approximation $F(x;t)$, value approximation $V(x;t)$
parameters learning rates , ; total optimization steps I; trajectory length T consisting of J timesteps of duration t each; temporal delay $M = = t$; number of trajectories N
initialize choose initial weights and , define iteration variables i and j, force and value function gradients $_P$, $_V$, stepwise rewards r representing the increments in return
$i \quad 0$
repeat
  Using chosen method to generate trajectories $X_{0;T}$ with configurations, times, noises, changes in Malliavin weights, value function gradients, temporal difference, rewards and cumulative rewards denoted by $x_j; t_j;$ $W_j;$ $y(t_j); \underline{z}(t_j);$ ^0_j, $r(x_{j+1}; x_j; t_j) = r_j$ and $R\ X_{t_j \quad ;t_j} = R_{j \quad M;j}$ respectively, and $r_j = V(x;t_j) = 0$ whenever $j < 0$ or $j \quad J$
  $j \quad 0$
  $_P \quad 0$
  $_V \quad 0$
  $R_{M;0} \quad 0$
  repeat
    $R_{j \quad M;j} \quad R_{j \quad M \quad 1;j \quad 1} + r_j \quad r_{j \quad M}$
    if $j < J$ then
      $y(t_j) \quad W_j \quad [G^{\ 1}r \quad F(x_j;t_j)]$
      $\underline{z}(t_j) \quad r \quad V(x_j;t_j)$
    end if
    if $j \quad M$ then
      $^0_j \quad V(x_j;t_j) + R_{j \quad M;j} \quad V(x_{j \quad M}; t_{j \quad M})$
      $_P \quad _P + {}^0_j \, y(t_{j \quad M})$
      $_V \quad _V + {}^0_j \, \underline{z}(t_{j \quad M})$
    end if
    $j \quad j + 1$
  until $j = J + M$
  average $_P$, $_V$ over N trajectories to get $\bar{}_P$, $\bar{}_V$
    $+ \quad \bar{}_P$
    $+ \quad \bar{}_V$
  $i \quad i + 1$
until $i = I$

4.4 Rare fluctuations in finite time

We have used the algorithms discussed above to examine rare fluctuations of trajectories of fixed duration, starting from a fixed point in configuration space. The specific observable we

have investigated is an indicator function for reaching a desired region, , in con guration space, $O[X_{0;T}] = h\ [x_T]$, where

$$h\ [x_T] = \begin{matrix} 1 & x_T\ 2 \\ 0 & \text{otherwise} \end{matrix} ;$$

at the nal time T. Rare trajectories reaching a target basin in con guration space are often of interest as transition paths for reactive events, and signi cant development has been undertaken to e ciently generate them [70, 287, 288, 289, 105]. Computing optimal drift forces for generating these rare trajectories enables the study of reactive dynamics in a direct manner. We expect these algorithms to nd use in the study of di usive dynamics where Monte Carlo approaches have di culty sampling [82, 290, 291, 292]. Further, as the modi ed force is used with the original noise from the SDE, we have access to the full reactive trajectory ensemble allowing the interrogation of the statistics of the reactive events in a way that other direct path methods like nudged elastic band and zero temperature string methods do not, as they represent only the dominant path [293, 294, 295, 296]. As a consequence, we expect out method will nd use when there is a large path space entropy.

The CGF for an indicator variable is given by

$$(s;T) = \ln\ \ e^{\ sh\ [x_T]}\ _p \tag{4.41}$$

as an average in the original reference dynamics. From Eq. (4.13), the KL divergence being nonnegative implies the average return is bounded above by the value of the CGF $(s;T)$. The bound can be saturated only by the unique optimal drift force. We compare the value of the optimized return to numerically exact estimates of the CGF given as

$$(s;T) = \ln\ \ 1 + (e^{\ s}\ \ 1)\ \ \int dx\ \ (x;T)\ \ ; \tag{4.42}$$

where the de nition of the indicator function and the nal time distribution $(x;T)$ evolved from a speci c initial condition has been used. This form demonstrates the statistics of a single-time indicator observable is described solely by its mean,

$$hh\ i_p = \int dx\ \ (x;T) : \tag{4.43}$$

For a rare uctuation such that $hh\ i_p < 0{:}5$, this form indicates that there are two distinct regimes in the biased ensemble with $s < 0$. For a small magnitude of the bias, the indicator function stays close to the unbiased value. Below a critical value of $s = \ \ln[hh\ i_p=(1\ \ hh\ i_p)]$ the indicator crosses over to being close to 1. For all of our calculations, we choose a xed value of s estimated to be smaller then the threshold. With this value of s, we compute the right side of Eq. (4.42) using an eigen-expansion of the propagator of the Fokker-Planck equation of the original dynamics, and compare with the value of the average return from the gradient descent algorithms having the same value of s. Details of this calculation and comparison to an approximate Kramers escape rate are in Section 4.5.

Figure 4.1: Softened Brownian bridges: (Left column) Smoothened learning curves showing running estimates of the CGF (a), average value of the indicator observable with the optimized dynamics (b), and the average cost function (c), as functions of optimization steps i, with the MCR(`A', yellow), MCVB(`B', green) and AC(`C',blue) algorithms. The horizontal dashed grey lines denote the numerically exact values. (Middle column) 100 trajectories obtained with the nal converged dynamics from the three di erent algorithms but with the same noise history.(Right column) (g) and (h) show the smoothened convergence of a time-slice of the force parameters, as a function of optimization steps i, in the absence (MCR) and presence (MCVB) of a value function. (i) shows the convergence of the KL divergence cost with ner basis sets optimized with the MCVB algorithm. Green ($3 \times 1 \quad 21$ t), black ($3 \times 1 \quad 41$ t), orange ($3 \times 1 \quad 81$ t) and brown ($4 \times 1 \quad 201$ t) curves show that in the increasing basis limit, the cost-function estimate approaches the value expected from the numerically exact CGF.

Softened Brownian bridges

The rst example we consider is a softened version of a so-called Brownian bridge [297, 109], in which a one-dimensional Brownian motion starting from the origin is biased to end near a particular point. The reference dynamics is simply given by free di usion,

$$dx = \sqrt{2}\,dW \tag{4.44}$$

where comparing to Eq. 4.1 we have $\epsilon = \sqrt{2}$. We consider the target well, (x), to be de ned as $f\,1 \quad x \quad 1 + \quad g$ with $= 0.1$. The dynamics is simulated with a discrete

timestep of 0.001. We use a tilting parameter $s = -100$ to bias the original ensemble towards higher occurrence of the rare event.

We optimize a force and value function parametrized by linear combinations of Gaussian distributions with fixed variance and mean. Given a set of means $\{(x_m, t_m)\}_{m=0}^{M}$ and variances $\{\sigma_m\}_{m=0}^{M}$, the force and value function of a position x at time t are given by the coefficients $\{f_m\}_{m=0}^{M}$ and $\{v_m\}_{m=0}^{M}$ as

$$F(x;t) = F(x) + \sum_{m=0}^{M} f_m e^{-\frac{(x-x_m)^2+(t-t_m)^2}{2\sigma_m}}$$
$$V(x;t) = \sum_{m=0}^{M} v_m e^{-\frac{(x-x_m)^2+(t-t_m)^2}{2\sigma_m}}; \tag{4.45}$$

where initially the basis sets are a grid of 31 21 Gaussians in the x-t space. The Gaussians in time are spaced uniformly between $t \in [0;T)$, with standard deviations equal to half the grid-spacing. A third of the Gaussians in space are placed between $x \in [-4;-0.5]$, a third in $x \in (-0.5;1.5)$ and a third in $x \in [1.5;5]$. These three families of Gaussians each have standard deviations half of the corresponding grid spacings. We initialize all $f_m = v_m = 0$.

We consider the performance of the three algorithms differing in the gradient used to optimize them. These include an algorithm that uses no value function (MCR), one that uses a value baseline (MCVB), and one that uses a value function for future returns with $\gamma = 0.1$ (AC). We evaluate the efficiency of the algorithms by comparing learning curves, convergence with respect to basis, and properties of the learnt dynamics, shown in Fig 4.1. All figures comparing different algorithms use the same noise history and the same amount of statistics, such that the differences are solely ascribed to the learned dynamics. The MCR algorithm uses a learning rate of $\eta = 0.4$. The MCVB algorithm learning rates $\eta = 0.4$; $\lambda = 50$, and the AC algorithm learning rates $\eta = 1$; $\lambda = 0.05$.

In Figs. 4.1(a-c), we show learning curves for the total return, the average of the indicator observable, and the KL divergence, generated with 12 trajectories at each optimization step for each of the three algorithms. We have compared the results obtained with this finite basis to the numerically exact value of the optimal return and the corresponding observable average and KL divergence, obtained from Eq. 4.42 where for free diffusion the distribution is known. We find that while all three algorithms quickly achieve a dynamics which mostly fulfills the indicator function conditioning, the MCR algorithm struggles to optimize the KL divergence cost, while the MCVB and AC algorithm achieve converged values efficiently. As expected, each algorithm provides a variational estimate to the CGF with the MCVB and AC outperforming MCR. Trajectories with the final learned dynamics for the three algorithms are plotted in Fig. 4.1(d-f). The MCR algorithm finds forces that constrain the bridge trajectories too excessively, which results in the suboptimal estimate of the KL divergence. The AC trajectories are closest to the optimal bridge trajectories [109] while the MCVB trajectories lie in between. The main reason for the difference in performance in the three algorithms is the resultant suppression in the statistical errors in the gradient

estimate. This is illustrated in Figs. 4.1(g-h) where the convergence of the gradients of the 31 Gaussian coe cients at a time slice of $t = 0:7$ is shown for both MCR and MCVB. Since the learning rate is same in both algorithms, the large suppression of uctuations in the MCVB learning curves results from a more statistically converged gradient estimate using a value function. This suppression of gradient errors at limited statistics in the MCVB and AC algorithms is directly illustrated in Section 4.6.

We have studied the convergence of the KL divergence estimate towards the optimal value extracted from the numerically exact CGF, using the MCVB algorithm with an increasing position and time basis. We increased the number of time Gaussians, from 21 to 41 to 81, to observe the KL divergence cost shrinking as the ner grained force can better support the singular indicator function condition at the end of the trajectory. We also ran the optimization with a much bigger basis of 41_x 201_t Gaussians, and used 248 trajectories at every optimization step and learning rates $= 5$; $= 1000$. The Gaussians in x have standard deviations equal to half the grid spacing, while the Gaussians in t have standard deviations equal to a third of the grid spacing. While the estimate increased, in this particular problem, obtaining the numerically exact KL divergence would require use of still ner-grained Gaussians in space and time in order to represent the singularities of the edges of the target region and of the last timestep.

Barrier crossing with multiple reaction pathways

We now investigate the ability of the three algorithms to nd the optimal dynamics in two-dimensional barrier-crossing problems, the rst involving a potential allowing for multiple reaction pathways. The two-dimensional potential $U(x)$ we consider[2] has two minima and two degenerate reaction pathways involving the upper and lower halves of the $x = (x; y)$ plane as illustrated in Fig. 4.2. Barrier-crossing from one well to another is a rare event occurring with one randomly chosen pathway [72]. Without prior knowledge of the possibility of multiple reaction paths, path sampling algorithms typically need special techniques to discover them [298]. We use our reinforcement learning algorithms to compute an optimal force $F(x; t)$ that reproduces unbiased and uncorrelated reaction paths.

The reference equation of motion we consider is

$$dx = r \quad U(x) + \sqrt{2}dW \tag{4.46}$$

where the matrix G is proportional to the identity. We use a discretization timestep of 0.001. The trajectories start from the minimum of the left well, at $(x; y) = (1:11; 0)$, and are allowed to run for a duration of $T = 1:5$ and checked for reaching the right target well de ned as $x > 0; U(x; y) < 0$. This small region centered around (1.11,0) is used as for de ning the indicator function observable. The value of T has been chosen to be slightly greater than the typical transition path timescale, such that the optimized force should reproduce trajectories that follow the natural steady-state uctuations of the system. As long as the choice of

[2] The potential we use is $U(x; y) = 4=3[4(1 \quad x^2 \quad y^2)^2 + 2(x^2 \quad 2)^2 + ((x + y)^2 \quad 1)^2 \quad ((x \quad y)^2 \quad 1)^2 \quad 2]$