Statistical Physics and Algorithms

Dana Randall
School of Computer Science, Georgia Institute of Technology, Atlanta, GA 30332-0765, USA
randall@cc.gatech.edu

Abstract

The field of randomized algorithms has benefitted greatly from insights from statistical physics. We give examples in two distinct settings. The first is in the context of Markov chain Monte Carlo algorithms, which have become ubiquitous across science and engineering as a means of exploring large configuration spaces. One of the most striking discoveries was the realization that many natural Markov chains undergo phase transitions, whereby they are efficient for some parameter settings and then suddenly become inefficient as a parameter of the system is slowly modified. The second is in the context of distributed algorithms for programmable matter. Self-organizing particle systems based on statistical models with phase changes have been used to achieve basic tasks involving coordination, movement, and conformation in a fully distributed, local setting. We briefly describe these two settings to demonstrate how computing and statistical physics together provide powerful insights that apply across multiple domains.

2012 ACM Subject Classification Theory of computation → Random walks and Markov chains; Mathematics of computing → Stochastic processes; Theory of computation → Self-organization

Keywords and phrases Markov chains, mixing times, phase transitions, programmable matter

Digital Object Identifier 10.4230/LIPIcs.STACS.2020.1

Category Invited Talk

Funding Funded in part by NSF awards CCF-1526900, CCF-1637031, CCF-1733812 and ARO MURI award #W911NF-19-1-0233.

1 Introduction

Statistical physics employs probabilistic techniques to study systems consisting of large populations. The underlying principles explain numerous physical phenomena, such as magnetism, changes in states of matter, thermal radiation, noise in electronic devices, and more (see, e.g., [16]). In addition, these scientific insights help explain collective behavior across disciplines, including interacting biological systems [22], colloidal mixtures from chemistry [3, 21], segregation models from economics [5, 25], and random graph models in combinatorics [9].

Throughout theoretical computer science, we also find many examples where a statistical physics perspective has enriched the design and analysis of algorithms. A significant example concerns the role of *phase transitions*, showing how micro-scale behavior can induce global, macro-scale changes to a system (see, e.g., [4, 7, 26]). For example, phase transitions in random structures allow us to identify emergent characteristics of a configuration space, such as the birth of the giant component [19]. Moreover, Markov chains have been shown to undergo phase changes in their convergence times, transitioning from *disordered phases*, where they converge to (near) stationarity in polynomial time, to *ordered phases* that require exponential time [7, 10, 26]. More recently, algorithms exhibiting particular phase changes from disordered gaseous or liquid phases to ordered solid phases have proven effective for the design of distributed algorithms for robot swarms and active matter, where we seek collective organization achieving certain tasks [1, 11, 12].

37th International Symposium on Theoretical Aspects of Computer Science (STACS 2020).
Editors: Christophe Paul and Markus Bläser; Article No. 1; pp. 1:1–1:6
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

Physical systems define probability measures favoring configurations that minimize energy. Each configuration σ has energy determined by a *Hamiltonian* $H(\sigma)$ and a corresponding weight $w(\sigma) = \mathrm{e}^{-B \cdot H(\sigma)}$, where $B = 1/T$ is inverse temperature. The *Gibbs* (or *Boltzmann*) distribution assigns probabilities proportional to their weight $w(\sigma)$, where configurations with the least energy $H(\sigma)$ have the highest weight and are most likely. However, if there are few of these higher weight configurations, the sample space may be dominated instead by those with small weight, simply because there are many more of them (i.e., there is higher *entropy*), giving these configurations much higher probability overall, even if they are individually less probable. The thermodynamic properties of a physical system, such as specific heat and free energy, are derived from statistical properties of these distributions, and discontinuities in any of these quantities indicate a phase transition between states of matter.

We will provide a small window into the rich marriage between statistical physics and algorithms in the context of Markov chains and programmable matter. Markov chains can become prohibitively slow once the energy from the Hamiltonian outweighs the effects of entropy and the system transitions to an ordered state. In contrast, for programmable matter, we purposefully design algorithms that achieve distinct collective behaviors in their disordered and ordered phases, leading to robust distributed algorithms for self-organizing particle systems.

2 Local Markov Chains

Markov chain Monte Carlo (MCMC) algorithms are ubiquitous throughout science and engineering, providing useful tools for approximate counting, combinatorial optimization and modeling. The main idea is to perform a random walk among a set of configurations so that samples drawn from the limiting distribution are meaningful. For this to be useful, these algorithms need to be efficient, and indeed bounding the converence time of a Markov chain is often the critical step in establishing the efficiency of approximation algorithms based on random sampling. For example, if $G = (V_1, V_2, E)$ is a bipartite graph with $E \subseteq V_1 \times V_2$, then sampling perfect matchings on G allows us to estimate the permanent of the adjacency matrix [20]. Calculating the permanent of a matrix was shown by Valiant to be $\#P$-complete [27], or as hard as counting solutions to any NP-complete problem, so solutions that efficiently produce estimates approximating the exact count are the best we can expect.

Markov chains based on local moves, known as *Glauber dynamics*, are common in practice, primarily because of ther simplicity. As an example, consider the following chain that can be used to sample from the set of independent sets in a given graph, known in statistical physics as the *hard-core lattice gas model.* Given a graph G, the state space Ω is the set of independent sets. We are also given an input parameter λ, known as the *fugacity* (or *activity*). Our goal is to sample from the Gibbs distribution

$$\pi(I) = \lambda^{|I|}/Z,$$

where $|I|$ is the size of independent set I and $Z = \sum_{J \in \Omega} \lambda^{|J|}$ is the normalizing constant known as the *partition function.* We define the Glauber dynamics so that we can move between pairs of configurations that differ by a single vertex, and the celebated Metropolis Algorithm tells us how to implement these moves so that we converge to the Gibbs distribution π, as follows. Starting at any configuration $\sigma \in \Omega$, say the empty independent set (with no vertices), we repeat the following: choose a vertex v at random; if v is in the current independent set, remove it with probability $\min(1, \lambda^{-1})/2$; if it is not in the independent set, add it with probability $\min(1, \lambda)/2$, if possible; in all other cases, the independent set remains unchanged. It is simple to show that this chain is ergodic and converges to π, so our goal is to determine if it is efficient.

An interesting phenomenon occurs as λ is varied. For small values of λ, Glauber dynamics converge quickly to stationarity, while for large values it is prohibitively slow. To see why, imagine the underlying graph G is an $n \times n$ region of $\mathbb{Z}^2$. Large independent sets dominate the stationary distribution π when λ is sufficiently large and lie on one of the two sublattices (corresponding to each of the two colors of the checkerboard coloring on the dual lattice). When λ is large, it will take exponential time to move from an independent set that lies mostly on the odd sublattice to one that is mostly even. Currently, the best known rigorous bounds verify the the Markov chain converges in polynomial time whenever $\lambda < 2.538$ [26] and requires exponential time when $\lambda > 5.365$ [7, 8].

This type of dichotomy is well known in the statistical physics community, where many models have been shown to abruptly transition from a disordered state to a predominantly ordered one. Physicists observe phase transitions when extending Gibbs distributions to infinite lattices and studying whether there is a unique limiting Gibbs measure, known as a *Gibbs state* (see, e.g., [14]). For the hard-core model on $\mathbb{Z}^2$, is believed that there exists a critical value λ_c such that for $\lambda < \lambda_c$ there is a unique Gibbs state, while for $\lambda > \lambda_c$ there are multiple Gibbs states. This has been verified for small and large values of λ bounded away from the conjectured critical point $\lambda_c \approx 3.79$ in both the computational and physics settings [7, 26].

Fortunately, insights from statistical physics can also allow us to design alternative approaches to sampling in the slow regimes, in some cases. One approach that has proven fruitful far below the critical point (in the slow regime) is based on the *cluster expansion* [18]; at sufficiently low temperatures, configurations have long-range order, and can be precisely defined as small, randomized perturbations from some ground state, or highest probability state. Then configurations can be sampled by first randomly picking a ground state, and then inserting random defects with the appropriate conditional probabilities. A second approach uses *simulated tempering* or *parallel tempering* to sample at low temperatures by dynamically adjusting temperatures up and down during each simulation. These algorithms can be effective when we can (i) generate random samples from a family of temperatures so that low temperature configurations of interest arise often enough, and with the correct conditional probabilities, and (ii) the composite Markov chain on the larger state space (including configurations at all temperatures) converges quickly, even if it is prohibitively slow at low temperatures [6]. Finally, in some contexts it may be possible to rewrite the partition function as a sum over a different family of configurations, and this new representation may suggest alternative Markov chains that are quickly converging, even at low temperatures (see, e.g., [17, 28]).

3 Programmable Matter

Systems of programmable matter can be viewed as collections of simple interacting components with constant-size memory and limited computational capacity. We are interested in how these systems can be made to self-organize to produce emergent behaviors, such as coordination and collective movement.

Using a stochastic approach based on Markov chains, we can design rigorous and robust distributed algorithms for programmable matter exhibiting various desirable properties. For example, for the *compression* problem, our goal is to design an algorithm that allows an interacting particle system to self-organize and gather together compactly. We say a connected particle system on a planar lattice is *α-compressed* if the perimeter of the ensemble is at most α times the minimum perimeter possible for the n particles, $p_{min} = \Theta(\sqrt{n})$. In [12], we gave a distributed, local Markov chain-based algorithm that solves the compression problem for connected particle systems under the *geometric amoebot model* [13], a formal distributed model in which particles move on the triangular lattice.

Our approach to this and other basic tasks proceeds as follows. We first choose a Hamiltonian $H(\sigma)$ over particle configurations that assigns lower values to preferable (compressed) configurations. The transitions of the Markov chain are then defined to favor configurations with small Hamiltonians. For compression, we let $H(\sigma) = -e(\sigma)$, where $e(\sigma)$ is the number of edges induced by configuration σ, i.e., the number of lattice edges with both endpoints occupied. Setting $\lambda = \mathrm{e}^B$, we get $w(\sigma) = \lambda^{e(\sigma)}$. It is easy to verify that the number of induced edges negatively correlated with the size of the perimeter, so the more induced edges, the more compressed a configuration will be.

Using a Metropolis filter, we design a Markov chain $\mathcal{M}$ that performs local moves and converges to a distribution that generates configurations proportional to their weight $w(\sigma)$. In particular, the probability of a configuration σ is $w(\sigma)/Z$, where $Z = \sum_{\sigma'} w(\sigma')$ is the normalizing constant known as the *partition function.* Using tools from both statistical physics and Markov chain analysis, we prove that, if we wait long enough, non-compressed configurations occupy an exponentially small fraction of probability distribution when λ is sufficiently large.

The Markov chain $\mathcal{M}$ for compression is defined as follows. Starting with an arbitrary configuration σ_0 of n simply connected particles, we define local rules that maintain connectivity throughout the algorithm. There is a *bias parameter* λ given as input, where $\lambda > 1$ corresponds to particles prefering more neighbors and $\lambda < 1$ corresponds to particles prefering fewer neighbors. The moves of the Markov chain $\mathcal{M}$ are carefully designed so that the particle system always remains simply connected, preventing the chain to disconnect or form holes, which still keeping the state space connected via allowable transitions (so $\mathcal{M}$ is ergodic). Moreover, the moves are defined locally so that they can be implemented in a fully distributed setting. Maintaining connectivity makes the analysis of the limiting distribution simpler, but showing ergodicity is more challenging.

Particles individually execute a distributed algorithm defined by $\mathcal{M}$, using Poisson clocks to define when to attempt local moves. We prove that for all $\lambda > 2 + \sqrt{2}$, there is a constant $\alpha = \alpha(\lambda) > 1$ such that at stationarity, with all but exponentially small probability, the particle system will be α-compressed. In fact, we show that for any $\alpha > 1$, there exists λ such that our algorithms achieve α-compression. Moreoever, when λ is small we achieve the inverse property of *expansion.* For all $0 < \lambda < 2.17$, there is a constant $\beta < 1$ such that at stationarity, with all but exponentially small probability, the perimeter will be β-expanded, i.e., the perimeter will be within a β fraction of the maximum perimeter $p_{max} = \Theta(n)$. This implies that for any $0 < \lambda < 2.17$, the probability that the particle system is α-compressed is exponentially small for any constant $\alpha > 1$.

The key ingredient used to establish compression and expansion is a careful *Peierls argument*, used in statistical physics to study non-uniqueness of limiting Gibbs measures and in computer science to establish slow mixing of Markov chains. Because we enforce connectivity throughout the Markov process, our Peierls arguments are significantly simpler than many standard arguments on configurations that are not required to be connected. In subsequent work, we extended these results to the disconnected setting where, in contrast, verifying ergodicity becomes trivial but analyzing the stationary distribution requires more sophisticated tools [15].

One appeal of such a stochastic, distributed algorithm is its robustness. The system can recover from deviations in Poisson clocks waking particles to perform moves, anomalies in our individual particle's movements, and even some particle failures. Moreoever, this stochastic approach provides a general framework that is applicable beyond compression – it has the potential to solve any problem where the objective can be described in terms of minimizing some energy function, provided changes in that energy function can be calculated using only local information. One example is an optimization problem inspired by ant behavior [23] known as *shortcut bridging* where particles maintain bridge structures that balance a efficiency-cost tradeoff [1, 2]. A second example is a self-organizing system achieving *separation*, where particles of different colors can be shown to either intermingle or segregate depending on the settings of parameters [11]. Distributed algorithms based on Markov chains also have provided a theoretical explanation of *phototaxing*, or directed collective motion towards or away from a light source, in an experimental system of swarm robots [24]. Finally, we have promising directions for *alignment* and *flocking*, where oriented particles coordinate to determine a preferred direction of movement. In many of these cases, the collective behavior can be controlled by adjusting whether a physical system is in a disordered (gaseous) or an ordered (solid) state by exploring the physical properties of these systems.

References

1 Marta Andrés Arroyo, Sarah Cannon, Joshua J. Daymude, Dana Randall, and Andréa W. Richa. A stochastic approach to shortcut bridging in programmable matter. In *DNA Computing and Molecular Programming*, DNA23, pages 122–138, 2017.

2 Marta Andrés Arroyo, Sarah Cannon, Joshua J. Daymude, Dana Randall, and Andréa W. Richa. A stochastic approach to shortcut bridging in programmable matter. *Natural Computing*, 17(4):723–741, 2018.

3 Akhilest K. Arora and Raj Rajagopalan. Applications of colloids in studies of phase transitions and patterning of surfaces. *Current Opinion in Colloid & Interface Science*, 2(4), 1997.

4 R. J. Baxter, I. G. Enting, and S. K. Tsang. Hard-square lattice gas. *Journal of Statistical Physics*, 22:465–489, 1980.

5 Prateek Bhakta, Sarah Miracle, and Dana Randall. Clustering and mixing times for segregation models on $\mathbb{Z}^2$. In *Proceedings of the 25th ACM/SIAM Symposium on Discrete Algorithms*, (SODA), 2014.

6 Nayantara Bhatnagar and Dana Randall. Simulated tempering and swapping on mean-field models. *Journal of Statistical Physics*, 164(3):495–530, 2016.

7 Antonio Blanca, Yuxuan Chen, David Galvin, Dana Randall, and Prasad Tetali. Phase coexistence for the hard-core model on $\mathbb{Z}^2$. *Combinatorics, Probability and Computing*, pages 1–22, 2018.

8 Antonio Blanca, David Galvin, Dana Randall, and Prasad Tetali. Coexistence and slow mixing for the hard-core model on $\mathbb{Z}^2$. In *Approximation, Randomization and Combinatorial Optimization (APPROX/RANDOM)*, volume 8096, pages 379–394, 2013.

9 Bela Bollobas. The evolution of random graphs. *Transactions of the American Mathematical Society*, 286(1):257–274, 1984.

10 Christian Borgs, Jennifer T. Chayes, Jeong Han Kim, Alan Frieze, Prasad Tetali, Eric Vigoda, and Van Ha Vu. Torpid mixing of some Monte Carlo Markov chain algorithms in statistical physics. In *Proceedings of the 40th Annual Symposium on Foundations of Computer Science*, FOCS '99, pages 218–229, Washington, DC, USA, 1999. IEEE Computer Society.

11 Sarah Cannon, Joshua J. Daymude, Cem Gökmen, Dana Randall, and Andréa W. Richa. A local stochastic algorithm for separation in heterogeneous self-organizing particle systems. In *Approximation, Randomization and Combinatorial Optimization (APPROX/RANDOM)*, pages 54:1–54:22, 2019.

12 Sarah Cannon, Joshua J. Daymude, Dana Randall, and Andréa W. Richa. A Markov chain algorithm for compression in self-organizing particle systems. In *Proceedings of the 2016 ACM Symposium on Principles of Distributed Computing*, PODC '16, pages 279–288, New York, NY, USA, 2016. ACM.

13 Zahra Derakhshandeh, Robert Gmyr, Andréa W. Richa, Christian Scheideler, and Thim Strothmann. An algorithmic framework for shape formation problems in self-organizing particle systems. In *Proceedings of the Second Annual International Conference on Nanoscale Computing and Communication*, NANOCOM '15, pages 21:1–21:2, 2015.

14 Roland L. Dobrushin. The problem of uniqueness of a gibbsian random field and the problem of phase transitions. *Functional Analysis and Its Applications*, 2:302–312, 1968.

15 Bahnisikha Dutta, Shengkai Li, Sarah Cannon, Joshua J. Daymude, Enes Aydin, Andrea W. Richa, Daniel I. Goldman, and Dana Randall. Programming robot collecitves using mechanics-induced phase changes, 2020. (In preparation).

16 Sacha Friedli and Yvan Velenik. *Statistical Mechanics of Lattice Systems: A Concrete Mathematical Introduction*. Cambridge University Press, Cambridge, 2017.

17 Vivek K. Gore and Mark R. Jerrum. The Swendsen–Wang process does not always mix rapidly. *Journal of Statistical Physics*, 97(1):67–86, 1999.

18 Tyler Helmuth, Will Perkins, and Guus Regts. Algorithmic pirogov-sinai theory. In *Proceedings of the 51st Annual ACM Symposium on the Theory of Computing*, pages 1009–1020, 2019.

19 Svante Janson, Donald Knuth, Tomasz Luczak, and Boris Pittel. The birth of the giant component. *Random Structures and Algorithms*, 4(3):231–358, 1993.

20 Mark R. Jerrum, Alistair Sinclair, and Eric Vigoda. A polynomial-time approximation algorithm for the permanent of a matrix with non-negative entries. *Journal of the ACM*, 51:671–697, 2004.

21 Sarah Miracle, Dana Randall, and Amanda Pascoe Streib. Clustering in interfering binary mixtures. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques*, APPROX '11, RANDOM '11, pages 652–663, 2011.

22 Gerald H. Pollack and Wei-Chun Chin (Eds.), editors. *Phase Transitions in Cell Biology*. Springer International Publishing, 2008.

23 Chris R. Reid, Matthew J. Lutz, Scott Powell, Albert B. Kao, Iain D. Couzin, and Simon Garnier. Army ants dynamically adjust living bridges in response to a cost–benefit trade-off. *Proceedings of the National Academy of Sciences*, 112(49):15113–15118, 2015.

24 William Savoie, Sarah Cannon, Joshua J. Daymude, Ross Warkentin, Shengkai Li, Andréa W. Richa, Dana Randall, and Daniel I. Goldman. Phototactic supersmarticles. *Artificial Life and Robotics*, 23(4):459–468, 2018.

25 Thomas C. Schelling. Models of segregation. *American Economic Review*, 59(2):488–493, 1969.

26 Alistair Sinclair, Piyush Srivastava, Daniel Stefankovic, and Yintong Yin. Spatial mixing and the connective constant: Optimal bounds. *Probabability Theory Relatated Fields*, 168(1–2):153–197, 2017.

27 Leslie Valiant. The complexity of computing the permanent. *Theoretical Computer Science*, 8:189–201, 1979.

28 Jian-Sheng Wang and Robert H. Swendsen. Nonuniversal critical dynamics in monte carlo simulations. *Physics Review Letters*, 58(2):86–88, 1987.