Distribution Agreement

In presenting this thesis as a partial fulfillment of the requirements for a degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis in whole or in part in all forms of media, now or hereafter now, including display on the World Wide Web. I understand that I may select some access restrictions as part of the online submission of this thesis. I retain all ownership rights to the copyright of the thesis. I also retain the right to use in future works (such as articles or books) all or part of this thesis.

Ginger Lau March 29, 2024

Extremal Optimization for Ground States of the Three-Spin Mean Field Spin Glass

By

Ginger Lau

Stefan Boettcher, Ph.D.
Advisor

Physics

Stefan Boettcher, Ph.D.
Advisor

Daniel Sussman, Ph.D.
Committee Member

Jed Brody, Ph.D.
Committee Member

2024

Extremal Optimization for Ground States of the Three-Spin Mean Field Spin Glass

By

Ginger Lau

Stefan Boettcher, Ph.D.
Advisor

An abstract of
a thesis submitted to the Faculty of Emory College of Arts and Sciences
of Emory University in partial fulfillment
of the requirements of the degree of
Bachelor of Science with Honors

Physics

2024

Abstract

Extremal Optimization for Ground States of the Three-Spin Mean Field Spin Glass
By Ginger Lau

While the ground state energy of the $p = 2$ mean field spin glass has been analytically obtained by Parisi, the analogous $p = 3$ spin glass has not yet been solved. Instead, a variety of numerical methods, simulations, and heuristics have been developed to approximate complex combinatorial optimization problems such as spin glasses. This honors thesis project applied the extremal optimization (EO) heuristic to estimate the ground state energy of the three-spin mean field spin glass. In this research, two key modifications were made to the existing version of EO. First, multiple spins are now flipped at each time step in a parallelized update scheme to increase efficiency. Second, an auxiliary matrix is now used to hold intermediate fitness calculations and reduce run duration. The heuristic contains two tuning parameters: t, total number of update steps, and τ, a parameter to determine how many spins to flip once the system reaches a local minimum. Values for these quantities were chosen to be $t = (0.3)N^3$ and $\tau = 1.4$. Applying the modified EO to system sizes up to $N \approx 100$, the ground state energy density of the three-spin mean field spin glass in the $N \to \infty$ limit was obtained to be $\langle e_\infty \rangle = -0.4708(1)$. This value is within a 0.001 error of previous studies in the literature. The finite-N scaling correction was obtained to be $\omega = 4/5$ through fitting methods.

Extremal Optimization for Ground States of the Three-Spin Mean Field Spin Glass

By

Ginger Lau

Stefan Boettcher, Ph.D.
Advisor

A thesis submitted to the Faculty of Emory College of Arts and Sciences
of Emory University in partial fulfillment
of the requirements of the degree of
Bachelor of Science with Honors

Physics

2024

Acknowledgments

I would like to thank my research advisor, Stefan, for all his support during my time at Emory. Approximately two years ago, I began research with him knowing Python, the Ising spin glass model, and that was about it. Fast forward to now, not only did I learn a few more programming languages, but I learned very key principles in both computing and physics. I will especially remember our discussions on how to represent data clearly – humans like looking at straight lines!

I would also like to thank Dr. Sussman and Dr. Brody for taking the time to serve on my thesis committee. I really appreciate Dr. Sussman's diligence in annotating my thesis draft with comments. They were very valuable, and I undoubtedly would have never caught that stray comma on line 10 of page 4 otherwise. Taking Dr. Brody's quantum computing and digital electronics/microprocessors classes has reinforced that I really enjoy the intersection between physics and computers. Furthermore, the complementary knowledge I gained in his classes helped me to think more creatively about my research codes. Lastly, I would like to express my appreciation to Ms. Susan for her unrelenting support to us physics undergraduates, and to my friends for attending my thesis defense.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Optimization problems permeate physics, biology, computer science, business, and many other fields. Whether in survival of the fittest or cost-benefit analysis, many methods have been developed to find the optima of their respective systems. As the complexity of a system increases, it becomes increasingly difficult to develop efficient optimization schemes.

One combinatorial optimization problem in physics is determining the ground state of a spin glass. These systems, characterized by geometric frustration, are therefore difficult to solve for analytically. This research focuses on the Sherrington-Kirkpatrick (SK) [19] spin glass model. In the SK model, spin orientation is binary. The model can also be described as a $p = 2$ mean field spin glass, meaning that bonds describe the relationship between exactly two spins. There exist two possible bond relations: spins *aligning* in the same direction *decrease* the energy of the system, or spins with *opposite* orientations decrease the energy of the system.

For the $p = 2$ mean field spin glass, an analytical solution was famously calculated by Parisi [18] to be $\langle e_\infty \rangle = -0.7633$. Many numerical methods have also been applied to the spin glass problem, such as gradient descent [5, 10], simulated annealing [12], and extremal optimization [7]. For the $p = 3$ mean field spin glass (one bond connecting

three spins), an analytical solution has not yet been determined. However, numerical studies [1] have obtained values of $\langle e_\infty \rangle = -0.46950(6)$.

In this honors thesis, the extremal optimization heuristic [6, 8] is applied to the $p = 3$ mean field spin glass optimization problem. The goal is to approximate the ground state energy (normalized by number of spins) in the limit $N \to \infty$. To accomplish this, two main modifications were made to the existing EO heuristic: improving efficiency through parallelizing updates, and reformulating the data storage structure of the heuristic to accommodate $p > 2$ problems. The data collected in this study were ground state energies for thousands of different bond instances, for system sizes up to $N \approx 250$. From results up to system sizes $N \approx 100$, a good approximation of the ground state energy per spin is extrapolated to the $N \to \infty$ limit. We calculated this value to be $\langle e_\infty \rangle = -0.4708(1)$.

In the following sections, Chapter 2 describes the EO heuristic and provides an introduction to spin glasses. Chapter 3 introduces the three-spin problem and details the EO code development process. In Chapter 4, experiments to determine optimal user-tuned parameters are performed, and comparisons to solvable models are made. Chapter 5 describes the data collection and analysis process to obtain an estimate for the ground state energy density of a $p = 3$ mean field spin glass. Lastly, Chapter 6 summarizes the EO modifications and three-spin SK spin glass optimization results, and provides suggestions for future work.

Chapter 2

Background

2.1 Extremal Optimization

The extremal optimization (EO) local search heuristic was developed by Boettcher and Percus [8] to solve combinatorial optimization problems. In this heuristic, each member of a system is assigned a fitness value to describe how well it optimizes the system. To dynamically update the system, low fitnesses are identified and replaced with new, random ones to evolve the system towards a local fitness maxima.

A key feature of EO is that instead of choosing only the *lowest* (worst) fitness to update, a tuning parameter τ is introduced to allow choosing of *any* fitness. A preference for lower fitnesses still remains. This key feature eliminates deterministic loops that may arise from only choosing the lowest fitness, while still maintaining that lower fitnesses are chosen at a higher rate. The occasional random replacement of an average or above average fitness, then, drives the system out of a local optimum and allows the exploration of many local optima, in an attempt to find the global optimum. The relatively-simple EO evolution can be applied to a wide variety of optimization problems, and this paper focuses on the application of EO to determine the ground state of a spin glass.

2.1.1 Self-Organized Criticality

Extremal optimization was inspired by the Bak-Sneppen model based on self-organized criticality (SOC). First proposed by Bak, Tang, and Wiesenfeld [4, 2], SOC describes a method in which complexity in large systems arises. SOC systems are open nonequilibrium systems driven into a steady state characterized by a broad distribution of fluctuations. Changes occur via intermittent "avalanches," rather than through a smooth and gradual evolution. In this framework, the stable states are called "critical steady states" in which small disturbances draw the system through broadly distributed chain reactions.

Following the landmark SOC paper, the Bak-Sneppen model [3, 15] applies SOC to describe the evolution of an ecosystem. To begin, each constituent species is assigned a fitness. This numerical value describes the likelihood of survival, analogous to its definition in biology. Fitness of a species is also correlated to the fitnesses of the other species in the system; predator and prey would have an inverse fitness relationship, for example. As a result of this complex coupling, changes in the fitness of one species affects the fitnesses *and* the potential future adaptations of other species in the ecosystem.

Through interactions within the system, the ecosystem will quickly settle into a local maximum of species' fitness values. The system can only leave optima states by going to a worse state before it can find new, potentially better, optima. This mechanism is unbiased and does not involve adaptive logic to identify *better* optima; rather, the mechanism searches many states and "learns" by adding explored states to memory.

2.2 Spin Glasses

Spin glasses are magnetic systems in which a property of the atoms forming the material, called *spin*, has a preferred orientation. This orientation need not be orderly or follow any specific pattern. Rather, the specific random couplings of the spin glass lead to a ground state in which a seemingly random spin configuration optimizes the energy.

One example of a spin glass is the Sherrington-Kirkpatrick (SK) model [19]. In this model, spin orientation is binary (Ising); the two possible states are described as "up" or "down." Interactions between spins are defined by "plus" and "minus" bonds. Two spins connected by a "plus" bond want to be oriented in the same direction, and two spins connected by a "minus" bond want to be oriented in opposite spin directions. Given a set of spins, the SK model describes a fully-connected system, where each spin has a bond relation with every other spin in the system. Therefore, in a system of N spins, each of the spins will have N bonds (counting the "bond" of a spin to itself as having a value of zero). The system can then be described by N^2 bonds. Recognizing that a bond value is shared between two spins, the final expression for number of bonds is $\frac{1}{2}N^2$.

The spin glass optimization problem is then as follows: *Given a specific set of bond relations in a frustrated spin glass system, what is the optimal orientation for each of the spins to obtain the ground state (lowest) energy?*

To describe the degree of bonds that are satisfied in a system, the Hamiltonian

$$H = -\frac{1}{2}\sum_{i}\sum_{j} J_{ij}\sigma_i\sigma_j, \tag{2.1}$$

can be used, where lower energies correspond to more favorable states. In the equation, σ_i and σ_j denote two spins, and J_{ij} denotes the bond between them. A summation is performed over all pairs, and then the value is divided by two to account for symmetric

couplings. In the literature [14], theoretical calculations use Gaussian-distributed bonds J_{ij}. However, when applying this problem to computational methods, a "plus" bond is represented as the integer 1 and "minus" as -1, for convenience. Both representations are identical in leading behavior and second moment, with differences arising only in higher orders. Therefore, integer bond values $J_{ij} = \pm 1$ are used in this computational study.

Chapter 3

Approach

3.1 Three-Spin SK Model

Another spin glass is the $p = 3$ mean field spin glass, otherwise known as the 3-spin Sherrington-Kirkpatrick spin glass. In this model, *three* spins (instead of two) are connected by *one* bond relation. Therefore, the meaning of a "satisfied" bond now changes. Table 3.1 shows each possible 3-spin configuration to satisfy a "plus" bond and a "minus" bond. The Hamiltonian for a 3-spin SK system is adjusted to

$$H = -\frac{1}{6}\sum_{i}\sum_{j}\sum_{k} J_{ijk}\sigma_i\sigma_j\sigma_k, \tag{3.1}$$

where the factor of $\frac{1}{6}$ divides out repeated counting of the same bond J_{ijk} in the summation. Again, the bonds are Gaussian-distributed in theoretical representations, but computational research can use $J_{ijk} = \pm 1$ due to identical leading behavior.

3.2 Three-Spin Extremal Optimization

To employ the EO heuristic to optimize the 3-spin SK model, firstly, spin orientations are randomly initialized. Next, an N^3 matrix is filled with values representing the

Bond satisfied	Spin configuration
+	↑ ↑ ↑
+	↑ ↓ ↓
+	↓ ↑ ↓
+	↓ ↓ ↑
−	↑ ↑ ↓
−	↑ ↓ ↑
−	↓ ↓ ↓
−	↓ ↑ ↑

Table 3.1: Spin configurations to satisfy a "plus" and "minus" bond in the 3-spin SK model.

bond relation between every grouping of three indices. Then, the fitness λ_i of each spin σ_i is calculated using

$$\lambda_i = \frac{1}{6}\sigma_i \sum_j \sum_k J_{ijk}\sigma_j\sigma_k, \tag{3.2}$$

such that the sum of all spin fitnesses equals the negative of the Hamiltonian in Eq. 3.1. Fitnesses in the 3-spin SK model contain one more level of summation compared to the 2-spin SK model. Next, to evolve the system, spins are flipped at each time step. The fitnesses of each spin are recalculated after each flip, and the total energy is adjusted accordingly.

The main dynamics of EO, selecting which spins to flip at each time step, differs from what was originally described in [7]. A parallelized update scheme is now introduced. Rather than flipping only *one* spin per time step, *multiple* spins can now be changed. To begin Phase I of the update procedures, the heuristic flips spins with negative fitness at a given probability rate. At each time step, the heuristic attempts to conduct Phase I. If at least one spin is flipped, the heuristic moves to the next time step once all spins under the criteria are flipped. If no spins are identified to flip, it means the system has reached a state where all (or almost all) fitnesses are positive, signaling a local energy minimum. At this point, the heuristic moves into Phase II to complete the time step.

Once a local energy minimum is reached (most or all spins have positive fitness), the tuning parameter τ is utilized in Phase II of the update procedures. This parameter decides how many spins to flip in order to "kick" the system out of a local minimum so it can continue traversing the energy landscape. This value is important because it balances the heuristic between not enough randomness (getting stuck in a local minimum) and too much randomness (a simple random walk). The exact mechanism to use τ is described in Sec. 3.4. Once Phase II is successfully performed, the heuristic resumes with Phase I at the next time step and the process repeats. The approximate split between how often Phase I versus Phase II occurs is 90% versus 10% of update time steps, respectively.

Two methods of choosing spins to flip are described later in Sec. 3.4.1, but the key concept is that the higher the fitness of the spin(s) flipped, the more radically the entire system is altered. There is no bias towards or away from visited states, but rather, the method of choosing spins includes mechanisms to avoid back-and-forth flipping of potentially coupled spins. Therefore, visiting the same state twice is unlikely. After this "kick," the system should have a considerable amount of negative fitnesses again, and the process of flipping and updating fitnesses is repeated until another local minimum is found, and so on. The heuristic terminates when a certain number of time steps have been completed. The best local minimum with the lowest energy is determined to be the global minimum, and the corresponding spin configuration is considered the ground state of the given bond configuration.

3.3 EO Heuristic Initialization

The EO algorithm, as described in [6], has been modified for the research in this honors thesis. Namely, a Python implementation is chosen to utilize the `numpy` package array operations. Furthermore, the EO heuristic must be adapted for 3-spin interactions,

since the existing EO was only for 2-spin interactions. To this end, the representation of bond relations and spin fitness definitions must be modified.

To establish the system, initial values for N spins and N^3 bond relations are chosen. The bond relations are represented as a matrix with dimensions $N \times N \times N$. The indices of each bond matrix entry identify which three spins the bond relation connects. The number of unique bonds can be represented with only a three-dimensional triangular matrix; however, all N^3 values are filled to create a symmetric matrix for ease of future calculations. For mutually distinct indices $i \neq j \neq k$, the bond value is $J_{ijk} = \pm 1$, and everywhere else with one or more identical indices is zero. As a result, an unrestricted summation can be used for the Hamiltonian, with nonsensical bond relations contributing zero to the total energy.

To initialize the N spins, each is randomly assigned a value $\uparrow = +1$ or $\downarrow = -1$ with equal probability. Once the bond matrix and spin array are initialized, the initial spin fitnesses can be calculated using Eq. 3.2. In adapting the 2-spin EO to the 3-spin EO, an auxiliary matrix, γ, is introduced to facilitate intermediate fitness calculations. The γ matrix is an $N \times N$ matrix, with each entry of the matrix taken to be

$$\gamma_{ij} = \sigma_i \sigma_j \sum_k J_{ijk} \sigma_k. \tag{3.3}$$

This allows fitness to be redefined as

$$\lambda_i = \frac{1}{2} \sum_j \gamma_{ij}, \tag{3.4}$$

which changes the computational complexity of the fitness update procedure, as will be shown in Sec. 3.4.3 via Fig. 3.1.

To fill the γ matrix, a nested `for` loop over $0 \leq j < i$ within a `for` loop over $0 \leq i < N$ is run. For each i, j combination, the bond submatrix J_{ij} is taken. This will be a row vector. Then, using `numpy` for array operations, the bond submatrix vector

is multiplied element-by-element to the spin vector (representing all σ_k). The sum of the elements of this resultant vector is taken, then multiplied by the corresponding σ_i and σ_j, and then stored to the γ_{ij} matrix element. Once the auxiliary matrix γ is filled, the fitness λ_i of each spin σ_i is calculated by summing over the row (or column; the matrix is symmetric) with the index corresponding to the spin index. Finally, the initial energy of the system is calculated by taking

$$H = -\frac{1}{3}\sum_i \lambda_i. \tag{3.5}$$

3.4 EO Heuristic Update Procedure

The dynamical update procedure used in this study differs from previous EO work in the literature [8, 7, 6] because a *parallel* spin flip update scheme is used instead. At each time step, rather than only choosing one spin to flip (using a rank-ordered scheme as in [7]), the parallel-update EO flips multiple spins. Therefore, a threshold to determine which spins to flip must be established.

Several methods were tested to determine how the heuristic should handle states at or near a local minimum, a state characterized by all or almost all the spins having positive fitnesses. The foundation of the update methodology is similar to the 2-spin EO. The degree of randomness in choosing spin updates is introduced through the τ-dependent power law:

$$y = \left[1 - x\left(1 - \frac{1}{N^{(\tau-1)}}\right)\right]^{-\left(\frac{1}{\tau-1}\right)}. \tag{3.6}$$

To use this to generate a probability density function for EO, x is restricted to a univariate probability density function with $0 \leq x < 1$. The range of outputs of Eq. 3.6 therefore lies in $1 \leq y < N$. Outputs y are then rounded to the nearest integer, and this value $q = \text{int}(y)$ is used in Phase II of the heuristic update procedure (described

in Sec. 3.4.2).

3.4.1 Code Development: Parallel Spin Updates

The first attempt at defining the update procedure involved flipping spins with negative fitness at a 50% probability rate for Phase I until the system gets close to a local minimum. When the system reaches a time step where no spins are chosen to flip (because no negative fitnesses exist or none were chosen due to the 50% probability rate), the heuristic moves into the Phase II treatment of a local minimum. In Phase II, a "threshold" value is chosen by randomly drawing a value from the distribution in Eq. 3.6 and rounding the output to the nearest integer. Now, all the spins with *fitnesses* below this threshold value are flipped with a 50% probability. If no spins were randomly chosen to flip, then the process of choosing a new threshold would be repeated, drawing a new value from Eq. 3.6 and *adding* it to the existing threshold. This ensures that even though a configuration is energetically favorable, at least one spin will be flipped each time step.

The main issue with this approach is in Phase II. The outputs of Eq. 3.6 are restricted between $1 \leq q < N$, while fitness values exist on a much wider scale. Therefore, the repeated drawing and adding of values from this distribution in order to reach usable threshold values is not computationally effective. This reveals that the heuristic treatment of states near local minima should not use the fitness scale to create a threshold using Eq. 3.6. It also reveals that the 50% flip rate may be too low, since the threshold needed to be increased constantly despite having spins already under the fitness threshold.

The next attempt involved first flipping 80% of spins with negative fitness during Phase I until a local minimum is approached. This 80% rate seems to strike a good balance between flipping enough spins to efficiently evolve the system, but does not force all candidate spins to be flipped and thus avoids deterministic loops. Then, when

the system is at or near a local minimum, Eq. 3.6 is used to determine the *number* of spins, of the N total, to flip in Phase II. This differs from the previous attempt by not using Eq. 3.6 on a scale of spin *fitnesses*, but instead using a scale of *number* of spins. Once this threshold value, $q = \text{int}(y)$, is chosen in Phase II (a random value is drawn from the distribution only once!), the goal is to flip the q number of spins with the lowest fitnesses.

3.4.2 Parallel Spin Update Scheme

The final method tested was successful. It eliminates dependence on a fitness scale by ordering the spin fitnesses into a histogram and selecting spins bin-by-bin. Each histogram has a bin width of one fitness unit. Spins to flip are identified bin-by-bin beginning with the lowest fitness, until the next bin contains more spins than the remaining number of spins to flip. With this final bin, each spin to flip is chosen with a probability calculated by dividing number of remaining spins to flip by number of spins in the final bin.

Once all q spins to flip are identified, they are flipped one at a time by multiplying spin value by -1. After each individual spin flip, all fitnesses, and subsequently the total energy, must be updated. Fitness updates must occur in parallel before another spin is flipped; otherwise, calculations will no longer match the given spin configuration.

3.4.3 Code Development: Auxiliary Matrix

In developing the 3-spin EO heuristic, the auxiliary matrix γ was hypothesized to cut down redundancy of calculations, and therefore reduce overall usage of computational resources. Two different code implementations were tested, one using the γ matrix and the other without (it only used summation loops). System sizes of $N = [16, 24, 32, 48, 64, 96, 128, 192, 256, 350, 512, 750]$ were tested. A Python profiler was used to

measure the duration (in seconds) for runs of each of the two different implementations. The fitness update procedure was run for 100 spin flips, repeated five times per system size, and then averaged.

Fig. 3.1 shows the time difference and confirms that using the γ matrix does help cut down total run duration in the regime that this research focuses on. However, as system size increases, run duration of the fitness update subroutine increases linearly if the γ matrix is not used. Contrary to the hypothesis, run duration increases more steeply for the runs that used the γ matrix.

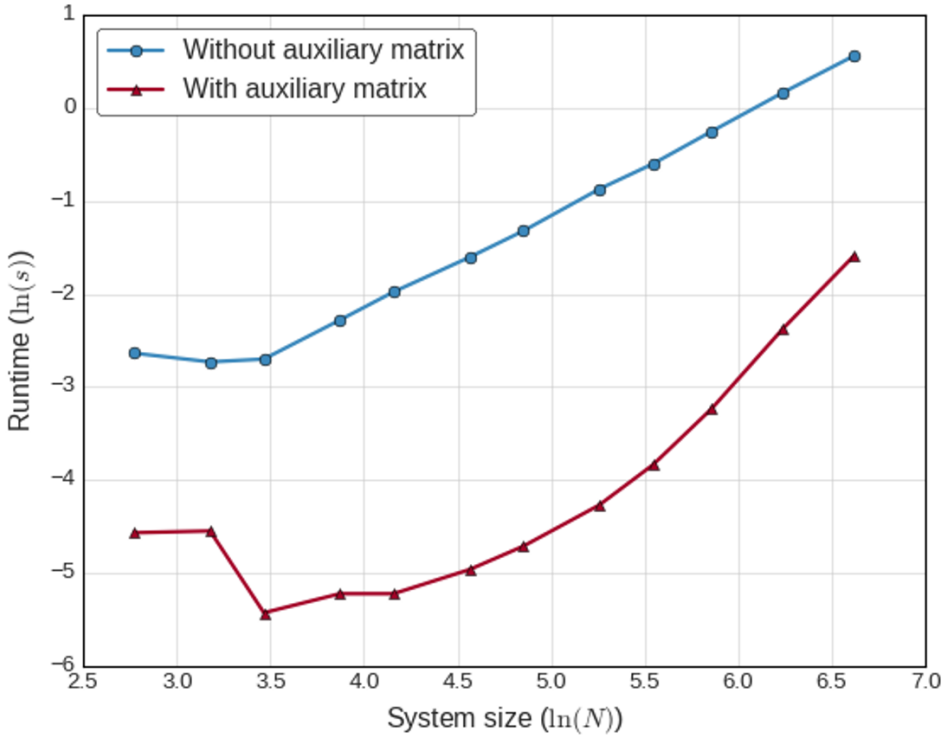


Figure 3.1: Comparison of using an auxiliary matrix γ in intermediate fitness calculations versus using only nested loops for the full summation. The blue line with circles represents not using the auxiliary matrix, and it lies above the red line with triangles representing the use of the auxiliary matrix.

3.4.4 Fitness Update Scheme

To update fitness values, rather than recalculating the Hamiltonian, a fitness adjustment is calculated instead. If spin σ_α is flipped,

$$\begin{aligned} \gamma'_{ij} &= \sigma_i \sigma_j \left[\sum_{k \neq \alpha} J_{ijk} \sigma_k + J_{ij\alpha} \sigma'_\alpha \right] \\ \gamma'_{ij} &= \gamma_{ij} + 2 \left[J_{ij\alpha} \sigma_i \sigma_j \sigma'_\alpha \right], \end{aligned} \tag{3.7}$$

where σ'_α represents the spin value after it is flipped and γ'_{ij} applies only where both $i \neq \alpha$ and $j \neq \alpha$. The row $i = \alpha$ and column $j = \alpha$ of the auxiliary matrix γ are simply changed as follows

$$\gamma'_{\alpha j} = \sigma'_\alpha \sigma_j \sum_k J_{ijk} \sigma_k = -\gamma_{\alpha j}, \tag{3.8}$$

$$\gamma'_{i\alpha} = \sigma_i \sigma'_\alpha \sum_k J_{ijk} \sigma_k = -\gamma_{i\alpha}, \tag{3.9}$$

by multiplying by -1 due to the spin flip.

The fitness values for each spin σ_i are then recalculated by summing over the ith row of the γ matrix. The total energy of the system is updated by subtracting λ'_α from the previous energy. Once all spins from the chosen candidates have been flipped (updating the fitness after every spin flip), the time step is completed and the process of identifying new spins to flip repeats.

Chapter 4

Experiments

4.1 Runtime and τ Test

While EO is a simple algorithm, two parameters must be chosen by the user. The first is the total runtime t. Runtime is measured in number of time steps; one time step is one pass of either Phase I or Phase II. The second parameter is the tuning parameter τ used in the Eq. 3.6 probability distribution to determine the degree of flipping at Phase II time steps. Larger τ corresponds to fewer spins flipped, and smaller τ leads more spins flipped in the time step. The quantity of interest in optimizing is the energy *density* of the 3-spin SK spin glass, not the Hamiltonian in Eq. 3.1. Energy density divides out N-dependence from the Hamiltonian for better comparison across system sizes.

To determine an optimal t and τ combination, a study on system sizes $N = [32, 64, 128]$ was conducted. Combinations of the values $\tau = [0.7, 1.2, 1.7, 2.2, 2.7, 3.2]$ and runtimes $t = [1N^2, 2N^2, 4N^2, 8N^2]$ were tested. In this study, 100 different random bond configuration instances were used. Each bond instance was run 30 times, with a different spin orientation initialization each run. The runtime $t = 1N^2$ was run 16 times. The runtime $t = 2N^2$ was repeated 8 times, $t = 4N^2$ was repeated 4 times, and

$t = 8N^2$ was repeated twice. The results are shown in Fig. 4.1.

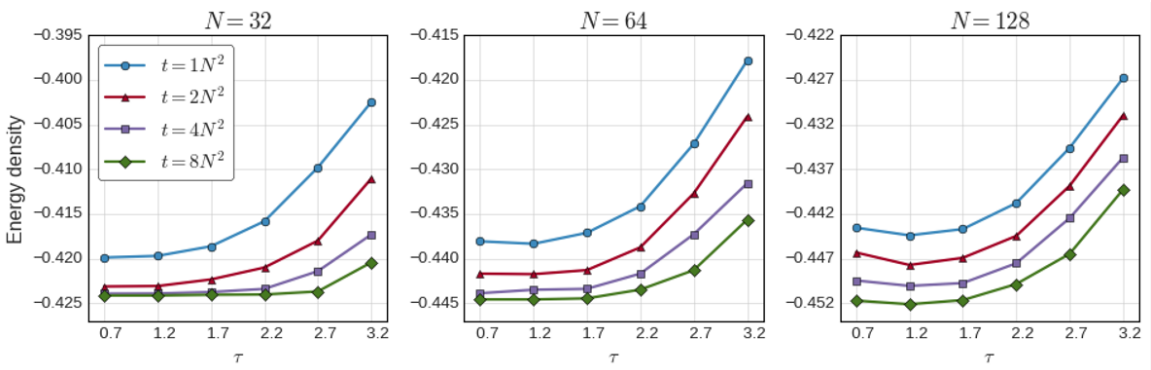


Figure 4.1: Results of the initial runtime and τ combination test. System sizes $N = [32,\ 64,\ 128]$ were each tested over four runtimes $t = [1N^2,\ 2N^2,\ 4N^2,\ 8N^2]$ and τ values between $\tau = [0.7,\ 3.2]$. For all runs, energy density increases as $\tau > 2.2$, and as more runtime is added, energy density decreases. Note that the energy scale for each system size N is different.

Based upon the results, it appears that increasing runtime lowers energy; therefore, a sufficient runtime has not yet been determined. In the next test, runtime based on N^3 time is tested. Combinations of $\tau = [0.8,\ 1.1,\ 1.4,\ 1.7,\ 2.0]$ and $t = [(0.25)N^3,\ (0.50)N^3,\ (0.75)N^3,\ (1.00)N^3]$ were tested for $N = 32$ and $N = 64$. Each runtime (rounded to the nearest integer) is repeated six times for each bond instance (each run again initialized with different spins), over a total of 10 bond instances. The results are shown in Fig. 4.2.

4.1.1 Optimal τ Value

The ideal τ value corresponds to a minimum in the energy density. It is worth noting that there exists an optimal *range* of τ, not just one value. This minimum range will therefore be represented by a small plateau in the energy density versus τ plot. Initially, Fig. 4.1 reveals that any τ values above $\tau = 2.2$ do not need to be considered. This range corresponds to not enough randomness introduced, and the system therefore gets stuck in a local minimum. For the lower bound on τ, the $N = 128$ curves reveal where τ is too small by showing an increase in energy density. It is most apparent in

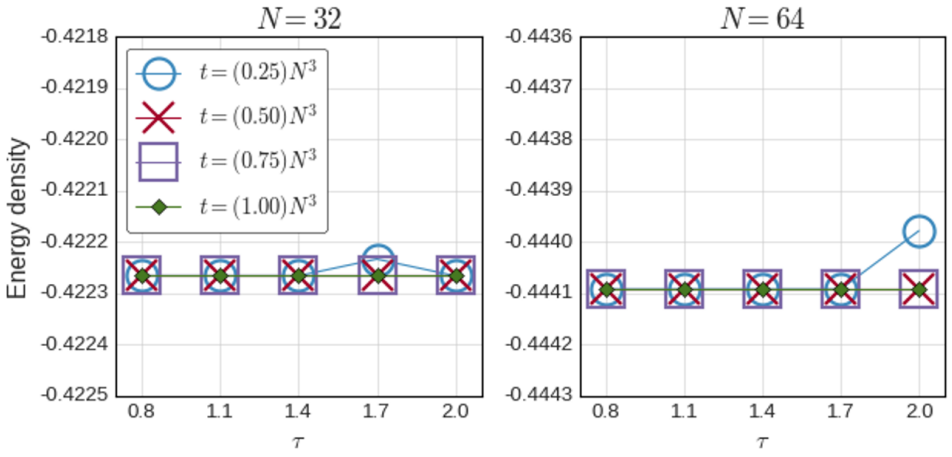


Figure 4.2: An extension of the first runtime and τ combination test. System sizes $N = [32,\ 64]$ were tested over four longer runtimes $t = [(0.25)N^3, (0.50)N^3, (0.75)N^3, (1.00)N^3]$ and τ in the narrower range $\tau = [0.8,\ 2.0]$. For nearly all the runtime and τ combinations, the same energy density value was obtained and the curves are therefore flat. Note that the energy scale for each system size N is different.

the $t = 2N^2$ line of the $N = 128$ plot of Fig. 4.1, where an increase in energy density near $\tau = 0.7$ establishes the existence of a minimum τ.

The results in Fig. 4.2 indicate that the τ parameter does not need to be tuned to very high precision. The four plots show that *any* τ in the $\tau = [0.8, 2.0]$ range is *capable* of obtaining the ground state configuration, so long as the runtime is sufficient. Therefore, for the study in the next chapter (Sec. 5.1), the value $\tau = 1.4$ is used because it lies in the approximate region near the energy density minimum of the $N = 128$ curves of Fig. 4.1.

4.1.2 Optimal Runtime

The optimal runtime was determined by analyzing a variety of factors. Since the energy density kept decreasing at longer runtimes in Fig. 4.1 (based on N^2 time), a cubic-based time scale was instead tested in Fig. 4.2. When using the N^3 runtimes, the energy densities were almost all the same across the increasing prefactors, indicating that the ground state was likely obtained. Additionally, observing that the energy

density stayed the same across the different τ (x-axis) values indicates that the runtimes significantly exceeded the minimum runtime required to obtain the ground state. Therefore, regardless of the exact value of τ, the ground state was still obtained in Fig. 4.2.

In the domain of system sizes used in this research, a value of $t = (0.3)N^3$ is chosen as the optimal runtime for the study in Sec. 5.1. In this regime, it appears to be the best tradeoff between success and efficiency. As shown later, this runtime may not be sufficient for larger system sizes, in which an adaptive scheme [6] should instead be used.

4.2 Testing Solvable Models: Ferromagnet

To test that the EO heuristic actually obtains ground state energies, a ferromagnetic and antiferromagnetic case are tested. The ground state energies of these specific systems can be calculated via the Hamiltonian in Eq. 3.1. For the ferromagnetic case, the EO bond matrix is manually set to all "plus" bonds between spins. The energy obtained by the EO heuristic indeed matches the expected ferromagnetic ground state of $H = -\binom{N}{3}$, with all spins pointing up. For the antiferromagnetic case, the EO bond matrix is set to all "minus" bonds between spins. The ground state energy obtained again matches the expected analytical solution $H = -\binom{N}{3}$, with all spins pointing down.

4.3 Branch-and-bound for Solvable Models

In the final phase of evaluating the EO heuristic for accuracy, small-N systems are solved using a branch-and-bound (BB) algorithm [13]. This method, popular for NP-hard optimization problems, systematically searches and tests many spin configurations to determine the ground state. It is therefore less efficient as system size grows large.

However, it serves as a powerful tool to corroborate the energies obtained by EO.

In this study, the same bond configurations were tested by both the EO heuristic and the BB algorithm. System size $N = 16$ was chosen to test because it was small enough for the BB algorithm to handle efficiently. The average ground state energy density over 300 given bond configuration instances was obtained by BB to be $\langle e_{16} \rangle = -0.3895$. Then, EO was run for the same 300 instances. For this test, a range of $\tau = [-3.2,\ 3.2]$ was run, at increments of $\tau = 0.5$. Each bond instance was repeated 24 times, with eight runs each at runtimes of $t = [(0.25)N^3,\ (0.50)N^3,\ (1.00)N^3]$. All EO results were averaged to obtain Fig. 4.3.

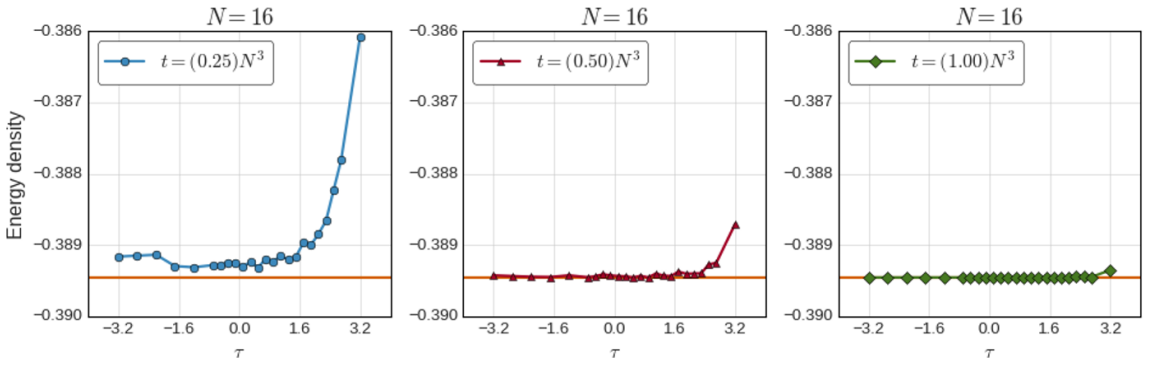


Figure 4.3: Extremal optimization and branch-and-bound results for $N = 16$ systems. Three runtimes $t = [(0.25)N^3,\ (0.50)N^3,\ (1.00)N^3]$ over a wide range of $\tau = [-3.2,\ 3.2]$ were tested. The ground state calculated by BB is shown as a red horizontal line near $\langle e_{16} \rangle = -0.3895$.

The branch-and-bound results offer a powerful insight into whether the EO heuristic is accurate and matches other predictions. As expected, it again confirms that runtime plays a large role in actually reaching the ground state. In this general domain, the ability to reach the ground state is shown to lie in the $t = (0.25)N^3$ to $t = (0.50)N^3$ range, given the proper τ chosen. It confirms that runtimes as long as $t = (1.00)N^3$ are not necessarily needed.

When evaluating τ, interesting behavior occurs because *negative* τ values seem to work just as well for the Phase II "kick." This study was performed before the authors

realized that larger N should be used to tune t and τ, hence why such a wide range of τ values is included. Fig. 4.3 seems to indicate that in the limit of performing a random walk ($\tau \to -\infty$), negative τ values work better. However, this conflicts with previous work [7] that shows τ should neither be too large nor too small. The authors therefore hypothesized that perhaps for small enough N, the system space is also small enough for a random walk to eventually find the ground state. To test this, the study in Fig. 4.1 was performed. Indeed, results for $N = 128$ reveal that too-small τ will not be able to obtain the ground state. Sec. 4.1.1 discusses this in more detail.

Regardless of the small system size N, the BB study results in Fig. 4.3 reveal that the ground state energy density approximated by the EO heuristic can match the true ground state energy density obtained by the branch-and-bound algorithm.

Chapter 5

Results and Analysis

5.1 Production Runs

The parameters chosen to produce the experimental data were determined based on the various studies in Sec. 4.1. For the production runs, the tuning parameter was $\tau = 1.4$ and maximum runtime was $t = (0.3)N^3$ rounded to the nearest integer. Each bond instance was run three times, each with a different spin initialization. The system sizes tested were $N = [16, 18, 21, 25, 32, 40, 51, 64, 90, 128, 192, 256]$. For each system size, Table 5.1 lists the total number of runs, the average energy density, standard deviation, and standard error.

5.2 Ground State Energy Density and Finite-Size Corrections

The aim of this study is to determine the ground state energy density for the 3-spin Sherrington-Kirkpatrick spin glass. Firstly, the data from Table 5.1 is plotted in Fig. A.1 with the average energy density of all runs per system size shown as large blue circles. As mentioned in Sec. 5.1, each bond instance is run three times. Of the three

N	I	$\langle e_N \rangle$	$\sigma(e_N)$	$\frac{\sigma(e_N)}{\sqrt{I}}$
16	160399	-0.3897	0.0274	0.00007
18	91562	-0.3965	0.0252	0.00008
21	126442	-0.4050	0.0224	0.00006
25	80922	-0.4136	0.0195	0.00007
32	25224	-0.4240	0.0160	0.00010
40	66522	-0.4316	0.0129	0.00005
51	19086	-0.4385	0.0103	0.00007
64	14838	-0.4440	0.0084	0.00007
90	3128	-0.4500	0.0060	0.00010
128	924	-0.4550	0.0040	0.00010
192	127	-0.4590	0.0030	0.00030
256	23	-0.4600	0.0030	0.00060

Table 5.1: Results of production runs for $N = [16, 18, 21, 25, 32, 40, 51, 64, 90, 128, 192, 256]$ to approximate the 3-spin SK ground state energy density. Column I represents number of bond instances run for the corresponding system size N. The average energy density is column $\langle e_N \rangle$, standard deviation of the energy density is $\sigma(e_N)$, and the standard error is $\frac{\sigma(e_N)}{\sqrt{I}}$. Note that this is one-third of the total data collected; from the three restarts per bond instance, only the best of the three is included in this set.

runs, the lowest energy density for each instance is taken, and the average for each system size is plotted as a small red square. Standard error bars for the set of best values per instance are included. The x-axis scale is chosen to be $1/N$ to identify the thermodynamic limit ($N \to \infty$) of energy density more easily; this value will therefore be the y-intercept.

The data set of best values per instance is fit to the linear equation

$$y = a + bx, \tag{5.1}$$

where y corresponds to energy density of a given system size $\langle e_N \rangle$, a corresponds to energy density in the thermodynamic limit $\langle e_\infty \rangle$, x equals $1/N$, and b is the slope of the line. This fitting method yields $\langle e_\infty \rangle = -0.4636$.

In Fig. A.1, the alignment of the data between the set of all values (blue circles) and the set of best values (red squares), as well as the narrowness of the error bars

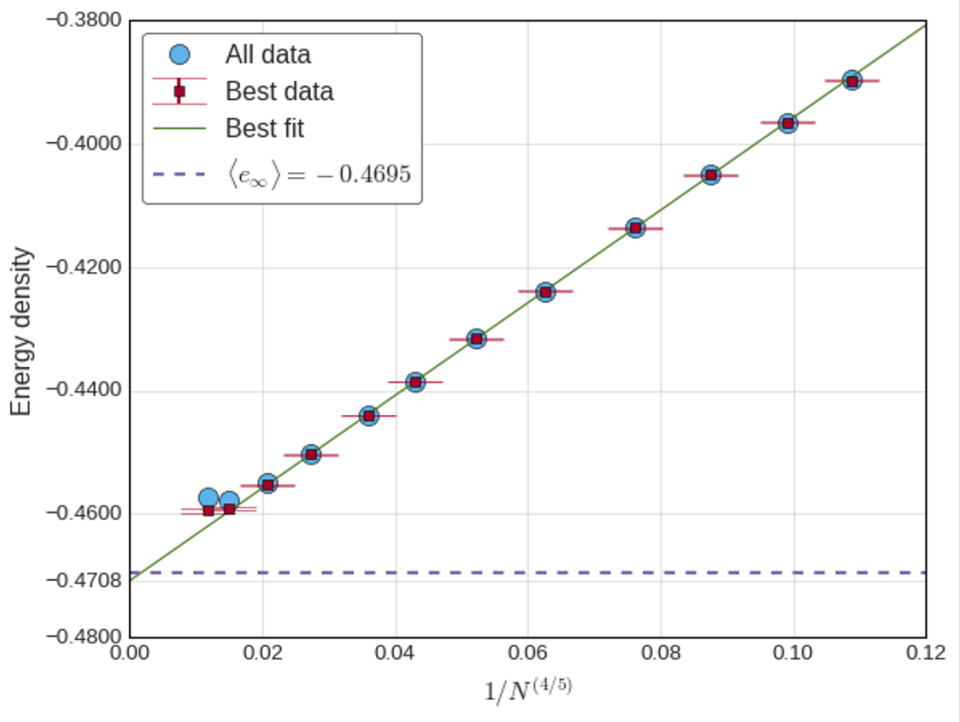


Figure 5.1: Extrapolation plot of energy density over $1/N^{\omega}$, where $\omega = 4/5$. The large blue circles represent the average of all data collected for each N, the red squares with error bars represent the average of the best results from each bond instance. The green line is the best fit line. The y-intercept of the best fit line represents $\langle e_\infty \rangle$, and is marked at $\langle e_\infty \rangle = -0.4708$ on the y-axis. The value obtained by Alaoui and Montanari [1] is shown as a horizontal dotted line at $\langle e_\infty \rangle = -0.4695$.

for systems up to $N = 128$, indicate that the results are good in this region. For the $N = 192$ and $N = 256$ points, the average of the best data deviates from the average of all data, and there exists a gap in error bars on this scale. These results likely have not reached the ground state yet given the chosen τ and runtime, and therefore should not be included in further analysis.

It must be noted that there is curvature of the data around the linear best fit line in Fig. A.1. Previous literature [17] states that for finite-N systems there exists a finite-size scaling correction to N. The next fit performed is to the asymptotic form

$$y = a + bx^c, \tag{5.2}$$

where c is the scaling correction ω for finite-size systems. In fitting, only the points for $N = [25, 32, 40, 51, 64, 90, 128]$ were used. The omission of the $N = 192$ and $N = 256$ due to excessive uncertainty has been explained above. The three smallest N have also been omitted to serve as a validation scheme. If the data from these three points align well to the results despite not being included while fitting, it is likely that the fit is accurate.

The results of fitting to Eq. 5.2 are $\langle e_\infty \rangle = -0.4708(1)$ and $\omega \approx -0.8$. This fit, plotted on an $x = 1/N^{0.8}$ scale, is shown in Fig. 5.1. Regarding the points $N = [16, 18, 21]$, they lie nicely on the line despite not being included to fit, which supports the validity of this fit.

5.3 General Mean Field Spin Glass Ground States

For the 2-spin ($p = 2$) Sherrington-Kirkpatrick spin glass, there exists an analytical solution. This was calculated by Parisi [18] to be $\langle e_\infty \rangle = -0.7633$. The finite-size scaling correction for a 2-spin SK spin glass has been largely agreed through analytic [17] and numerical [9, 16, 7] studies to be $\omega = 2/3$.

Such solutions for the 3-spin ($p = 3$) Sherrington-Kirkpatrick spin glass have not yet been analytically determined. However, other numerical methods have been employed by Alaoui and Montanari [1], using Gaussian-distributed bonds instead. As described in Sec. 2.2, the leading behavior is identical to our Hamiltonian. However, their Hamiltonian definition does differ by a factor of $-\sqrt{3}$, so adjusting their value $\frac{\langle e_\infty \rangle}{-\sqrt{3}} = -0.46950(6)$. This reveals that our approximation of $\langle e_\infty \rangle = -0.4708(1)$ is within 0.001 of the value in [1]. Alaoui and Montanari also state that local search algorithms generally have difficulty with the 3-spin SK problem, often getting stuck at ≈ 0.800 unable to cross the local minimum energy barrier. This suggests that our parallelized EO heuristic could potentially be more than just a *local* search heuristic.

Chapter 6

Conclusion

The goal of this honors thesis research was to approximate the ground state energy density of a $p = 3$ mean field spin glass. To accomplish this, an implementation of the extremal optimization heuristic was created in Python, with a few modifications. A parallel update scheme increased efficiency by flipping multiple spins in each time step. The introduction of an auxiliary matrix γ to aid in fitness calculations was shown to cut down run duration in the $N \approx 100$ domain.

The main tuning of the heuristic involved finding the right runtime and τ combination. The range of τ that could obtain the ground state turned out to be quite large. More importantly, sufficient runtime was required. The runtime and τ values ultimately used in this study were $t = (0.3)N^3$ and $\tau = 1.4$. During the testing phase, EO was applied to analytically-solvable models in order to check the accuracy of energies obtained. These solvable models included the (anti)ferromagnet and small systems solved by a branch-and-bound algorithm.

The results of this study obtained a ground state energy density of $\langle e_\infty \rangle = -0.4708(1)$, which is within a 0.001 error of existing estimations in the literature. The finite-size scaling correction was obtained to be $\omega = 4/5$ through fitting methods. In the $p \to \infty$ limit, the random-energy model [11] gives a finite-size scaling correction of

$\omega = \ln(N)/N$. Therefore, our value of $\omega \approx 4/5$ seems plausible for the 3-spin model as p increases from the 2-spin model.

Future work could include further parallelization via GPU to increase efficiency in optimizing larger systems. Additionally, this parallelized EO heuristic can be applied to other systems to further characterize performance and potentially solve other complex optimization problems.

Stefan Boettcher is an inventor on a patent related to this work filed by Emory University regarding the parallel-updated version of EO used in this work called Thresholded Extremal Optimization (TEO) with Patent Application No. 63/558,691, filed on 2/28/2024.

Appendix A

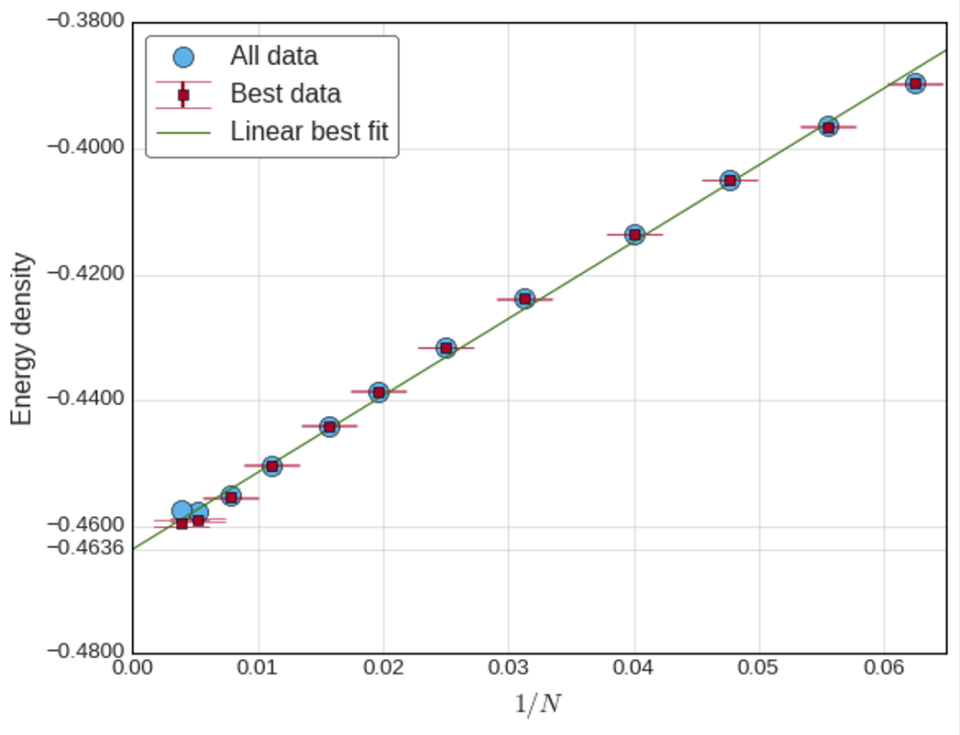


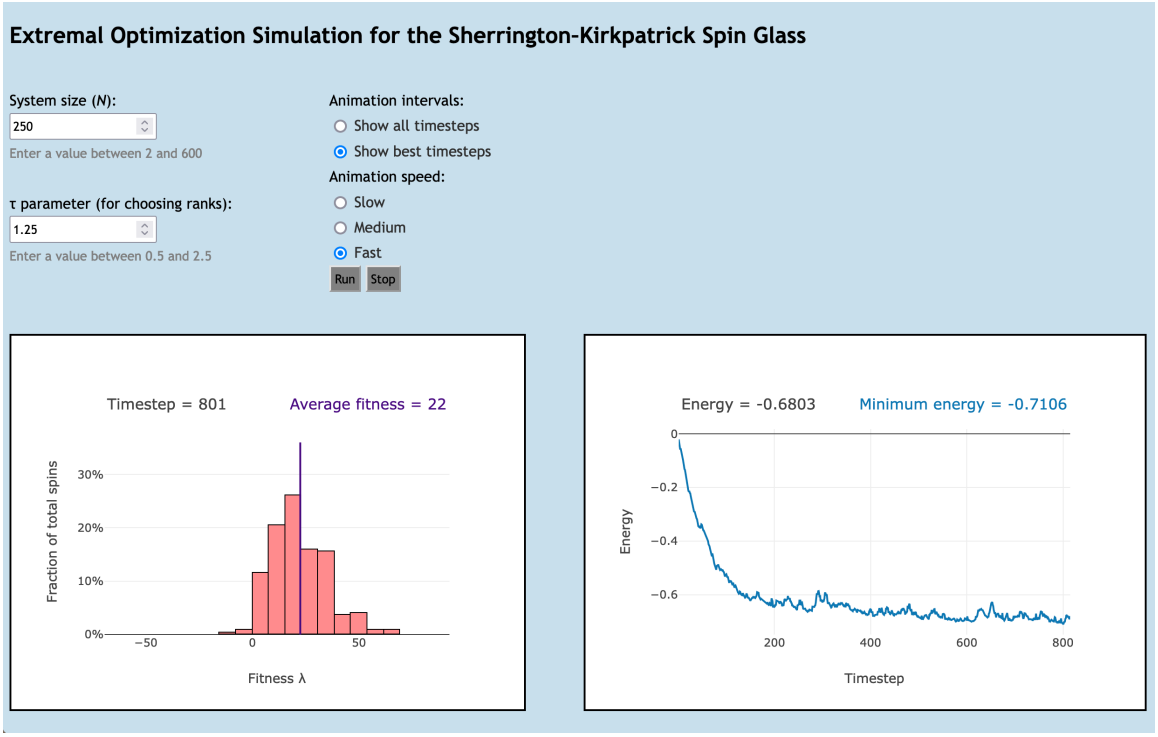
Figure A.1: Extrapolation plot of energy density over $1/N^{\omega}$, where $\omega = 1$. The large blue circles represent the average of all data collected for each N, and the red squares with error bars represent the average of the best results from each bond instance. The green line is the best fit line for the data represented by the red squares. The y-intercept of the best fit line represents $\langle e_{\infty} \rangle$, and is marked at $\langle e_{\infty} \rangle = -0.4636$ on the y-axis. The plot shows curvature around the linear fit, indicating that a scaling factor should be applied to N.

Appendix B

When I began working with Dr. Boettcher, my first project was to create a website illustrating EO dynamics, based on the original EO described in [8, 6]. The website

`https://gingyy.github.io/`

offers an interactive way to explore extremal optimization and the role of τ.



Bibliography

[1] Ahmed El Alaoui and Andrea Montanari. Algorithmic thresholds in mean field spin glasses. *arXiv preprint arXiv:2009.11481*, 2020.

[2] Per Bak. *How nature works: the science of self-organized criticality*. Springer Science & Business Media, 2013.

[3] Per Bak and Kim Sneppen. Punctuated equilibrium and criticality in a simple model of evolution. *Phys. Rev. Lett.*, 71:4083–4086, Dec 1993. doi: 10.1103/PhysRevLett.71.4083. URL `https://link.aps.org/doi/10.1103/PhysRevLett.71.4083`.

[4] Per Bak, Chao Tang, and Kurt Wiesenfeld. Self-organized criticality. *Physical review A*, 38(1):364, 1988.

[5] FT Bantilan Jr and RG Palmer. Magnetic properties of a model spin glass and the failure of linear response theory. *Journal of Physics F: Metal Physics*, 11(1): 261, 1981.

[6] Stefan Boettcher. *Extremal Optimization*, chapter 11, pages 227–251. John Wiley Sons, Ltd, 2004. ISBN 9783527603794. doi: https://doi.org/10.1002/3527603794.ch11. URL `https://onlinelibrary.wiley.com/doi/abs/10.1002/3527603794.ch11`.

[7] Stefan Boettcher. Extremal optimization for sherrington-kirkpatrick spin glasses.

The European Physical Journal B-Condensed Matter and Complex Systems, 46: 501–505, 2005.

[8] Stefan Boettcher and Allon G. Percus. Optimization with extremal dynamics. *Physical Review Letters*, 86(23):5211–5214, June 2001. ISSN 1079-7114. doi: 10.1103/physrevlett.86.5211. URL http://dx.doi.org/10.1103/PhysRevLett.86.5211.

[9] J-P Bouchaud, Florent Krzakala, and Olivier C Martin. Energy exponents and corrections to scaling in ising spin glasses. *Physical Review B*, 68(22):224404, 2003.

[10] S Cabasino, E Marinari, P Paolucci, and G Parisi. Eigenstates and limit cycles in the sk model. *Journal of Physics A: Mathematical and General*, 21(22):4201, nov 1988. doi: 10.1088/0305-4470/21/22/021. URL https://dx.doi.org/10.1088/0305-4470/21/22/021.

[11] Bernard Derrida. Random-energy model: Limit of a family of disordered models. *Physical Review Letters*, 45(2):79, 1980.

[12] Gary S Grest, CM Soukoulis, and K Levin. Cooling-rate dependence for the spin-glass ground-state energy: Implications for optimization by simulated annealing. *Physical review letters*, 56(11):1148, 1986.

[13] A. H. Land and A. G. Doig. An automatic method of solving discrete programming problems. *Econometrica*, 28(3):497–520, 1960. ISSN 00129682, 14680262. URL http://www.jstor.org/stable/1910129.

[14] Marc Mézard, Giorgio Parisi, and Miguel Angel Virasoro. *Spin glass theory and beyond: An Introduction to the Replica Method and Its Applications*, volume 9. World Scientific Publishing Company, 1987.

[15] Maya Paczuski, Sergei Maslov, and Per Bak. Avalanche dynamics in evolution, growth, and depinning models. *Phys. Rev. E*, 53:414–443, Jan 1996. doi: 10.1103/PhysRevE.53.414. URL `https://link.aps.org/doi/10.1103/PhysRevE.53.414`.

[16] Matteo Palassini. Ground-state energy fluctuations in the sherrington–kirkpatrick model. *Journal of Statistical Mechanics: Theory and Experiment*, 2008(10): P10005, 2008.

[17] G Parisi, F Ritort, and F Slanina. Several results on the finite-size corrections in the sherrington-kirkpatrick spin-glass model. *Journal of Physics A: Mathematical and General*, 26(15):3775, 1993.

[18] Giorgio Parisi. A sequence of approximated solutions to the sk model for spin glasses. *Journal of Physics A: Mathematical and General*, 13(4):L115, 1980.

[19] David Sherrington and Scott Kirkpatrick. Solvable model of a spin-glass. *Physical review letters*, 35(26):1792, 1975.