Parallel Markovian Algorithms and their application to combinatorial optimization

Alessio Troiani

Department of Mathematics and Computer Science
Univeristy of Perugia

Perugia - February 8, 2023

The quadratic Hamiltonian

Consider the quadratic form

$$H(\eta) = K \sum_{i,j} J_{ij} \eta_i \eta_j$$

- $K \in \mathbb{R}$
- $\eta \in \mathcal{A}^N$
- J_{ij} interacton between η_i and η_j

Possible questions of interest are

- finding the minimizer of H
- determine the statistical properties of $\min_\eta H(\eta)$

Different "models" depending on the values of J_{ij} and $\mathcal{A}$

Let $G = (V, E)$ be a graph. It is convenient to think

- $N = |V|$
- $J_{ij} = w(\{ij\})$.

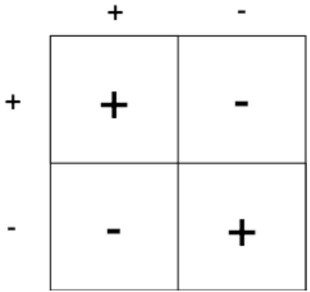
The quadratic Hamiltonian

Sherrington–Kirkpatrick model

$$H(\eta) = \frac{1}{\sqrt{N}} \sum_{i,j} J_{ij}\eta_i\eta_j$$

- $\eta_i \in \{-1, 1\}$
- J_{ij} i.i.d. Standard Normal Random Variables

SK model (max cut)



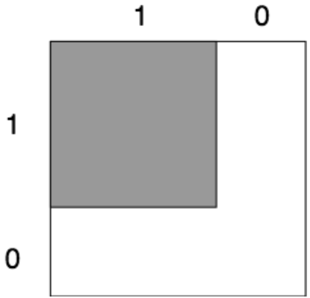
The quadratic Hamiltonian

Unconstrained Binary Quadratic Programming (main topic of this talk)

UBQP / QUBO
(Gaussian Mean Field Lattice Gas)

$$H(\eta) = \frac{1}{\sqrt{N}} \sum_{i,j} J_{ij} \eta_i \eta_j$$

- $\eta_i \in \{0, 1\}$
- $J_{ij} \in \mathbb{R}$



Can be used to represent a large class of discrete optimization problems.
Find the subgourp of data points with strongest correlation.

The quadratic Hamiltonian

Linear Programming

$$\min_{\eta} \sum_{i=1}^{N} c_i \eta_i$$
$$\eta_i \in \{0, 1\}$$

$$J_{ij} = \begin{cases} c_i & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$$

The quadratic Hamiltonian

The maximum clique problem

A clique, C, in an undirected graph $G = (V, E)$ is a subset of the vertices, $C \subset V$, subgraph of G induced by C is a complete graph.

A maximum clique of G, is a clique, such that there is no clique with more vertices.

$$J_{ij} = \begin{cases} -1 & \text{if } i = j \\ 0 & \text{if } (i, j) \in E \\ +M & \text{otherwise} \end{cases}$$

Linked to the cluster detection in large data sets and social networks.

The quadratic Hamiltonian

Other "classical" (lattice) problems in Statistical Mechanics, e.g:

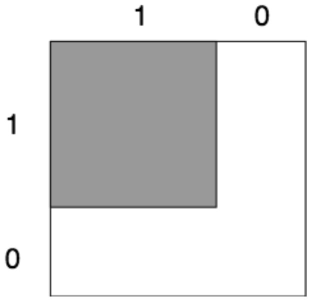
- Curie–Weiss model
  - $\eta_i \in \{-1, 1\}$
  - $J_{ij} = -J$ for $i \neq j$
- Ising model on $\mathbb{Z}^d$
  - $\eta_i \in \{-1, 1\}$
  - $J_{ij} = \begin{cases} -J & i, j \text{ neighboring sites on the lattice} \\ 0 & \text{otherwise} \end{cases}$
- Edward–Anderson model
  - $\eta_i \in \{-1, 1\}$
  - $J_{ij} \neq 0 \Leftrightarrow i, j$ neighboring sites on the lattice

Unconstrained Binary Quadratic Programming

$$H(\eta) = \frac{1}{\sqrt{N}} \sum_{i,j} J_{ij}\eta_i\eta_j$$

- $\eta_i \in \{0, 1\}$
- $J_{ij} \in \mathbb{R}$

UBQP / QUBO
(Gaussian Mean Field Lattice Gas)



Focus on

- finding the minimizer of H
- determine (some) statistical properties of $\min_\eta H(\eta)$

Gibbs Measure

How can we find te minima of H?

Idea: Sample from a probability distribution such that $P(H(\eta^\star)) \approx 1$ with $\eta^\star = \arg\max_\eta H(\eta)$.

Gibbs measure

$$\pi_\beta(\eta) = \frac{e^{-\beta H(\eta)}}{Z_\beta}$$

with $Z_\beta = \sum_\eta e^{-\beta H(\eta)}$ and $\beta > 0$.

It is straightforward to check that

$$\lim_{\beta \to \infty} \pi_\beta(\eta^\star) = 1.$$

Markov Chain Monte Carlo

How can we sample from the Gibbs measure?

Let X_t be a irreducible and aperiodic Markov Chain on $\mathcal{X}$ with transition probability on P. Then there is a unique probability distribution π on $\mathcal{X}$ such that $\pi P = \pi$ and

$$\lim_{t \to \infty} \|\mu^{(n)} - \pi\|_{\mathrm{TV}} = 0$$

The condition (detailed balance)

$$\pi_i P_{ij} = \pi_j P_{ji} \qquad \forall i, j \in \mathcal{X}$$

is enough to ensure π stationary for P.

Markov Chain Monte Carlo

Metropolis algorithm

$$P_{\eta,\tau} = \begin{cases} \frac{1}{N} e^{-\beta[H(\tau)-H(\eta)]_+} & \text{if } \eta \sim \tau \\ 1 - \sum_{\tau\sim\eta} \frac{1}{N} e^{-\beta[H(\tau)-H(\eta)]_+} & \text{if } \tau = \eta \\ 0 & \text{otherwise} \end{cases}$$

It is immediately checked that

$$\pi_\eta P_{\eta,\tau} = \pi_\tau P_{\tau,\eta}$$

Markov Chain Monte Carlo

Metropolis algorithm

$$P_{\eta,\tau} = \begin{cases} \frac{1}{N} e^{-\beta[H(\tau)-H(\eta)]_+} & \text{if } \eta \sim \tau \\ 1 - \sum_{\tau\sim\eta} \frac{1}{N} e^{-\beta[H(\tau)-H(\eta)]_+} & \text{if } \tau = \eta \\ 0 & \text{otherwise} \end{cases}$$

It is immediately checked that

$$\pi_\eta P_{\eta,\tau} = \pi_\tau P_{\tau,\eta}$$

Can we do "better"?

Probabilistic Cellular Automata

We are interested in *natively parallel* Markovian algorithms (instead of "single spin flip") to

- find the minimizers of $H(\eta)$
- draw samples from the Gibbs measure $\pi(\eta) = \frac{e^{-\beta H(\eta)}}{Z}$

A Probabilistic Cellular Automaton (PCA), is Markov Chain $(X_n)_{n \in \mathbb{N}}$ with state space $\mathcal{X} = \{1, \ldots, k\}^N$ whose transition probabilities are such that

$$P\{X_n = \tau | X_{n-1} = \sigma\} = \prod_{i=1}^{N} P\{(X_n)_i = \tau_i | X_{n-1} = \sigma\}.$$

- Possible fast(er) convergence to equilibrium measure
- Well adapted to be simulated on (massively) parallel processors

Probabilistic Cellular Automata

We will be interested in PCA defined as follows.

Let $G = (V, E)$ be a graph and let $H(\eta, \tau) = \sum_{i \in V} h_i(\eta)\tau_i$. We will consider transition probabilities of the type

$$P_{\eta,\tau} = \frac{e^{-H(\eta,\tau)}}{\sum_\tau e^{-H(\eta,\tau)}} = \prod_i \frac{e^{h_i(\eta)\tau_i}}{Z_i}$$

e.g.:

$$V = \Lambda \subset \mathbb{Z}^2 \quad E = \{\{i,j\} : i,j \in V, |i-j| = 1\}$$
$$\eta, \tau \in \{-1,1\}^V \quad h_i(\eta) = J(\eta_{i\downarrow}, \eta_{i\rightarrow}, \eta_{i\uparrow}, \eta_{i\leftarrow}) + q\eta_i + \lambda$$

Several results have been obtained for PCA of this type in the context of Ising models when each $h_i(\eta)$ depends on all neighbors of the spin at site x, e.g.:

- stationary measure of PCA
- relation of stationary measure of PCA with Gibbs measure

PCA - Application to UBQP

We define a transition matrix on $\{0,1\}^N$ as

$$P_{\eta,\tau} = \frac{e^{-H(\eta,\tau)}}{\sum_\tau e^{-H(\eta,\tau)}}$$

with

$$H(\eta,\tau) = \beta \sum_i h_i(\eta)\tau_i + q \sum_i [\eta_i(1-\tau_i) + \tau_i(1-\eta_i)]$$

where

- $h_i(\eta) = \frac{1}{\sqrt{N}} \sum_j J' ij \eta_j$,
- $J' = \frac{J+J^T}{2}$
- β is the inverse temperature
- q is a positive constant (inertial term)

Note that $H(\eta,\eta) = \beta H(\eta)$.

PCA - Application to UBQP

The transition matrix can be rewritten in the form

$$P_{\eta,\tau} = \prod_i \frac{e^{-\beta h_i(\eta)\tau_i - q[\eta_i(1-\tau_i)+\tau_i(1-\eta_i)]}}{Z_i}$$

which yields

$$P(\tau_i = 1|\eta) = \frac{e^{-\beta h_i(\eta) - q(1-\eta_i)}}{Z_i}$$

and

$$P(\tau_i = 0|\eta) = \frac{e^{-q\eta_i}}{Z_i}$$

where $Z_i = e^{-\beta h_i + q(1-\eta_i)} + e^{q\eta_i}$.

PCA - Application to UBQP

The reversible equilibrium measure of this PCA is

$$\pi(\eta) = \frac{\sum_\tau e^{-H(\eta,\tau)}}{\sum_{\eta,\tau} e^{-H(\eta,\tau)}}$$

since, because of the simmetry of J', the detailed balance condition is satisfied:

$$\frac{\sum_\tau e^{-H(\eta,\tau)}}{\sum_{\eta,\tau} e^{-H(\eta,\tau)}} P_{\eta,\tau} = P_{\tau,\eta} \frac{\sum_\eta e^{-H(\eta,\tau)}}{\sum_{\eta,\tau} e^{-H(\eta,\tau)}}$$

As q gets "large"

$$\pi(\eta) = \frac{H(\eta,\eta) + \sum_{\tau\neq\eta} e^{-H(\eta,\tau)}}{\sum_{\eta,\tau} e^{-H(\eta,\tau)}} \approx \pi_G(\eta)$$

PCA - Application to UBQP

N	Instance id	PCA	Metropolis
500	500a	0.415129916	0.415129916
	500b	0.424031186	0.424031186
1000	1000a	0.414470925	0.414470925
	1000b	0.412802104	0.412802104
2000	2000a	0.424186053	0.424186053
	2000b	0.416673303	0.416588939
4000	4000a	0.424745169	0.42479645
	4000d	0.415214004	0.415233809
8000	8000a	0.416367988	0.416174887
	8000d	0.421539704	0.421421773

PCA - Application to UBQP

Computation times for 10000 iterations

N	Metropolis (CPU-1 core)	PCA (CPU-4 cores)	PCA (GPU-P100)
500	616 ms	401 ms	1.8 s
1000	3.7 s	835 ms	2.0 s
2000	19.2 s	6.9 s	2.5 s
4000	79 s	23.9 s	3.6 s
8000	296 s	92 s	7.4 s
16000	$\approx$ 1200 s	356 s	23 s

UBQP: Theoretical results

Determine the statistical properties of $\min_\eta H(\eta)$

Conjecture

Let $\min_{\eta\in\{0,1\}^N} H(\eta) := H(\eta^*) := -M_N := -m_N N$
Then there exist $\bar{m} > 0$ *and* $0 < \bar{\alpha} < 1$ *such that for almost all* J

$$\lim_{N\to\infty} \frac{M_N}{N} = \lim_{N\to\infty} m_N = \bar{m} \qquad \lim_{N\to\infty} \frac{\sum_{i=1}^N \eta_i^*}{N} = \bar{\alpha}$$

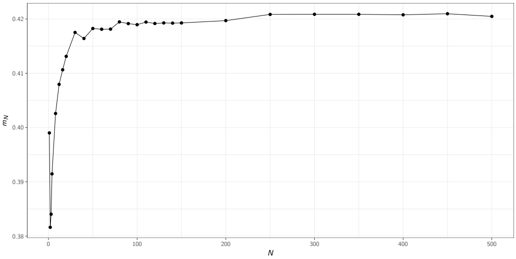
UBQP: Theoretical results

Determine the statistical properties of $\min_\eta H(\eta)$

Conjecture

Let $\min_{\eta\in\{0,1\}^N} H(\eta) := H(\eta^*) := -M_N := -m_N N$
Then there exist $\bar{m} > 0$ *and* $0 < \bar{\alpha} < 1$ *such that for almost all* J

$$\lim_{N\to\infty} \frac{M_N}{N} = \lim_{N\to\infty} m_N = \bar{m} \qquad \lim_{N\to\infty} \frac{\sum_{i=1}^N \eta_i^*}{N} = \bar{\alpha}$$



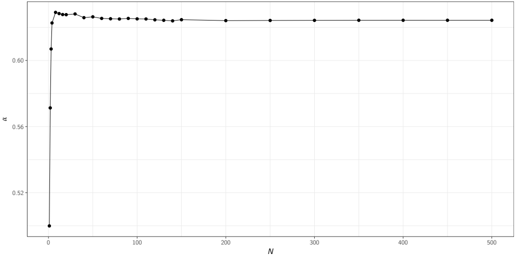
UBQP: Theoretical results

Determine the statistical properties of $\min_\eta H(\eta)$

Conjecture

Let $\min_{\eta\in\{0,1\}^N} H(\eta) := H(\eta^*) := -M_N := -m_N N$
Then there exist $\bar{m} > 0$ *and* $0 < \bar{\alpha} < 1$ *such that for almost all* J

$$\lim_{N\to\infty} \frac{M_N}{N} = \lim_{N\to\infty} m_N = \bar{m} \qquad \lim_{N\to\infty} \frac{\sum_{i=1}^N \eta_i^*}{N} = \bar{\alpha}$$



UBQP: Theoretical results

Determine the statistical properties of $\min_\eta H(\eta)$

Conjecture

Let $\min_{\eta\in\{0,1\}^N} H(\eta) := H(\eta^*) := -M_N := -m_N N$
Then there exist $\bar{m} > 0$ *and* $0 < \bar{\alpha} < 1$ *such that for almost all* J

$$\lim_{N\to\infty} \frac{M_N}{N} = \lim_{N\to\infty} m_N = \bar{m} \qquad \lim_{N\to\infty} \frac{\sum_{i=1}^N \eta_i^*}{N} = \bar{\alpha}$$

Theorem (Lower bound for the Ground State Energy)

$$\bar{m} < .562...$$

UBQP: Theoretical results

Determine the statistical properties of $\min_\eta H(\eta)$

Conjecture

Let $\min_{\eta\in\{0,1\}^N} H(\eta) := H(\eta^*) := -M_N := -m_N N$
Then there exist $\bar{m} > 0$ *and* $0 < \bar{\alpha} < 1$ *such that for almost all* J

$$\lim_{N\to\infty} \frac{M_N}{N} = \lim_{N\to\infty} m_N = \bar{m} \qquad \lim_{N\to\infty} \frac{\sum_{i=1}^N \eta_i^*}{N} = \bar{\alpha}$$

Let $\Delta := M_N - E(M_N)$.

Theorem (Small Fluctuations of Minimum per Particle)

For some $C > 0$ *and for all* $z > 0$

$$P(|\Delta| > Nz) \leq e^{-CNz^2}$$

Lower bound for the Ground State Energy - Naive approach

Let $\nu(m, \alpha) = \sum_{\eta : |\eta| = N\alpha} 1_{(H(\eta) < -mN)}$

Lower bound for the Ground State Energy - Naive approach

Let $\nu(m,\alpha) = \sum_{\eta:|\eta|=N\alpha} 1_{(H(\eta)<-mN)}$ Then

$$E(\nu(m,\alpha)) = \binom{N}{\alpha N} \frac{1}{\sqrt{2\pi\alpha^2 N}} \int_{-\infty}^{-mN} e^{-\frac{x^2}{2\alpha^2 N}} dx$$

and

$$E(\nu(m)) = \sum_{\alpha N} \binom{N}{\alpha N} \frac{1}{\sqrt{2\pi\alpha^2 N}} \int_{-\infty}^{-mN} e^{-\frac{x^2}{2\alpha^2 N}} dx$$

Denoting by $I(\alpha) = -\alpha\log(\alpha) - (1-\alpha)\log(1-\alpha)$, we have

$$E(\nu(m)) \asymp \max_{\alpha\in[0,1]} e^{N\left(I(\alpha)-\frac{Nm^2}{2\alpha^2}\right)} := \max_{\alpha\in[0,1]} e^{NF(\alpha,m)}$$

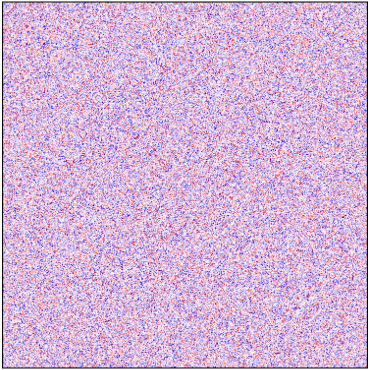
and $F(\alpha,m)=0$ for $\alpha \approx 0.788$ and $m \approx 0.801$

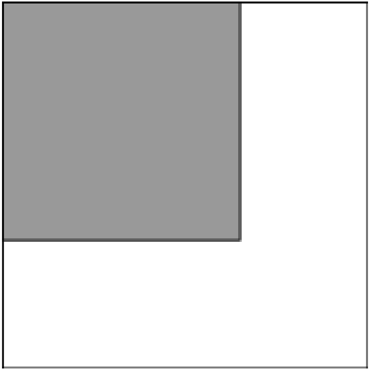
Lower bound for the Ground State Energy

Let's try to take into account the correlations between the sum of the J_{ij} selected by η and the ones not selected. We want to estimate $P\left(\nu(m,\alpha)=0\right)$.

Lower bound for the Ground State Energy

Let's try to take into account the correlations between the sum of the J_{ij} selected by η and the ones not selected. We want to estimate $P\left(\nu(m,\alpha)=0\right)$.





Lower bound for the Ground State Energy

Let's try to take into account the correlations between the sum of the J_{ij} selected by η and the ones not selected. We want to estimate $P\left(\nu(m,\alpha)=0\right)$.

Let $0<\gamma<\frac{1}{2}$. Then

$$\begin{aligned}
P\left(\nu(m,\alpha)=0\right) &\geq P\left(\nu(m,\alpha)=0, |H(1)|\leq N^{\frac{1}{2}+\gamma}\right)\\
&\geq 1-P(|H(1)|>N^{\frac{1}{2}+\gamma})-P\left(\bigcup_{|\eta|=\alpha N} H(\eta)<-mN, |H(1)|\leq N^{\frac{1}{2}+\gamma}\right)\\
&\geq 1-P(|H(1)|>N^{\frac{1}{2}+\gamma})-\sum_{|\eta|=\alpha N} P\left(H(\eta)<-mN, |H(1)|\leq N^{\frac{1}{2}+\gamma}\right)\\
&= 1-P(|H(1)|>N^{\frac{1}{2}+\gamma})-\binom{N}{\alpha N} P\left(H(\eta_\alpha)<-mN, |H(1)|\leq N^{\frac{1}{2}+\gamma}\right)
\end{aligned}$$

Lower bound for the Ground State Energy

We have

- $P(|H(1)| > N^{\frac{1}{2}+\gamma}) \asymp 0$
- $P\left(H(\eta) < -mN, |H(1)| \leq N^{\frac{1}{2}+\gamma}\right) \asymp e^{-N\frac{m^2}{2\alpha^2(1-\alpha^2)}}$

and hence

$$P\left(\nu(m,\alpha) = 0\right) \geq 1 - G(m,\alpha)$$

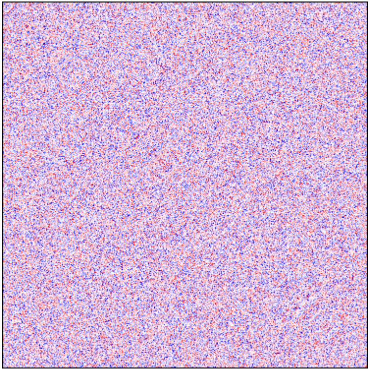
with

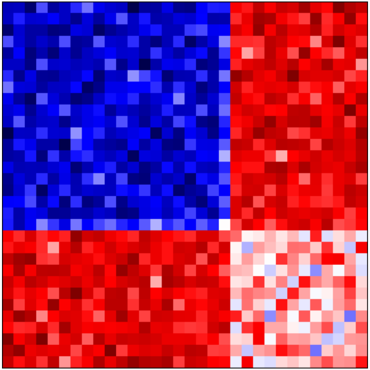
$$G(m,\alpha) \asymp +e^{N\left(I(\alpha)-\frac{m^2}{2\alpha^2(1-\alpha^2)}\right)}$$

Let $F_1(m,\alpha) := I(\alpha) - \frac{m^2}{2\alpha^2(1-\alpha^2)}$. Then,

- for $m > \bar{m} = .562$, $F_1(\alpha,\bar{m}) < 0 \ \forall \ \alpha$.
- for $m = \bar{m}$, $F_1(\bar{\alpha},\bar{m}) = 0$ for $\alpha = \bar{\alpha} = .644$
- $F_1(\alpha,\bar{m}) < 0 \ \forall \ \alpha \neq \bar{\alpha}$.

Lower bound for the Ground State Energy





Small fluctuations of Minimum per Particle

We want to show $P(|\Delta| > Nz) \leq e^{-CNz^2}$

Consider $P(\Delta > Nz)$. The evaluation of $P(\Delta < -Nz)$ is done in the same way.

By exponential Markov inequality we have that for all $t > 0$

$$P(\Delta > Nz) \leq e^{-tNz} E(e^{t\Delta})$$

Choose an arbitrary ordering on theJ_{ij}s so to have

$$M = \min_{\eta} \frac{1}{\sqrt{N}} \sum_{k=1}^{N^2} \eta_{i(k)} \eta_{j(k)} J_k$$

Small fluctuations of Minimum per Particle

Let $E_I(\cdot)$, with $I \subset \{1, 2, ..., N^2\}$, denote the expectation with respect to the J_k's with $k \in I$. Then

$$\Delta = M - E_{\{1\}}(M) + E_{\{1\}}(M) - E_{\{1,2\}}(M) + E_{\{1,2\}}(M) \cdots - E(M)$$

Call $\Delta_i = E_{\{1,2,\ldots,i-1\}}(M) - E_{\{1,2,\ldots,i\}}(M)$. and hence

$$E(e^{t\Delta}) = E\left(\prod_{i=1}^{N^2} e^{t\Delta_i}\right)$$

This expression can be estimated iteratively, showing that

$$E\left(\prod_{i=l}^{N^2} e^{t\Delta_i}\right) \leq E\left(\prod_{i=l+1}^{N^2} e^{t\Delta_i}\right) L(N)$$

with $L(N) \leq e^{\frac{ct^2}{N}}$.

For this purpose note that

$$E\left(\prod_{i=l}^{N^2} e^{t\Delta_i}\right) = E\left(\left(\prod_{i=l+1}^{N^2} e^{t\Delta_i}\right) e^{t\Delta_l}\right) =$$

$$= E_{\{l+1,\dots,N^2\}}\left(\left(\prod_{i=l+1}^{N^2} e^{t\Delta_i}\right) E_{\{l\}}(e^{t\Delta_l})\right)$$

and estimate

$$E_{\{l\}}(e^{t\Delta_l}) = 1 + \frac{t^2}{2} E_{\{l\}}(\Delta_l^2) + R_3(\Delta_l)$$

where

$$R_3(\Delta_l) = \frac{\tilde{t}^3}{3!} E_{\{l\}}\left(e^{\tilde{t}\Delta_l}\ \Delta_l^3\right) \qquad 0 \le \tilde{t} \le t$$

Small fluctuations of Minimum per Particle

Iterating on all indices,

$$E(e^{t\Delta}) \leq \left(e^{\frac{ct^2}{N}}\right)^{N^2} \leq e^{cNt^2}$$

and hence we get that for all $t > 0$

$$P(\Delta > Nz) \leq e^{-tNz} e^{cNt^2}.$$

Choosing $t = \frac{z}{2c}$ we get

$$P(\Delta > Nz) \leq e^{-N\frac{z^2}{4c}}$$

Work in progress

- Find better bounds for m
- Exploit "Shaken Dynamics" to find the minima of H
  - Introduced in the search of efficient dynamics to draw samples from Gibbs measure on spin system
  - Different set of neighbors considered at each step
  - Connects Ising models defined on different lattices

THANK YOU!