The sign problem and the spectral gap of quantum many-body systems

Elizabeth Crosson

University of New Mexico

Collaborators: Tameem Albash (UNM), Itay Hen (USC/ISI), Peter Young (UCSC)

arXiv: 2004.07681

The sign problem

At a high level, the sign problem refers to the general case of path integrals in quantum physics that involve interference between paths of different phases.

The special case of models without a sign problem are those for which we can define Euclidean path integrals over a nonnegative measure.

Surprisingly, many natural systems of interest do not have a sign problem. At a fundamental level, the sign problem arises from the existence of Fermions.

This talk primarily focuses on quantum spin systems that involve interactions between qubit degrees of freedom, where the sign problem can be characterized in terms of the Hamiltonian matrix.

These qubit systems can arise as effective interactions between molecules, or in engineered devices, or can be abstractly simulated with a quantum computer.

Stoquastic Hamiltonians

A Hamiltonian $H = \sum_i H_i$ is stoquastic if in some local basis $\mathcal{B}$ the terms H_i all have off-diagonal matrix elements that are zero or negative,

$$\langle x|H_i|y\rangle \leq 0 \quad \forall i, \forall x \neq y \in \mathcal{B}$$

This implies special properties for equilibrium states (ground state and Gibbs state) that put them on the border of quantum and classical complexity,

$$\rho_\beta = \frac{e^{-\beta H}}{Z} \quad , \quad Z = \mathrm{Tr}\left(e^{-\beta H}\right) \quad , \quad \lim_{\beta\to\infty} \rho_\beta = |\psi\rangle\langle\psi|$$

Stoquastic Hamiltonians

If H has all real and non-positive matrix entries in a particular basis, then

$$e^{-\beta H} = I + (-\beta H) + \frac{(-\beta H)^2}{2!} + \dots$$

is a sum of matrices with nonnegative entries. Therefore $\lim_{\beta\to\infty} \rho_\beta = |\psi\rangle\langle\psi|$ is also a matrix with nonnegative entries (for some choice of global phase).

The fact that the principal eigenvector of a matrix with nonnegative entries can be chosen to have nonnegative components is part of the Perron-Frobenius theorem.

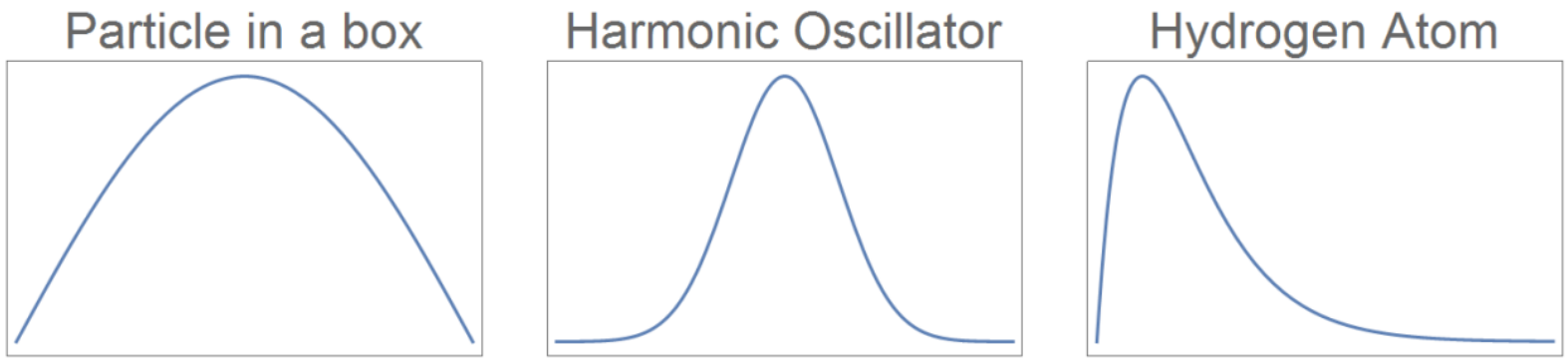
If H is irreducible then the components are strictly positive (no nodes). In the general (reducible) case the ground space of a stoquastic H is spanned by wave functions with nonnegative amplitudes.

Example: spinless particles

- H for a spinless particle in 1D is stoquastic in position basis,

$$H = -\partial_x^2 + V(x) \quad , \quad -\partial_x^2 \approx \begin{bmatrix} \ddots & & & & \\ -1 & 2 & -1 & & \\ & -1 & 2 & -1 & \\ & & & \ddots & \end{bmatrix}$$

- Every H we teach in first year QM is stoquastic!



- Any system of interacting bosons or distinguishable particles with $H = K + U$ is stoquastic.

Example: Experimental Qubit Devices

The effective qubit Hamiltonians produced in most (all?) current experimental quantum computing platforms are stoquastic! Trapped Ions, neutral Rydberg atoms, and flux qubits all implement quantum transverse Ising models (TIM),

$$H = \sum_{i,j} J_{ij} Z_i Z_j + \sum_i (b_i Z_i + \Gamma_i X_i)$$

Transmon qubits on bipartite graphs (e.g. square lattices) implement stoquastic Hamiltonians called XY models / Bose-Hubbard models,

$$H_{\text{sycamore}} = \sum_{\langle i,j \rangle} J_{ij}(X_i X_j + Y_i Y_j) + \frac{J_{ij}^2}{\eta} Z_i Z_j$$

The real-time evolution generated by these Hamiltonians is universal for quantum computing. Their semi-classical features only apply to equilibrium states.

Stoquastic Path Integrals

We can expand the partition function as a "path integral",

$$Z = \mathrm{tr}\left(e^{-\beta H}\right) = \sum_{x \in \mathcal{B}} \langle x | \left(e^{-\frac{\beta H}{L}}\right)^L | x \rangle = \sum_{x_1, \ldots, x_L \in \mathcal{B}} \prod_{i=1}^{L} \langle x_i | e^{-\frac{\beta H}{L}} | x_{i+1} \rangle$$

Since the "propagator" $e^{-\beta H/L}$ is a nonnegative matrix, every path $(x_1, \ldots, x_L)$ makes a positive contribution to the sum. Define a probability distribution over paths,

$$\pi(x_1, \ldots, x_L) = \frac{1}{Z} \prod_{i=1}^{L} \langle x_i | e^{-\frac{\beta H}{L}} | x_{i+1} \rangle$$

Evidently, some paths are more important than others. Path Integral Monte Carlo is a classical simulation algorithm that does important sampling on these paths.

PIMC and Suzuki's Quantum-to-Classical Mapping

PIMC aims to sample from a distribution over nonnegative path amplitudes that contribute to the partition function:

$$\pi(z_1, ..., z_L) = \frac{1}{Z} \prod_{i=1}^{L} \langle z_{i-1} | e^{-\beta H/L} | z_i \rangle \qquad z_i \in \{-1, 1\}^n$$

To bring each amplitude into an efficiently computable form we Trotterize. The general principle is to split H up into commuting layers. For TIM, $H = H_Z + H_X$,

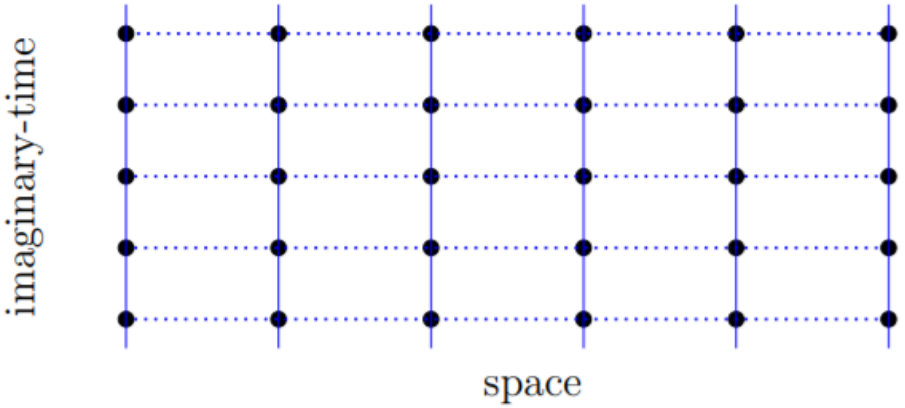
$$\prod_{i=1}^{L} \langle z_{i-1} | e^{-\beta H/L} | z_i \rangle \approx \exp\left(-\frac{\beta}{L} \sum_{i=1}^{L} H_Z(z_i) \right) \prod_{i=1}^{L} \langle z_{i-1} | e^{-\beta H_X/L} | z_i \rangle$$

The exponential out front looks like L copies of a classical system that contains the Z interactions, while the product couples this copies together.

PIMC and Suzuki's Quantum-to-Classical Mapping

This product expression only involves local couplings between the classical spins:

$$\prod_{i=1}^{L} \langle z_{i-1} | e^{-\epsilon \sum_i X_i} | z_i \rangle = \prod_{i,j=1}^{L,n} \langle z_{i-1,j} | e^{-\epsilon X_j} | z_{i,j} \rangle$$



At the end of all this, we can interpret the distribution over paths as a classical model spin model, with interactions that depend on the quantum temperature,

$$\pi(z_1, ..., z_L) = \frac{1}{Z} \exp\left(-\frac{\beta}{L} \sum_{i=1}^{L} H_Z(z_i) \right) \prod_{i,j}^{L,n} \langle z_{i-1,j} | e^{-\beta \Gamma_j X_j / L} | z_{i,j} \rangle$$

Suzuki developed this quantum-to-classical mapping to do MCMC on quantum spin systems, and to prove dualities between quantum and classical phase transitions.

Adiabatic Optimization

Adiabatic optimization is a heuristic quantum algorithm that aims to minimize $f : \{0,1\}^n \to \mathbb{R}$ by remaining close to the instantaneous ground state of a parameterized Hamiltonian $H(s), s \in [0,1]$, such that the ground state of H(0) is easy to initialize, and $H(1) = f(Z_1, ..., Z_n)$.

If $\Delta = \min_{s\in[0,1]} E_s^1 - E_s^0$ is the smallest spectral gap to the first excited state along the path then the algorithm will find the minimum of *f* in time $\mathrm{poly}(n, \Delta^{-1})$.

While the endpoint *H(s=1)* is a fixed classical Hamiltonian, we may consider various Hamiltonian paths *H(s)* that reach this endpoint, with the goal of making Δ as large as possible (subject to hardware constraints).

The original proposal was transverse-field adiabatic optimization: $H(s) = (1-s)H_x + sH_f \ , \ H_X = -\sum_{i=1}^{n} X_i$

However, these transverse Ising models are stoquastic Hamiltonians. The stoquastic local Hamiltonian problem is the 3rd level of the PH, and classical simulations based on Monte Carlo are effective in practice. These limitations have motivated the study of adiabatic optimization with nonstoquastic Hamiltonians.

Adiabatic Optimization

As it became increasingly unlikely 2000-2015 to find a quantum speedup with stoquastic adiabatic optimization, attention shifted to the potential use of nonstoquastic (general) Hamiltonians.

The main arguments for this were:

1. Adiabatic computation with nonstoquastic H can be universal for QC (but note the final Hamiltonian is more complicated in these constructions than a classical optimization Hamiltonian).

2. Physicists have long known that the practically efficient Monte Carlo simulations of transverse Ising ground states break down in the nonstoquastic case due to the sign problem.

Today's message: nonstoquastic Hamiltonians may make adiabatic optimization more difficult to classically simulate, but they also almost always decrease performance by shrinking the spectral gap. Therefore **randomly chosen nonstoquastic adiabatic paths are not a promising direction for a quantum speedup.**

De-signed Hamiltonians

We introduce new ways to systematically compare stoquastic and nonstoquastic Hamiltonians in adiabatic optimization:

$$H = \sum_{i,j \in \{0,1\}^n} H_{ij} |i\rangle\langle j|$$

Original H may be nonstoquastic.

$$\bar{H} = \sum_{i \in \{0,1\}^n} H_{ii} |i\rangle\langle i| - \sum_{\substack{i,j \in \{0,1\}^n \\ i \neq j}} |H_{ij}| |i\rangle\langle j|$$

De-signed and shifted Hamiltonians are always stoquastic.

Mathematically, we simply wish to compare the spectral gaps at the top of the spectrum for a Hermitian matrix G and its entry-wise absolute value $\bar{G}$.

We observe that de-signing almost always increases the spectral gap (for any reasonable ensemble). Any general theorem must be probabilistic because there are exceptions, but the fraction of exceptional cases (where the de-signed matrix has a smaller spectral gap) goes to zero with the dimension of the matrix.

A map from nonstoquastic to stoquastic H

De-signing is a nonlinear map that changes all the eigenvalues and eigenvectors. But it leaves classical Hamiltonians unchanged, so it preserves the endpoints of the adiabatic path. Therefore the de-signed H(s) should be regarded as a new path through the space of Hamiltonians connecting those endpoints.

Our message about the typical behavior of the spectral gap is supported by:

(1) a rigorous proof in the setting of random matrix ensembles.

(2) special cases: de-signing $H = f(X_1, ..., X_n)$ in the computational basis always increases the gap.

(2) results based on the spectral theory of signed graph Laplacians that help explain the phenomenon

(3) A clear trend in numerical simulations that already emerges at low qubit number.

All of our analytical arguments support the main conclusion, but random local Hamiltonians are difficult to analyze. We see a clear numerical trend in that case but can't prove the apparent result.

De-signing and the ground state energy

Initial observation: de-signing always decreases the ground energy

$$\bar{E}_0 \leq E_0$$

Proof: if $|\psi\rangle = \sum_x \psi_x |x\rangle$ is the ground state of H, then $|\bar{\psi}\rangle = \sum_x |\psi_x||x\rangle$ obeys

$$\bar{E}_0 \leq \langle\bar{\psi}|\bar{H}|\bar{\psi}\rangle \leq E_0$$

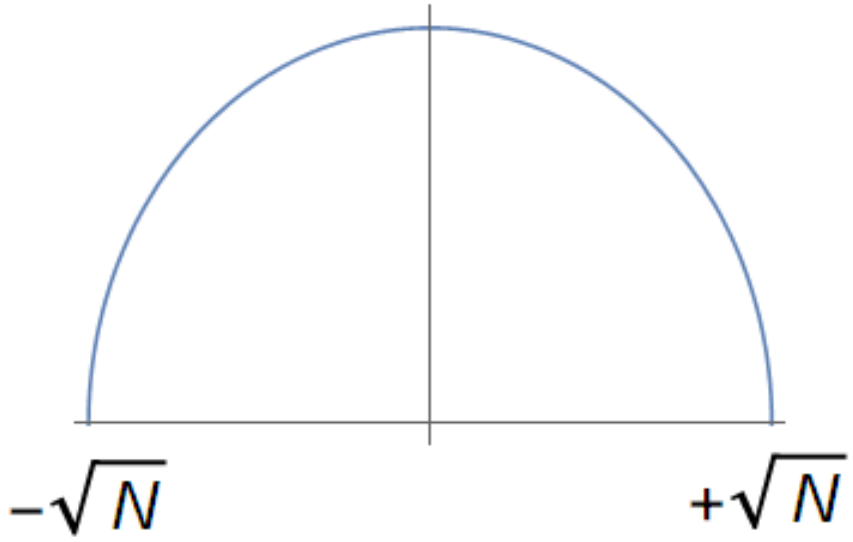
But notice that if $H \succeq 0$ then $\mathrm{tr}(H) = \mathrm{tr}(\bar{H})$, so de-signing always increases the spacing between eigenvalues s*omewhere* in the spectrum.

Notice that if $\bar{E}_1 \geq E_1$ we would be done, but this isn't true in general.

Dense Random Matrices

Random Matrices: consider an ensemble of $N \times N$ random symmetric matrices with i.i.d. entries of mean 0 and variance 1. If the entries are Gaussian distributed this is called the Gaussian Orthogonal Ensemble.

Wigner showed that the distribution of eigenvalues of these matrices is a semicircle between $\pm\sqrt{N}$. This also holds for entries chosen uniformly in [-1,1].

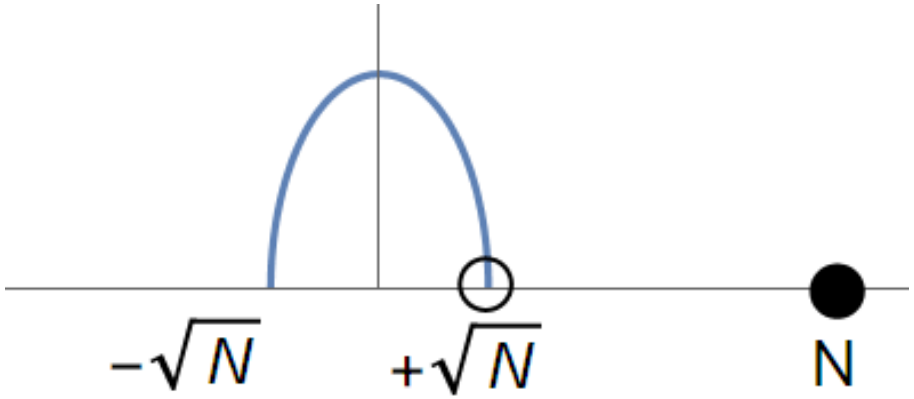


Since there are N eigenvalues, the typical spacing between them is $1/\sqrt{N}$, and large deviation bounds for this spacing are also known.

Dense Random Matrices

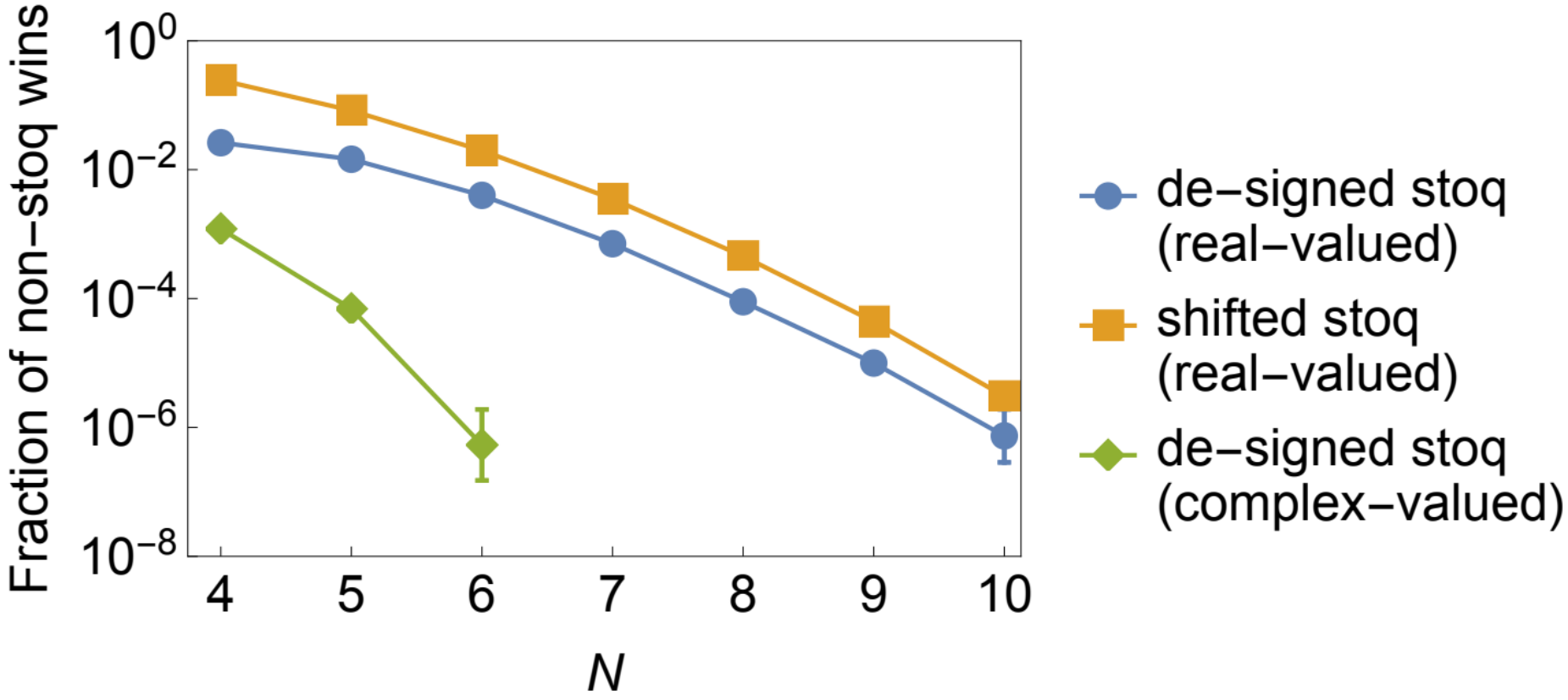
To compare this with stoquastic Hamiltonians, define an ensemble of symmetric matrices with nonnegative iid entries chosen uniformly from [0,1].

Each matrix in this ensemble can be obtained from a matrix in the previous ensemble by adding a matrix with a 1 in each entry. This rank 1 shift turns the largest eigenvalue into an outlier.



Therefore, in this normalization this increases the spectral gap dramatically, from $\mathcal{O}(N^{-1/2})$ to $N - \mathcal{O}(N^{1/2})$!

Dense Random Matrices



N is the dimension of the matrix. 10^7 random matrices for each data point. Error bars equal to twice standard error.

Spectral Graph Theory

For a combinatorial graph G = (V,E) the adjacency matrix is a |V| x |V| matrix with

$$A_{uv} = 1 \quad , \quad \text{if } (u,v) \in E \quad , \quad A_{uv} = 0 \quad , \quad \text{o.w.}$$

If D is a diagonal matrix with the degree of each vertex along the diagonal, then the random walk on the graph is described by a probability-conserving transition matrix called the random walk Laplacian:

$$L_{RW} = I - D^{-1}A$$

This RW Laplacian can be mapped by a similarity transform to a symmetric matrix called the normalized Laplacian:

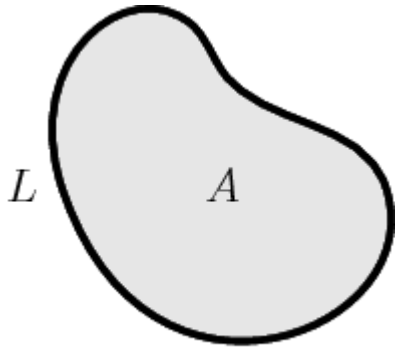
$$L = D^{1/2} L_{RW} D^{-1/2} = I - D^{-1/2} A D^{-1/2}$$

If G is connected, the ground state of L is uniform over the vertices. Generalizing these definitions to include weighted edges and vertices means that every stoquastic Hamiltonian can be mapped to a normalized Laplacian by a combination of an affine transformation and a similarity transformation, knowledge of the ground state.

SGT relates the spectral properties of L to the geometry of the graph, and in the same way we can learn about the spectrum of a Hamiltonian from its ground state geometry and the connectivity of its matrix elements. For nonstoquastic Hamiltonians we have also adapted results for graphs with *signed* adjacency matrices.

Isoperimetric Inequalities

$$\frac{\text{measure on the boundary of a set}}{\text{measure inside a set}}$$

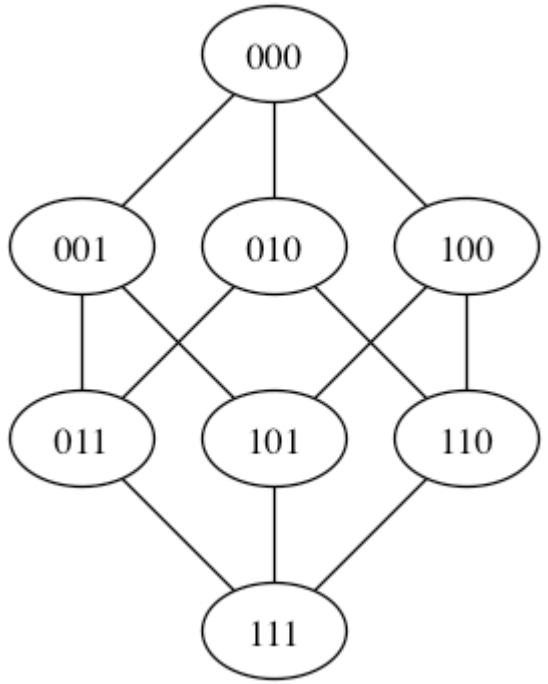


$$L^2 \geqslant 4\pi A$$

Discrete space: “vertex expansion”

$$S = \{000, 100, 010, 001\} \implies \partial S = \{010, 010, 001\}$$

$$\implies \frac{|\partial S|}{|S|} = \frac{3}{4}$$



Ground States and Weighted Graphs

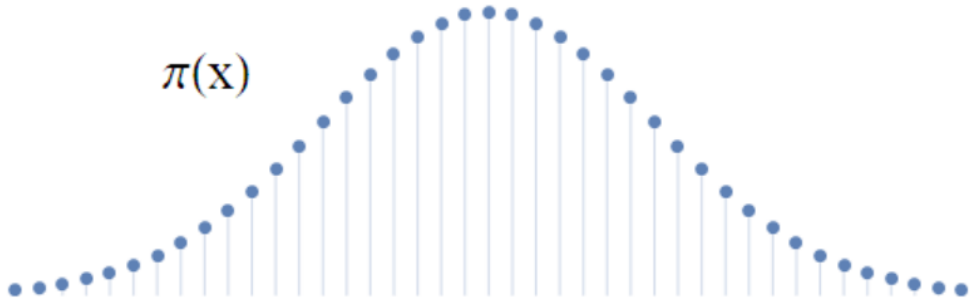
Think of the elements of a basis *B* as vertices of a graph, and the ground state as a probability distribution on this graph.

$$\pi(x) = |\psi(x)|^2 \ , \ x \in \mathcal{B}$$

Connect *x,y* by an edge if $\langle x|H|y\rangle \neq 0$, and define

$$\partial S = \{x \in S : y \notin S \text{ with } \langle x|H|y\rangle \neq 0\}$$

Basis *B* could be (discrete) position,
or could be the computational basis.



Ground State Isoperimetric Inequality

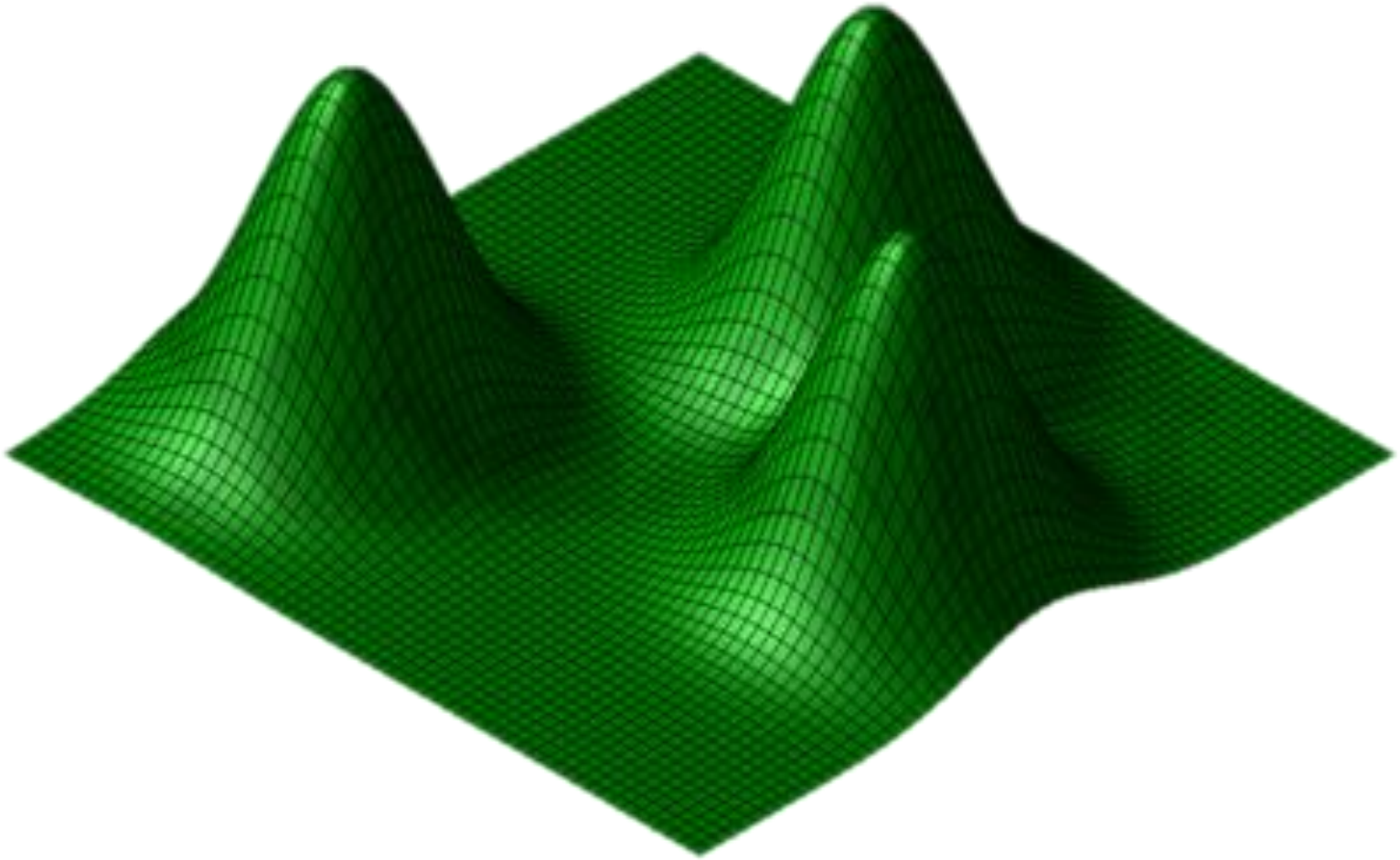
Define: ground energy E, spectral gap Δ, largest energy $\|H\|$

Theorem: any H, any basis B, any $S \subset \mathcal{B}$ with $0 < \pi(S) \leq 1/2$,

$$\Delta \leq 2\left(\|H\| - E\right) \frac{\pi(\partial S)}{\pi(S)}$$

Depends on the ground state, but not the details of H. For a given π it constrains the spectral gap of any H with that ground state!

Implications: QAO inevitably encounters small spectral gaps when preparing complicated multi-modal distributions.



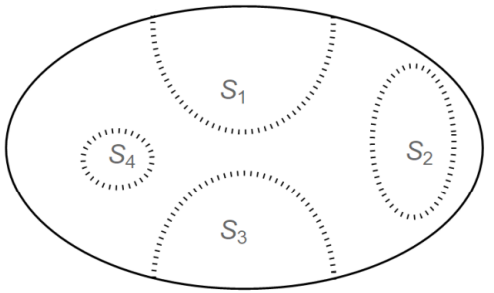
Spectral Graph Theory

The idea of relating the spectrum to geometric bottlenecks with the isoperimetric inequality can be greatly generalized. Higher-order Cheeger inequalities partition the graph into k disjoint subsets and that minimize the maximum expansion of the k sets, yielding upper and lower bounds on λ_k .

Define the edge expansion of a subset S of vertices, $\Phi(S) = \frac{|\mathcal{E}(S, \bar{S})|}{\mathrm{vol}(S)}$

Consider all partitions of the graph into k-subsets, minimizing the largest expansion of any subset:

$$h_k^+ = \min_{\substack{S_1,\ldots,S_k \subseteq \mathcal{V} \\ S_i \cap S_j = \emptyset}} \max\{\Phi(S_1), \ldots, \Phi(S_k)\}$$



This k-th order Cheeger constant of the graph characterizes the k-th eigenvalue of the graph Laplacian:

$$\frac{\lambda_k^+}{2} \leq h_k^+ \leq Ck^3\sqrt{2D_{\max}\lambda_k^+}$$

Spectral Graph Theory

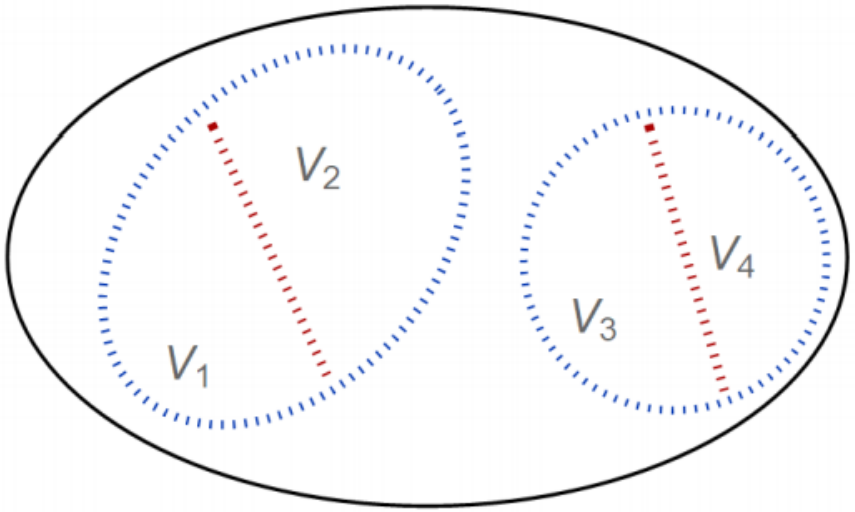
If signs +/- 1 are placed on the edges of a graph, then it is called a signed graph. The resulting signed graph Laplacian is now a nonstoquastic Hamiltonian (in general), and its low-energy spectrum can be related to the geometry of the signed graph by the signed Cheeger inequalities.

In addition to the edge expansion, a new ingredient is the "frustration index" caused by an inability to divide the vertices into two sets, with only positive edges internal to each set, and negative edges between them.

$$F(S) := \frac{1}{\mathrm{vol}(S)} \cdot \min_{\substack{S_1 \cap S_2 = \emptyset \\ S_1 \cup S_2 = S}} \left(2|\mathcal{E}^+(S_1, S_2)| + |\mathcal{E}^-(S_1)| + |\mathcal{E}^-(S_2)|\right)$$

$$h_k^\sigma = \min_{\substack{S_1, \ldots, S_k \subseteq \mathcal{V} \\ S_i \cap S_j = \emptyset}} \max\left\{F(S_1) + \Phi(S_1), \ldots, F(S_k) + \Phi(S_k)\right\}$$

$$\frac{\lambda_k^\sigma}{2} \leq h_k^\sigma \leq Ck^3\sqrt{2D_{\max}\lambda_k^\sigma}.$$

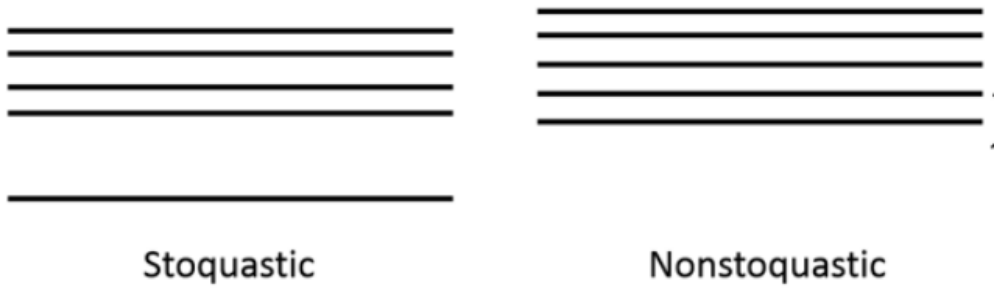


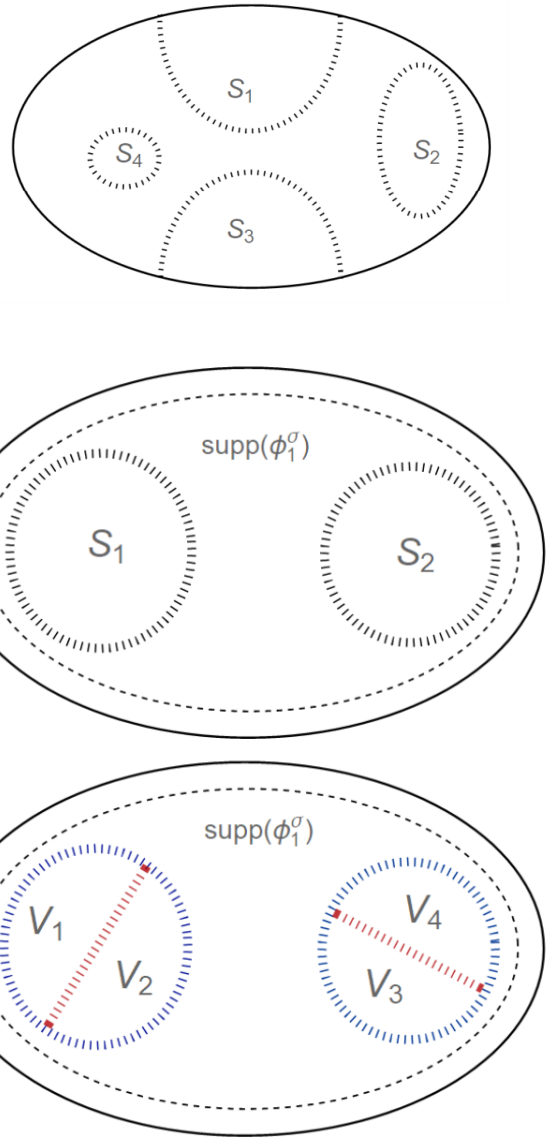
$$S_1 = V_1 \cup V_2 \quad , \quad S_2 = V_3 \cup V_4$$

Spectral Graph Theory

Cheeger inequalities for stoquastic graphs tells us that the energy of excited states comes from partitioning the vertices into modal regions.

Cheeger inequalities for nonstoquastic graphs generalize this picture, adding frustration from the phases on the edges.





Using the stoquastic ground and first excited states, and the nonstoquastic ground state, we can construct a variational nonstoquastic 1st excited state to upper bound the gap.

MaxCut Numerical Experiments

Our numerical experiments with local Hamiltonians are for MaxCut problems on 3-regular graphs with n = 6,…,20 vertices.

$$H = (1-s)H_D + s(1-s)H_C + sH_I$$

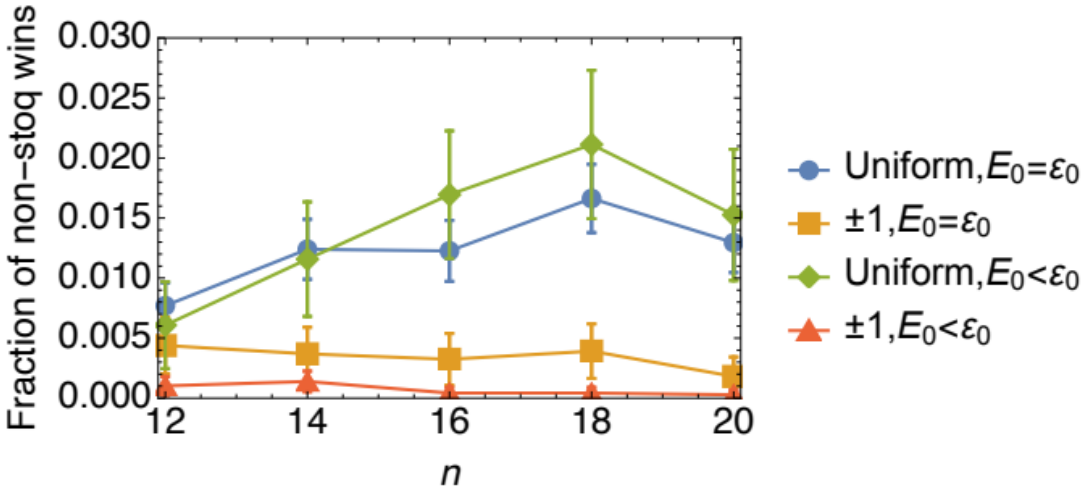
$$H_D = -\sum_i X_i \quad , \quad H_C = \sum_{i\sim j} \alpha_{ij} X_i X_j \quad , \quad H_I = \sum_{i\sim j} Z_i Z_j$$

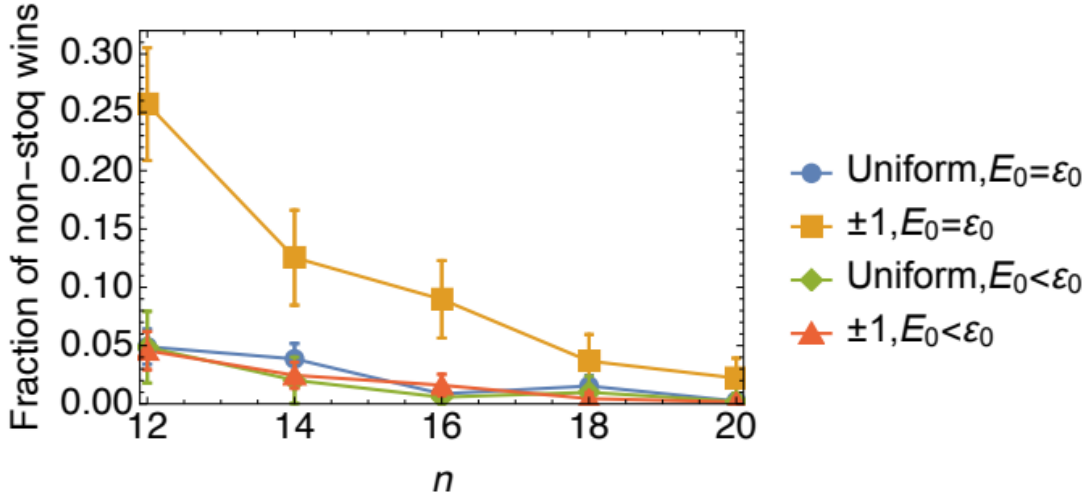
We distinguish the case where the α_{ij} coefficients are randomly +/- 1, and the case where they are uniformly random in [-1,1].

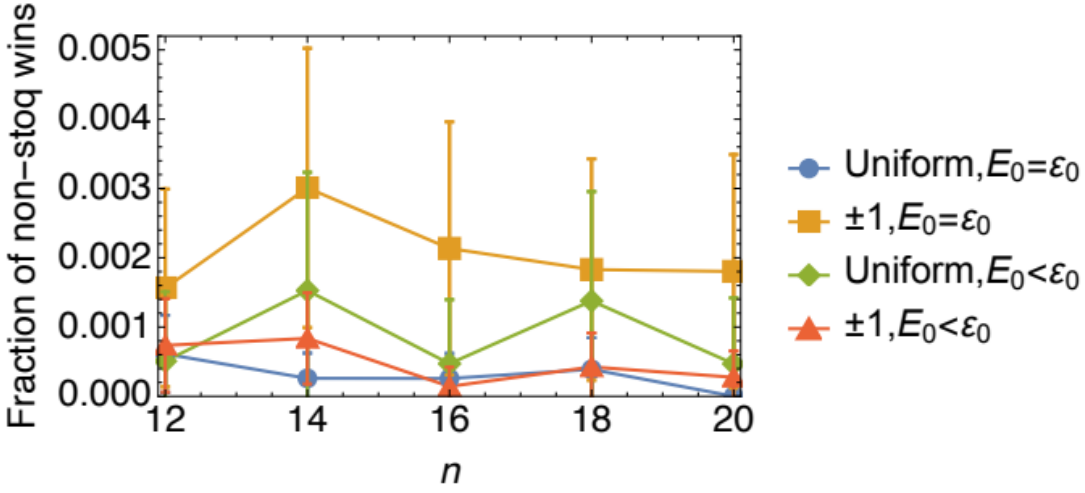
The Hamiltonian has a global bit flip symmetry, $P = \prod_i X_i$, and the time evolution is constrained to this subspace. We hypothesize that nonstoquastic paths are more likely to out perform stoquastic ones when the true ground state energy at the point of the minimum gap is outside this +1 subspace, written

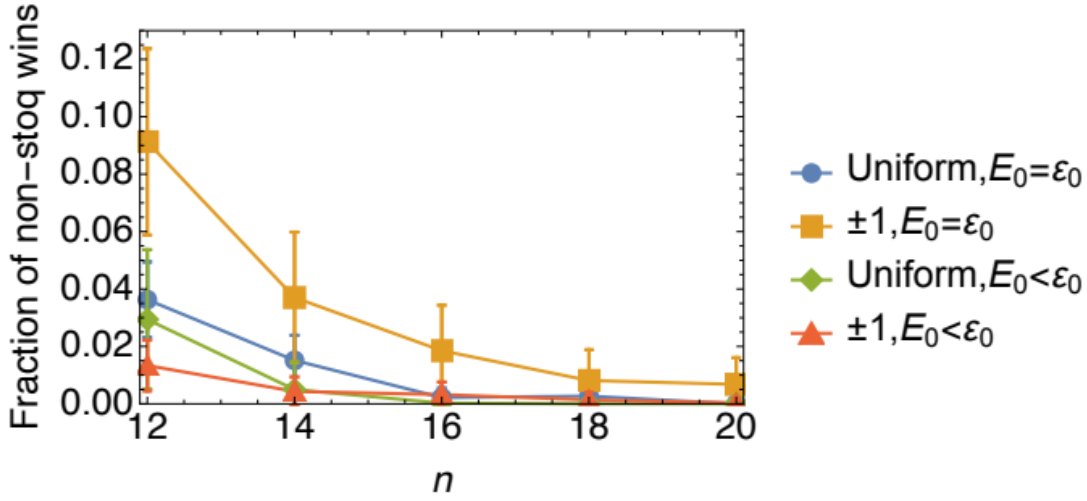
$$E_0(s*) < \epsilon_0(s*)$$

MaxCut Numerical Experiments









Left: comparing the spectral gap, Right: comparing TTS. Top: de-signed, Bottom: shifted.
Error bars represent 2 sigma over 100-1000 instances. Fraction of nonstoq wins is always small.

Conclusion and Open Questions

Nonstoquastic Hamiltonians were one of the last great hopes for the original proposal of adiabatic optimization. This can be a cautionary tale about equivocating hardness of classical simulation and useful algorithmic enhancement.

One of the main general problems is that Hamiltonians with disordered interactions representing hard optimization problems run into “many-body localized” phases with exponentially small spectral gaps.

If we’re forced to leave the ground state, we can still consider “diabatic quantum annealing”, which is recently found to be related to QAOA. Lukin et al, 1812.01041, “our results indicate that QAOA is closely related to a cleverly optimized diabatic QA path that can overcome limitations set by the adiabatic theorem.”

The main question we sought to answer is still open: prove that for any reasonable ensemble of Hermitian matrices (e.g. local Hamiltonians, lattice models, etc) that the entrywise absolute value of the matrix increases the gap at the top of the spectrum.

The same question can be asked for comparing Bosonic and Fermionic systems with the same kind of interactions.