Quantum Boltzmann Machines: a bridge between semidefinite optimisation and quantum thermodynamics

Nana Liu

Shanghai Jiao Tong University

10th December 2025

International Workshop on Quantum Boltzmann Machines

Mark Wilde

Michele Minervini

Dhrumil Patel

Quantum thermodynamics and semi-definite optimization

arXiv: 2505.04514

Boltzmann Machine

$$\frac{e^{-\beta \bar{H}(\mu)}}{Z_T(\mu)}$$

$\bar{H}(\mu)$ is linear in μ

μ: tuneable parameters

$Z_T(\mu)$: partition function

Classical thermodynamics: Grand canonical ensemble

$$\frac{e^{-\beta(H-\mu\cdot Q)}}{Z_T(\mu)} \qquad [H, Q_i] = 0 \qquad [Q_i, Q_j] = 0$$

where Q_i are conserved quantities (e.g. particle number)
μ: chemical potential

Thermal state is a minimiser of free energy

$$\frac{e^{-\beta(H-\mu\cdot Q)}}{Z_T(\mu)}$$

Quantum thermodynamics

$$\frac{e^{-\beta(H-\mu\cdot Q)}}{Z_T(\mu)} \qquad [Q_i, Q_j] \neq 0$$

Non-Abelian thermal states of thermodynamic systems with conserved, non-commuting charges

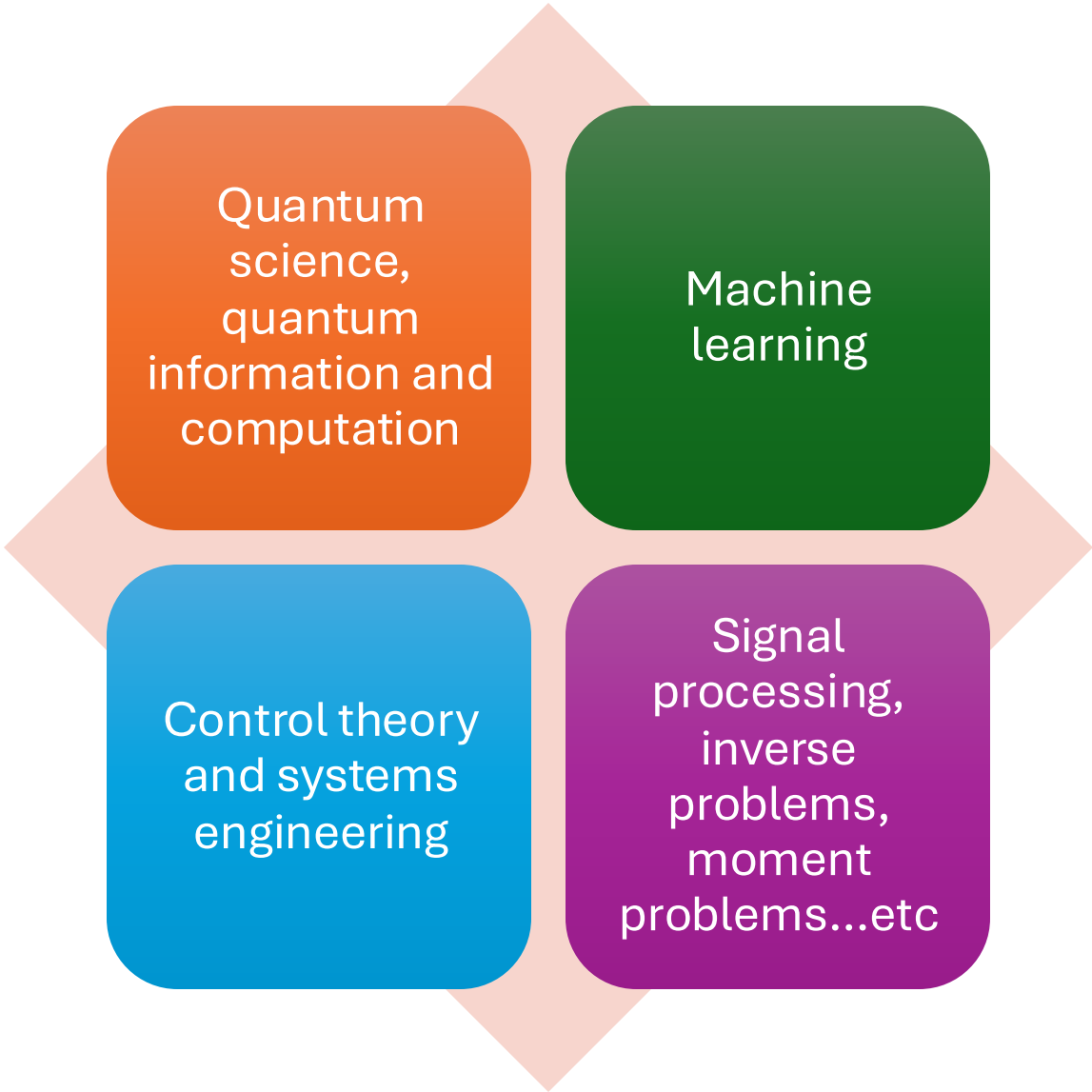
Quantum thermodynamics

$$\frac{e^{-\beta(H-\mu\cdot Q)}}{Z_T(\mu)}$$

$$[Q_i, Q_j] \neq 0$$

...also a minimiser of free energy: a natural ansatz to solving an optimisation problem!

Semidefinite optimisation: broad applications



Semidefinite optimisation

Given $d \times d$ Hermitian matrices C, $\{A_i\}$

Solve

$$\alpha := \min_{X \geq 0} \{\mathrm{Tr}[CX] : \mathrm{Tr}[A_i X] = b_i \ \forall i \in [c]\}$$

$$b_i \in \mathbb{R}$$

Modified semidefinite optimisation

Solve

SDP

$$\alpha_R = \min_{X \geq 0} \ \mathrm{Tr}[CX]$$
$$\text{s.t.} \ \ \mathrm{Tr}[A_i X] = b_i, \forall i$$
$$\mathrm{Tr}[X] \leq R$$

This is equivalent to solving for the density matrix ρ

$$\alpha_R = R \min_{\rho \in \mathcal{D}_{d+1}} \{\mathrm{Tr}(H\rho) : \quad \mathrm{Tr}(Q_i \rho) = q_i, \quad \forall i \in [c]\}$$

$$H = C \oplus [0], \quad Q_i = A_i \oplus [0], \quad q_i = \frac{b_i}{R}$$

Quantum thermodynamic system

Given a tuple $(H, Q_1, \cdots, Q_c)$ of $d \times d$ Hermitian matrices

Describes a quantum thermodynamical system

- Hamiltonian H
- Conserved non-commuting charges Q_i, so $[Q_i, Q_j] \neq 0$
- e.g. spin, angular momentum

Energy minimisation

Solve

Energy Minimization

$$E(\mathcal{Q}, q) = \min_{\rho \in \mathcal{D}_d} \langle H \rangle_\rho$$

$$\text{s.t. } \langle Q_i \rangle_\rho = q_i, \forall i$$

Energy minimisation problem in quantum thermodynamics equivalent to semidefinite optimisation

Energy minimisation and free energy minimisation

Real physical systems operate in *finite temperature*, so free energy minimisation is much more natural

$$F_T(\mathcal{Q}, q) \coloneqq \min_{\rho \in \mathcal{D}_d} \left\{ \langle H \rangle_\rho - TS(\rho) : \langle Q_i \rangle_\rho = q_i \ \forall i \in [c] \right\}$$

where $\langle H \rangle_\rho - TS(\rho)$ is the free energy and the von Neumann entropy is defined as

$$S(\rho) \coloneqq -\operatorname{Tr}[\rho \ln \rho].$$

Energy minimisation and free energy minimisation

Free energy very closely approximates energy at low temperature

$$0 \leq S(\rho) \leq \ln d \implies E(\mathcal{Q}, q) \geq F_T(\mathcal{Q}, q) \geq E(\mathcal{Q}, q) - T \ln d.$$

Zero temperature limit: $\lim_{T \to 0} F_T(\mathcal{Q}, q) = E(\mathcal{Q}, q)$

To have error ϵ, use $T \sim \frac{\epsilon}{\ln d}$, $d = 2^n$, $n =$ number of qubits

Free energy minimisation subject to constraints same as maximisation of function of chemical potential

$$F_T(\mathcal{Q}, q) = \min_{\rho \in \mathcal{D}_d} \{\mathrm{Tr}[H\rho] - TS(\rho) : \mathrm{Tr}[Q_i\rho] = q_i \ \forall i \in [c]\}$$

(Lagrange multipliers)

$$= \min_{\rho \in \mathcal{D}_d} \left\{ \mathrm{Tr}[H\rho] - TS(\rho) + \sup_{\mu \in \mathbb{R}^c} \sum_{i=1}^{c} \mu_i \left(q_i - \mathrm{Tr}[Q_i\rho]\right) \right\}$$

(Sup outside)

$$= \min_{\rho \in \mathcal{D}_d} \sup_{\mu \in \mathbb{R}^c} \left\{ \mu \cdot q + \mathrm{Tr}[H\rho] - TS(\rho) - \sum_{i=1}^{c} \mu_i \,\mathrm{Tr}[Q_i\rho] \right\}$$

(Minimax)

$$= \sup_{\mu \in \mathbb{R}^c} \min_{\rho \in \mathcal{D}_d} \left\{ \mu \cdot q + \mathrm{Tr}[H\rho] - TS(\rho) - \sum_{i=1}^{c} \mu_i \,\mathrm{Tr}[Q_i\rho] \right\}$$

Free energy minimisation subject to constraints same as maximisation of function of chemical potential

Some algebra:

$$\min_{\rho\in\mathcal{D}_d}\left\{\mu\cdot q+\mathrm{Tr}[H\rho]-TS(\rho)-\sum_{i=1}^{c}\mu_i\,\mathrm{Tr}[Q_i\rho]\right\}$$

$$=T\min_{\rho\in\mathcal{D}_d}\left\{\frac{\mu\cdot q}{T}-S(\rho)-\mathrm{Tr}[\rho\ln\rho_T(\mu)]-\ln Z_T(\mu)\right\}$$

$$=T\min_{\rho\in\mathcal{D}_d}\left\{\frac{\mu\cdot q}{T}+D(\rho\|\rho_T(\mu))-\ln Z_T(\mu)\right\}$$

$$=\mu\cdot q-T\ln Z_T(\mu).$$

$\rho=\rho_T$ is the minimiser!

Recall:

$$Z_T(\mu):=\mathrm{Tr}\left[\exp\left(-\frac{1}{T}\left(H-\sum_{i=1}^{c}\mu_iQ_i\right)\right)\right],$$

$$\rho_T(\mu):=\frac{1}{Z_T(\mu)}\exp\left(-\frac{1}{T}\left(H-\sum_{i=1}^{c}\mu_iQ_i\right)\right).$$

Quantum (Umegaki) relative entropy:

$$D(\omega\|\tau):=\mathrm{Tr}[\omega\,(\ln\omega-\ln\tau)]$$

$$D(\omega\|\tau)\geq 0$$

$$D(\omega\|\tau)=0\iff\omega=\tau$$

Free energy minimisation subject to constraints same as maximisation of function of chemical potential

Since

$$F_T(\mathcal{Q}, q) \coloneqq \min_{\rho \in \mathcal{D}_d} \left\{ \langle H \rangle_\rho - TS(\rho) : \langle Q_i \rangle_\rho = q_i \ \forall i \in [c] \right\}$$

$$= \sup_{\mu \in \mathbb{R}^c} \{ \mu \cdot q - T \ln Z_T(\mu) \}$$

therefore our task is to prepare the state

$$\rho_T(\mu) \coloneqq \frac{1}{Z_T(\mu)} \exp\left(-\frac{1}{T} \left(H - \sum_{i=1}^{c} \mu_i Q_i \right) \right)$$

where μ is determined by maximising the function $f(\mu) = \mu \cdot q - T \ln Z_T(\mu)$

Optimising value of chemical potential

$$F_T(\mathcal{Q}, q) = \sup_{\mu \in \mathbb{R}^c} \{\mu \cdot q - T \ln Z_T(\mu)\}$$

Objective function: $f(\mu) := \mu \cdot q - T \ln Z_T(\mu)$

Gradient: $\frac{\partial}{\partial \mu_i} (\mu \cdot q - T \ln Z_T(\mu)) = q_i - \langle Q_i \rangle_{\rho_T(\mu)}$

Optimising value of chemical potential

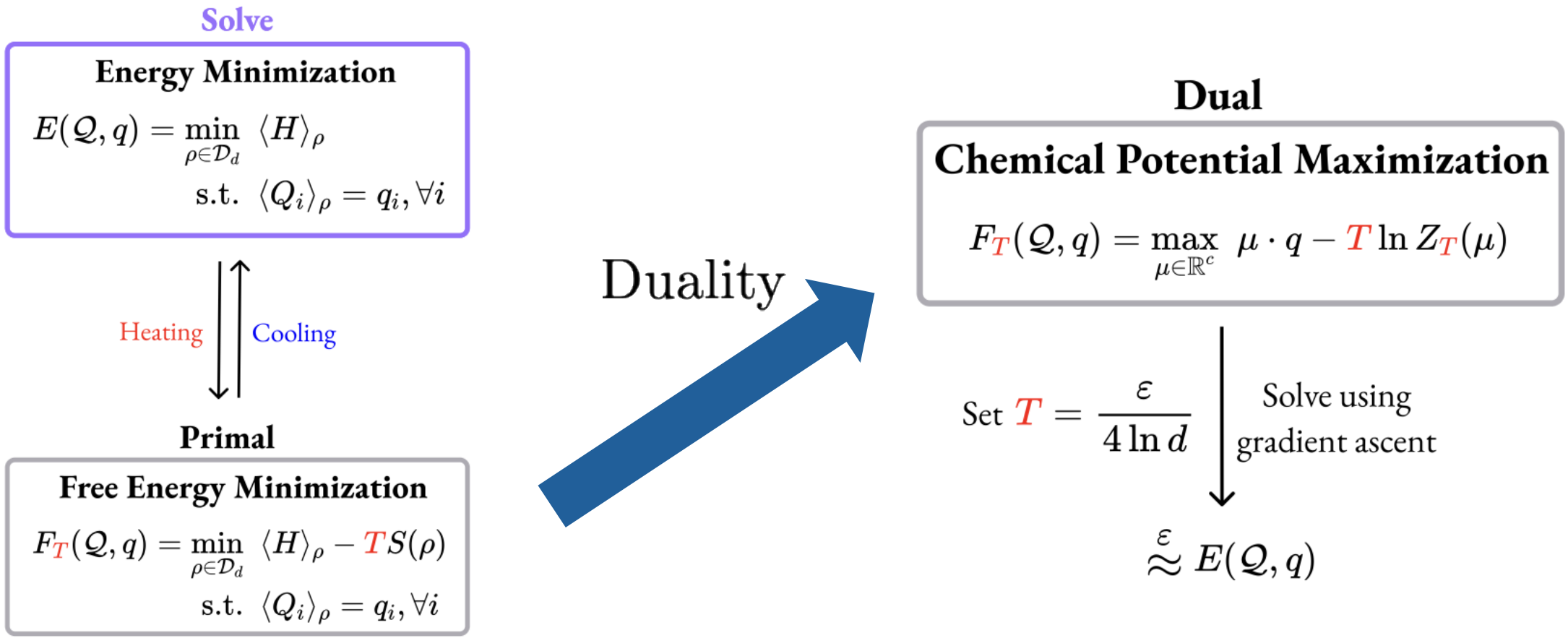
$$\frac{\partial^2}{\partial \mu_i \partial \mu_j} f(\mu) = -I_{ij}^{\mathrm{KM}}(\mu),$$

$$I_{ij}^{\mathrm{KM}}(\mu) = \frac{1}{T}\int_0^1 ds\ \mathrm{Tr}[\rho_T(\mu)^s Q_i \rho_T(\mu)^{1-s} Q_j]\ - \frac{1}{T}\langle Q_i \rangle_{\rho_T(\mu)} \langle Q_j \rangle_{\rho_T(\mu)}$$

$$\text{Kubo-Mori information matrix } I_{KM}(\mu) \geq 0 \implies$$

- Hessian of $f(\mu)$ is negative semi-definite and $f(\mu)$ is concave in μ
- maximisation of $f(\mu)$ can be solved (globally) via gradient ascent w.r.t μ

Overview of basic idea



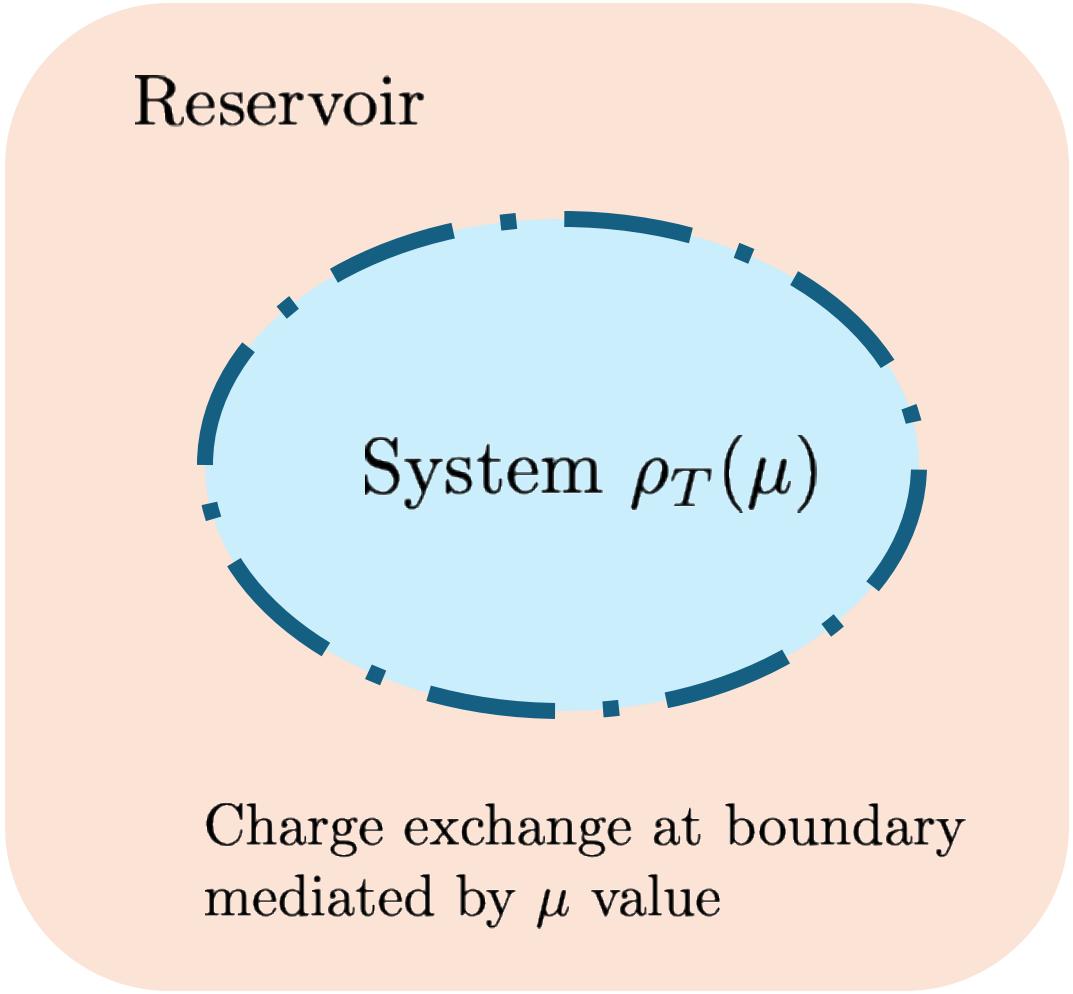
Gradient ascent algorithm (classical algorithm)

Algorithm 1 `minimize_energy`$(H, (Q_i)_i, (q_i)_i, L, d, \varepsilon, r)$

1: **Input:**

- Observables H and $(Q_i)_{i\in[c]}$
- Constraint values $(q_i)_{i\in[c]}$
- Smoothness parameter L
- Hilbert space dimension d
- Desired error $\varepsilon > 0$
- Radius r: An upper bound on $\|\mu^*\|$, where μ^* is the optimal solution to (14) for $T = \frac{\varepsilon}{4 \ln d}$

2: Set $T \leftarrow \frac{\varepsilon}{4 \ln d}$
3: Initialize $\mu^0 \leftarrow (0, \ldots, 0)$
4: Set learning rate $\eta \in (0, 1/L]$
5: Choose $M = \lceil Lr/\varepsilon \rceil$
6: **for** $m = 1$ to M **do**
7: $\quad \mu^m \leftarrow \mu^{m-1} + \eta \left(q - \langle Q \rangle_{\rho_T(\mu^{m-1})}\right)$
8: **end for**
9: **return** $\mu^M \cdot q + \langle H - \mu^M \cdot Q \rangle_{\rho_T(\mu^M)}$



Gradient ascent algorithm (classical algorithm)

Algorithm 1 `minimize_energy`$(H, (Q_i)_i, (q_i)_i, L, d, \varepsilon, r)$

Input:

- Observables H and $(Q_i)_{i\in[c]}$
- Constraint values $(q_i)_{i\in[c]}$
- Smoothness parameter L
- Hilbert space dimension d
- Desired error $\varepsilon > 0$
- Radius r: An upper bound on $\|\mu^*\|$, where μ^* is the optimal solution to (14) for $T = \frac{\varepsilon}{4 \ln d}$

Set $T \leftarrow \frac{\varepsilon}{4 \ln d}$

Initialize $\mu^0 \leftarrow (0, \ldots, 0)$

Set learning rate $\eta \in (0, 1/L]$

Choose $M = \lceil Lr/\varepsilon \rceil$

for $m = 1$ to M **do**

$\mu^m \leftarrow \mu^{m-1} + \eta \left(q - \langle Q \rangle_{\rho_T(\mu^{m-1})}\right)$

end for

return $\mu^M \cdot q + \langle H - \mu^M \cdot Q \rangle_{\rho_T(\mu^M)}$

Smoothness parameter $L \geq$ largest value of Hessian of $f(\mu)$

Can set: $$L = \frac{2c}{T} \max_{i\in[c]} \|Q_i\|^2$$

Number of iterations (complexity):

$$M = \left\lceil \frac{Lr}{\varepsilon} \right\rceil$$

$$= \left\lceil \frac{8cr \ln d}{\varepsilon^2} \max_{i\in[c]} \|Q_i\|^2 \right\rceil$$

Hybrid classical-quantum algorithm

Suppose we have an n-qubit system with large n

Need to estimate $\langle H \rangle_{\rho_T}$ and $\langle Q \rangle_{\rho_T}$ on quantum device

Requirements:
1) Efficient preparation of $\rho_T(\mu)$ at $T \sim \epsilon/n$
2) Efficient measurement of observables $\langle H \rangle_{\rho_T}$ and $\langle Q \rangle_{\rho_T}$

Assumptions

Each observable H, Q_i are efficiently measurable on quantum device when

$$H = \sum_{\vec{J}} h_{\vec{J}} \sigma_{\vec{J}},$$

$$Q_i = \sum_{\vec{J}} a_{i,\vec{J}} \sigma_{\vec{J}},$$

for which the number of non-zero terms in each sum is polynomial in n and $\|h\|_1$, $\|a_i\|_1$ are poly(n)

Sampling complexity to estimate observables to precision ϵ with success probability $1-\delta$: $O(\ln(1/\delta)\|h\|_1^1/\epsilon^2)$, $O(\ln(1/\delta)\|a_i\|_1^1/\epsilon^2)$

Gradient ascent algorithm (hybrid classical-quantum algorithm)

Algorithm 2 `minimize_energy`$(H, (Q_i)_i, (q_i)_i, d, \varepsilon, \delta, r)$

Input:

- Observables H and $(Q_i)_{i\in[c]}$ (as given in (39)–(40))
- Constraint values $(q_i)_{i\in[c]}$
- Hilbert space dimension d
- Accuracy $\varepsilon > 0$
- Error probability $\delta \in (0, 1)$
- Radius r: An upper bound on $\|\mu^*\|$, where μ^* is the optimal solution to (14) for $T = \frac{\varepsilon}{4 \ln d}$

Initialize $\mu^0 \leftarrow (0, \ldots, 0)$
Set learning rate η as in (51)
Set number of iterations, M, as in (52)
for $m = 1$ to M **do**
  for $i = 1$ to c **do**
    $\tilde{Q}_i \leftarrow$ `estimate_obs`$(\mu^{m-1}, (a_{i,\vec{\jmath}})_{\vec{\jmath}}, \varepsilon, \delta)$
    $\overline{g}_i(\mu^{m-1}) \leftarrow q_i - \tilde{Q}_i$
  end for
  Update: $\mu^m \leftarrow \Pi_{\mathcal{X}}(\mu^{m-1} + \eta\, \overline{g}(\mu^{m-1}))$
end for
Set $\overline{\mu^M} \leftarrow \frac{1}{M} \sum_{m=1}^{M} \mu^m$
Set $g_{\vec{\jmath}} \leftarrow h_{\vec{\jmath}} - \sum_{i\in[c]} \left[\overline{\mu^M}\right]_i a_{i,\vec{\jmath}}$, for all $\vec{\jmath}$
$\tilde{G} \leftarrow$ `estimate_obs`$(\overline{\mu^M}, (g_{\vec{\jmath}})_{\vec{\jmath}}, \varepsilon/4, \delta)$
return Output $\overline{\mu^M} \cdot q + \tilde{G}$

Do not have perfect $\langle Q \rangle$ values for evaluation of $f(\mu)$ gradient: use *stochastic gradient ascent*

Gradient ascent algorithm (hybrid classical-quantum algorithm)

Theorem: Sample complexity (number of ρ_T preparations) required to estimate the optimal $\langle H \rangle$ to precision ϵ and success probability $\geq 1 - \delta$

$$O\left(\frac{c \max\{c, r^2\} \ln d \, \|h\|_1^2 \max_i \|a_i\|_1^4 \ln\left(\frac{1}{\delta}\right)}{\varepsilon^4}\right)$$

Second-order algorithms also possible using quantum estimator for elements of Kubo-Mori I_{KM} matrix

Original semidefinite optimisation problem

Back to solving original semidefinite optimisation problem with $\mathrm{Tr}(X) \leq R$

Solve

SDP

$$\alpha_R = \min_{X \geq 0} \ \mathrm{Tr}[CX]$$
$$\text{s.t.} \ \ \mathrm{Tr}[A_i X] = b_i, \forall i$$
$$\mathrm{Tr}[X] \leq R$$

Sample complexity:

$$O\left(\frac{cR^4 \max\{c, r^2\} n \|h\|_1^2 \max_i \|a_i\|_1^4 \ln\left(\frac{1}{\delta}\right)}{\varepsilon^4} \right)$$

Examples

Constrained free energy minimization for the design of thermal states and stabilizer thermodynamic systems

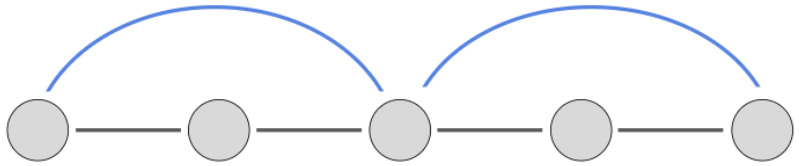
(arXiv: 2508.09103)

Michele Minervini, Madison Chin, Jacob Kupperman, Nana Liu, Ivy Luo, Meghan Ly, Soorya Rethinasamy, Kathie Wang, and **Mark M. Wilde**

Quantum Heisenberg model

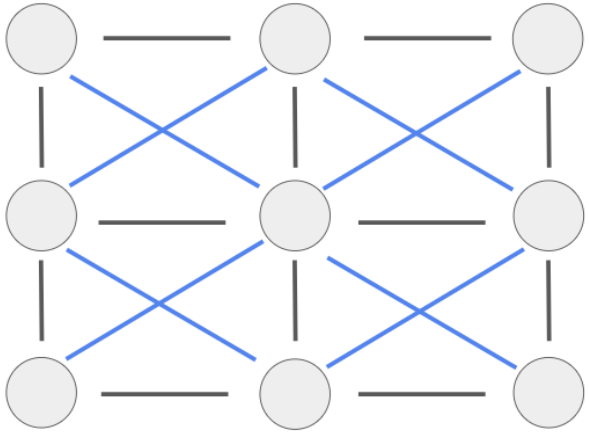
Describes interacting spin systems on a lattice and are essential for understanding magnetic materials

$1D$



$$H(G, \vec{J}) \coloneqq \sum_{\{i,j\} \in E} J_{i,j} \left(X^{(i)} \otimes X^{(j)} + Y^{(i)} \otimes Y^{(j)} + Z^{(i)} \otimes Z^{(j)} \right)$$

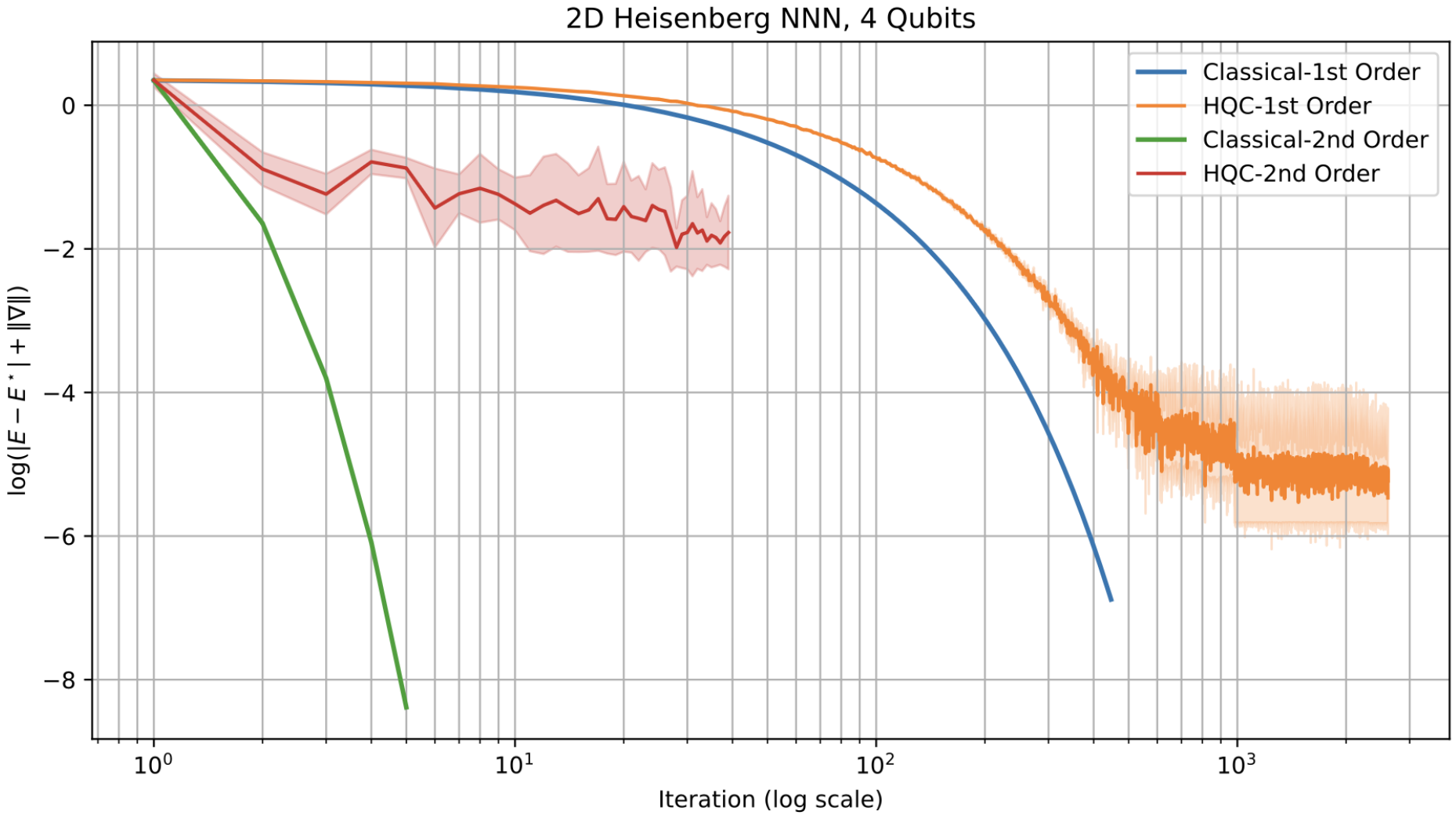
$2D$



Total magnetisation:

$$X(G) \coloneqq \sum_{i \in V} X^{(i)}, \quad Y(G) \coloneqq \sum_{i \in V} Y^{(i)}, \quad Z(G) \coloneqq \sum_{i \in V} Z^{(i)}.$$

These all commute individually with $H(G, J)$
but not each other, so form conserved non-commuting charges



The figure depicts the logarithm of the error metric in (118) versus the logarithm of the number of iterations, for the task of constrained energy minimization for the two-dimensional, four-qubit quantum Heisenberg model with nearest- and next-to-nearest-neighbor interactions and constraints on the total magnetizations in the x, y, and z directions set to be 1, 0, and 1, respectively. All of the algorithms converge, but the HQC algorithms, shown as the average over five independent runs with shaded regions denoting one standard deviation, require more iterations to converge due to sampling noise inherent in them.

Stabiliser thermodynamic systems

Definition (Stabilizer thermodynamic system)**.** Let $\mathcal{S}$ denote a stabilizer code that encodes k logical qubits into n physical qubits, has commuting stabilizer generators $S_1, \ldots, S_{n-k}$, and has the set $\mathcal{L}$ of logical operators. We define a stabilizer thermodynamic system to have a Hamiltonian H given by

$$H := -\sum_{i=1}^{n-k} \gamma_i S_i,$$

where $\gamma_i > 0$, and conserved, non-commuting charges given by

$$Q_i := \sum_j \alpha_{i,j} L_{i,j},$$

where $L_{i,j} \in \mathcal{L}$ and $\alpha_{i,j} \in \mathbb{R}$, for all $i \in [c]$ and j.

A stabiliser $\mathcal{S}$ is an abelian subgroup of the Pauli group not containing $-I$

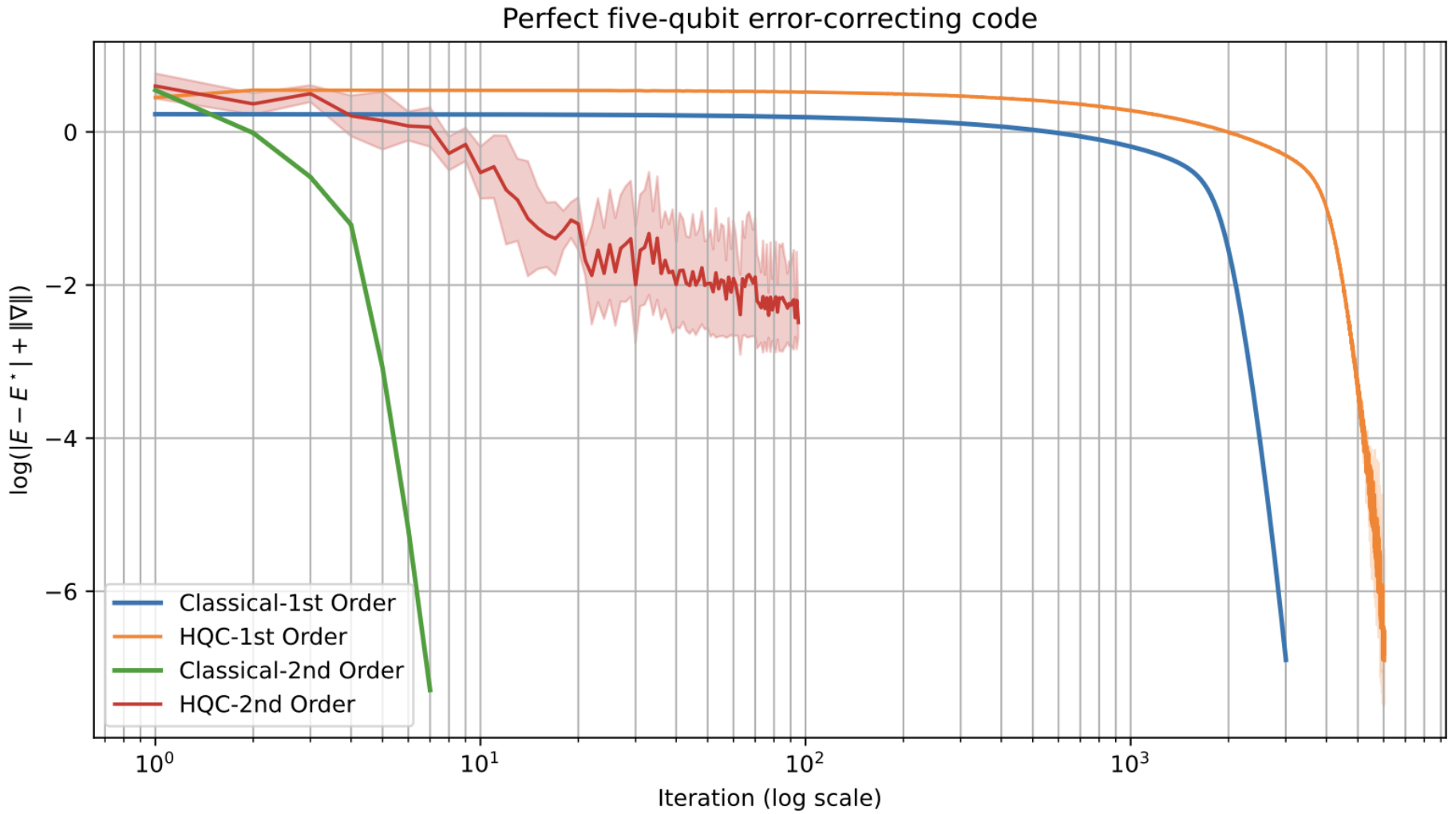
Code space:

$$\mathcal{C} := \text{span}\left\{|\psi\rangle \in (\mathbb{C}^2)^{\otimes n} : S|\psi\rangle = |\psi\rangle \ \ \forall S \in \mathcal{S}\right\}$$

Set of logical operators:

$$\mathcal{L} = \{L \in \text{Pauli group} : \ L\mathcal{S} = \mathcal{S}L\}$$

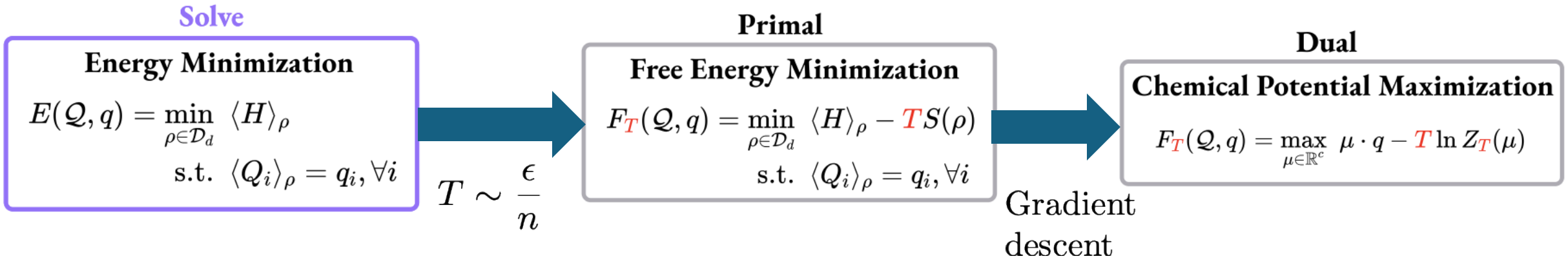
$$[H, Q_i] = 0 = [S, L_i],$$
$$[Q_i, Q_j] \neq 0$$



The figure depicts the logarithm of the error metric in (118) versus the number of iterations, for all of the LMPW algorithms (1st- and 2nd-order, and classical and HQC) for the task of constrained energy minimization for a stabilizer thermodynamic system formed from the perfect five-qubit error-correcting code. The constraints on the logical operators $\overline{X}$, $\overline{Y}$, and $\overline{Z}$ were set to 0.2, 0, and 0.5, respectively. All of the algorithms converge, but the HQC algorithms, shown as the average over five independent runs with shaded regions denoting one standard deviation, require more iterations to converge due to sampling noise inherent in them. For this simulation, we did not warm-start the algorithm according to the recipe from Section V A, but we instead started with $\mu_x = 1$, $\mu_y = 1$, and $\mu_z = 1$.

Summary

- *Physical basis* for *why* quantum thermal states with non-commuting charges (a quantum Boltzmann machine) are important and *natural* for solving semidefinite optimisation



- Extension to continuous-variable systems? Broader classes of SDPs? Thermodynamically-inspired implementation?

Mark Wilde

Michele Minervini

Dhrumil Patel

1. Quantum thermodynamics and semi-definite optimization
arXiv: 2505.04514

2. Constrained free energy minimization for the design of thermal states and stabilizer thermodynamic systems
arXiv: 2508.09103

Thank you and welcome to visit us in Shanghai Jiao Tong!

Shanghai Jiao Tong University:

Institute of Natural Sciences

Global College (formerly called University of Michigan-Shanghai Jiao Tong Joint Institute)

Open postdoc positions available!

UnitaryLab internship and more senior positions in company also open!

Nana.Liu@quantumlah.org

www.nanaliu.weebly.com

Group website:
https://www.quantumquintet.com/