Limitations of optimization on noisy quantum devices

Daniel Stilck França (QMATH, University of Copenhagen)
Joint work with Raúl García-Patrón (University of Edinburgh)

arXiv:2009.05532

UNIVERSITY OF COPENHAGEN

"Ever heard of D-Wave?" Slade asks as Helena takes a sip of a white burgundy, the best wine she has ever tasted.

"Sorry, I haven't."

"It's a company out of British Columbia. A year ago, they released a prototype quantum processor. Its application is highly specific, but ideal for the sort of enormous data-set mapping problem we've run up against."

"How much are they?"

"Not cheap, but I was interested in the technology, so I ordered a few of their advanced prototypes for future projects last summer."

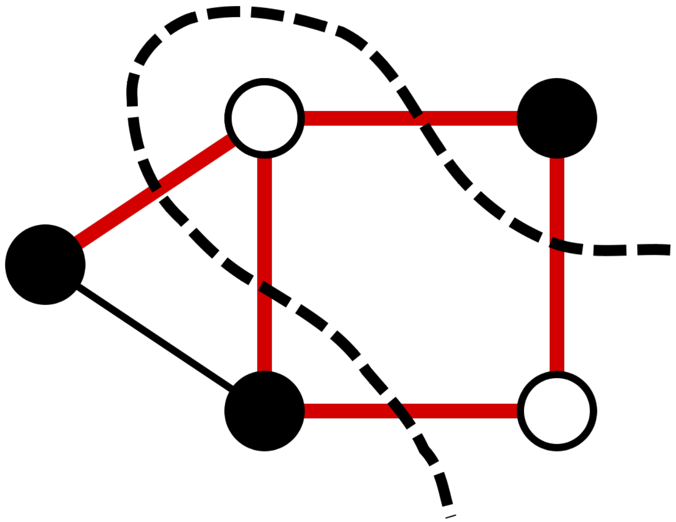
Excerpt From: Blake Crouch. "Recursion."

- The main character reshapes reality.
- All based on computations done on a D-Wave machine!
- Expectations are high!
- **Guiding question:** is it realistic to expect NISQ devices to achieve that? Or maybe just a *unequivocal quantum speedup for optimization?*
- **Main focus:** limitations imposed by the noise.
- **Conclusion:** opportunity window for reshaping of reality slim.

Setting: problem types

- Given some Hamiltonian *H* whose ground state encodes the solution to a problem.
- Can be classical or quantum. Will focus on classical problems here for simplicity, so think of

$$H_I = -\sum_{i<j}^{n} a_{i,j} Z_i Z_j - \sum_{i=1}^{n} b_i Z_i$$

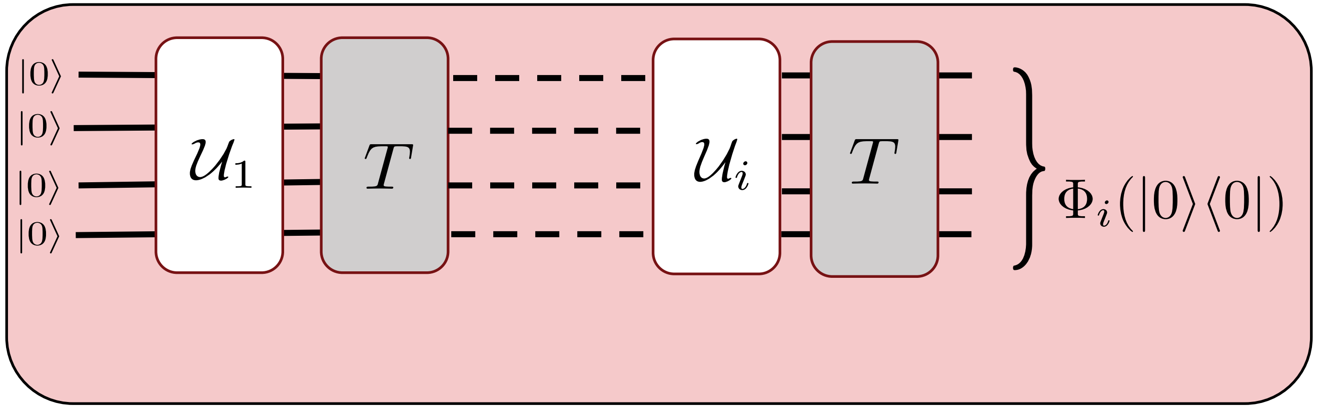


- Equivalent to MAX-CUT.
- Minimize $\mathrm{tr}\left[H_I \Phi_D(\rho)\right]$

No error correction!

Setting: noise model

- Want to implement unitary circuit.
- Assume they are interspersed by noisy channel.
- This talk: 1-qubit depolarizing noise.
- Paper: more general, also nonunital and continuous time.



How does noise constrain the performance of such algorithms?

- Noisy quantum circuits: assume each gate is followed by 1-qubit depolarizing noise with probability *p*.
- Intuition: lose advantage at 1/p.
- **Rigorously confirm this intuition. At depth**

$$\sim p^{-1}$$

we perform as well as high temperature Gibbs sampling.

How does noise constrain the performance of such algorithms?

- At which depth do noisy quantum computers lose advantage against *nontrivial* classical algorithms?
- And can we estimate this without having to simulate the noisy circuit?
- **Achieve this and can show that when a good classical algorithm is available, advantage is lost at**

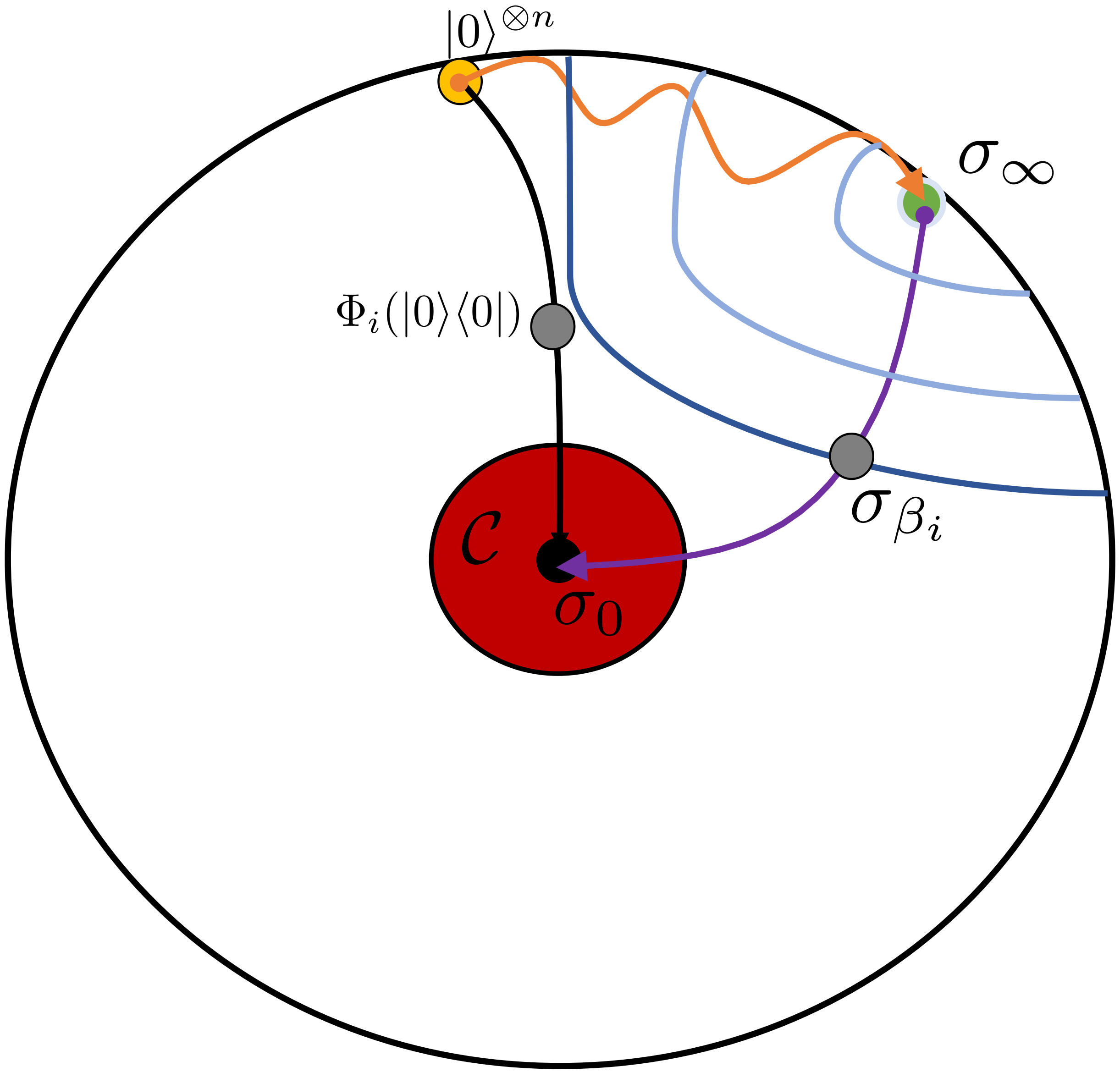
$$cp^{-1}, c \ll 1$$

Comparison to previous work

- Aharanov et al (96), Gottesman et al (13), Müller-Hermes et al (15) and others: used entropic methods to establish when quantum computers lose advantage under noise.
- Different setups, but obtain less stringent bounds on maximal depth.
- Razborov (03), Plenio et al (10), Kempe et al (08) and others: reverse threshold theorem.
- Stronger statements, but do not apply to current noise rates.

Alternative path imposed by the noise

- Quantify convergence to easy state in relative entropy.
- Assign equivalent inverse temperature to each stage of the noisy computation.
- Find inverse temperature region that can be simulated efficiently.



Framework: entropic convergence

- **Step 1:** have a quantum circuit whose noise channel T is driving it to a classically "easy" state .
- **Crucial:** can quantify how fast this happens in relative entropy:

$$\forall \rho : D(T(\rho)\|\sigma) \leq \mu D(\rho\|\sigma)$$

- Example: for 1-qubit depolarizing noise this holds with $\mu = (1-p)^2$

- For any noisy circuit of depth D Φ_D we have:

$$D(\Phi_D(\rho)\|I/2^n) \leq (1-p)^{2D} \ln(2) n$$

Framework: from relative entropy to inverse temperature

- **Step 2.A:** relate the expected output's energy to that of a Gibbs state.
- Let Φ_D be a noisy quantum circuit of depth D. Then for all initial states ρ:

$$\mathrm{tr}\left[\Phi_D(\rho)H\right] \geq \mathrm{tr}\left[\sigma_\beta H\right] - \epsilon\|H\|, \quad \sigma_\beta = \frac{e^{-\beta H}}{\mathcal{Z}_\beta}, \quad \beta = \frac{D(\Phi_D(\rho)\|\sigma_0)}{\|H\|\epsilon}$$

- Optimization: only care about lower bound.

Framework: threshold inverse temperature

- **Step 3.A:** efficient classical Gibbs sampling for small enough $\beta \leq \beta_c$ for H.

- Let Φ_D be a noisy quantum circuit of depth D. Then for all initial states ρ:

$$\operatorname{tr}\left[\Phi_D(\rho)H\right] \geq \operatorname{tr}\left[\sigma_\beta H\right] - \epsilon\|H\|, \quad \sigma_\beta = \frac{e^{-\beta H}}{\mathcal{Z}_\beta}, \quad \beta = \frac{D(\Phi_D(\rho)\|\sigma_0)}{\|H\|\epsilon}$$

- If $\frac{D(\Phi_D(\rho)\|\sigma_0)}{\|H\|\epsilon} \leq \beta_c$ there exists a classically easy state that performs almost as well.

Example: Ising Hamiltonians

- Let $H_I = -\sum_{i<j}^{n} a_{i,j} Z_i Z_j - \sum_{i=1}^{n} b_i Z_i$ and $\|A\|$ be the operators of $A = (a_{ij})$
- **Step 3.A: (Zeitouni et al)** we can sample efficiently for

$$\beta \leq \frac{1}{2\|A\|}$$

Example: Ising Hamiltonians

- Under 1-qubit depolarizing after depths $D \geq D_{\max}$:

$\simeq 0$

$$D_{\max} = \frac{\ln(2\ln(2)\epsilon^{-1}) + \ln\left(\|H_I\|^{-1}\|A\|n\right)}{2p}$$

there is a classical Gibbs state σ_β that can be sampled in $\mathcal{O}(n\log(n))$ time on a classical computer such that:

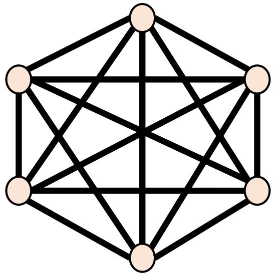
$$\operatorname{tr}\left[H_I\Phi_D(\rho)\right] \geq \operatorname{tr}\left[H_I\sigma_\beta\right] - \|H_I\|\epsilon.$$

Bounding layers of QAOA

- If the device's topology does not match the problem's, the physical depth of QAOA scales superlogarithmically with system size.
- This makes speedups impossible even with *very* small error rates.

a) SK model

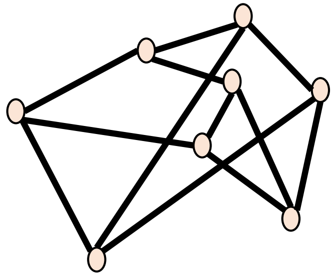
Gate type	Gates per QAOA layer
2 qubit	3N
All	7N



		Gate error probability $p_2 \neq 0$ and $p_1=0$		
SK model	**ε**	**6 10^{-3}**	**10^{-4}**	**10^{-5}**
QAOA layers (n=50)	0.1	1	88	876
Max prob. size (L=10)	0.1	7	438	4,380

b) 3-regular graph

Gate type	Gates per QAOA layer
2 qubit	$\sim\sqrt{7N}$
All	7N

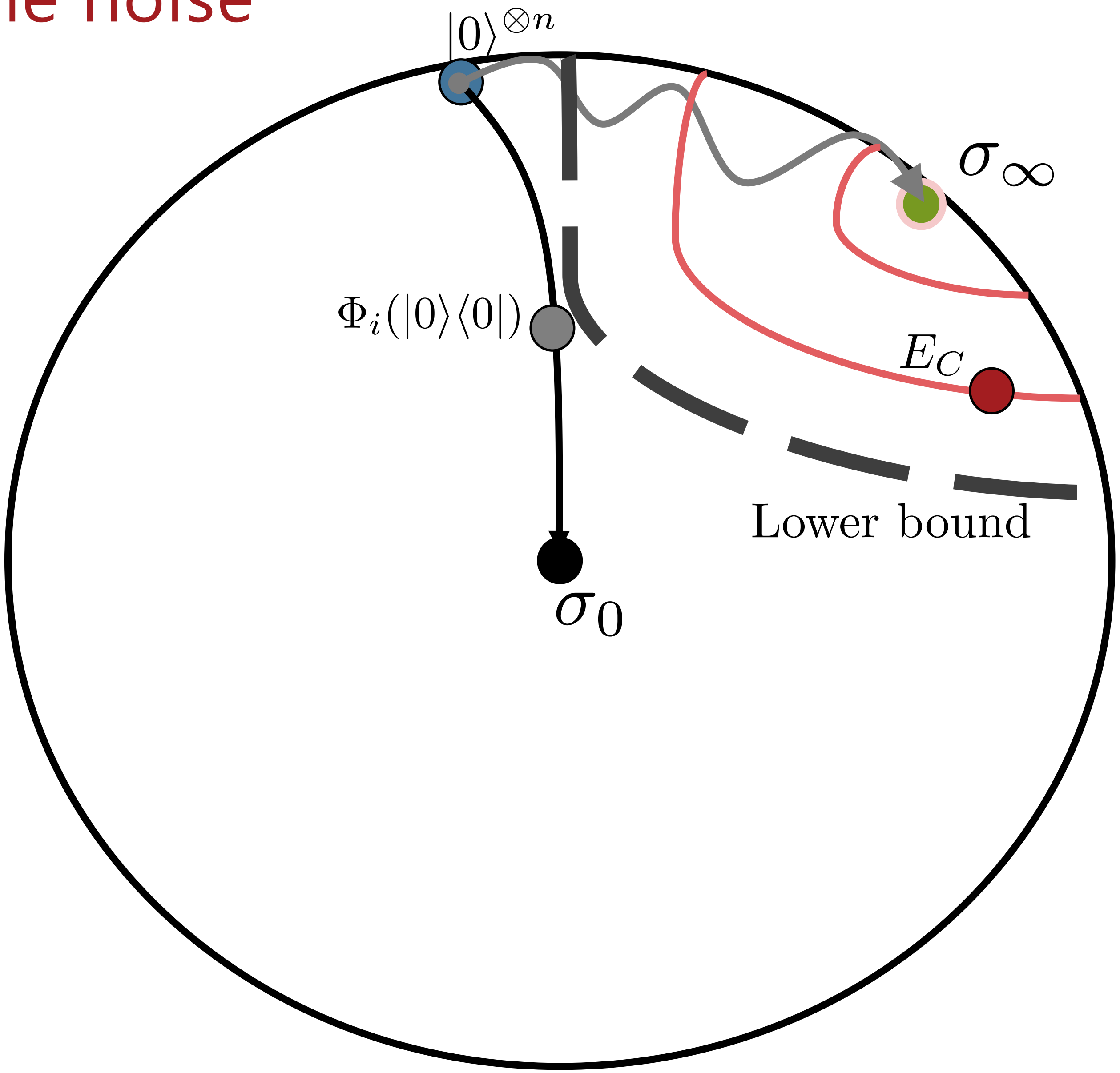


		Gate error probability $p_2=p_1\neq 0$		
3-regular graph	**ε**	**10^{-3}**	**10^{-4}**	**10^{-5}**
QAOA layers (n=50)	0.1	4	38	376
Max prob. size (L=10)	0.1	19	188	1,880

But how to compare to practical algorithms?

Alternative path imposed by the noise

- Quantify convergence to easy state in relative entropy.
- Compute lower bound on output energy by estimating the partition function.
- Compare to output of classical algorithm.



Framework

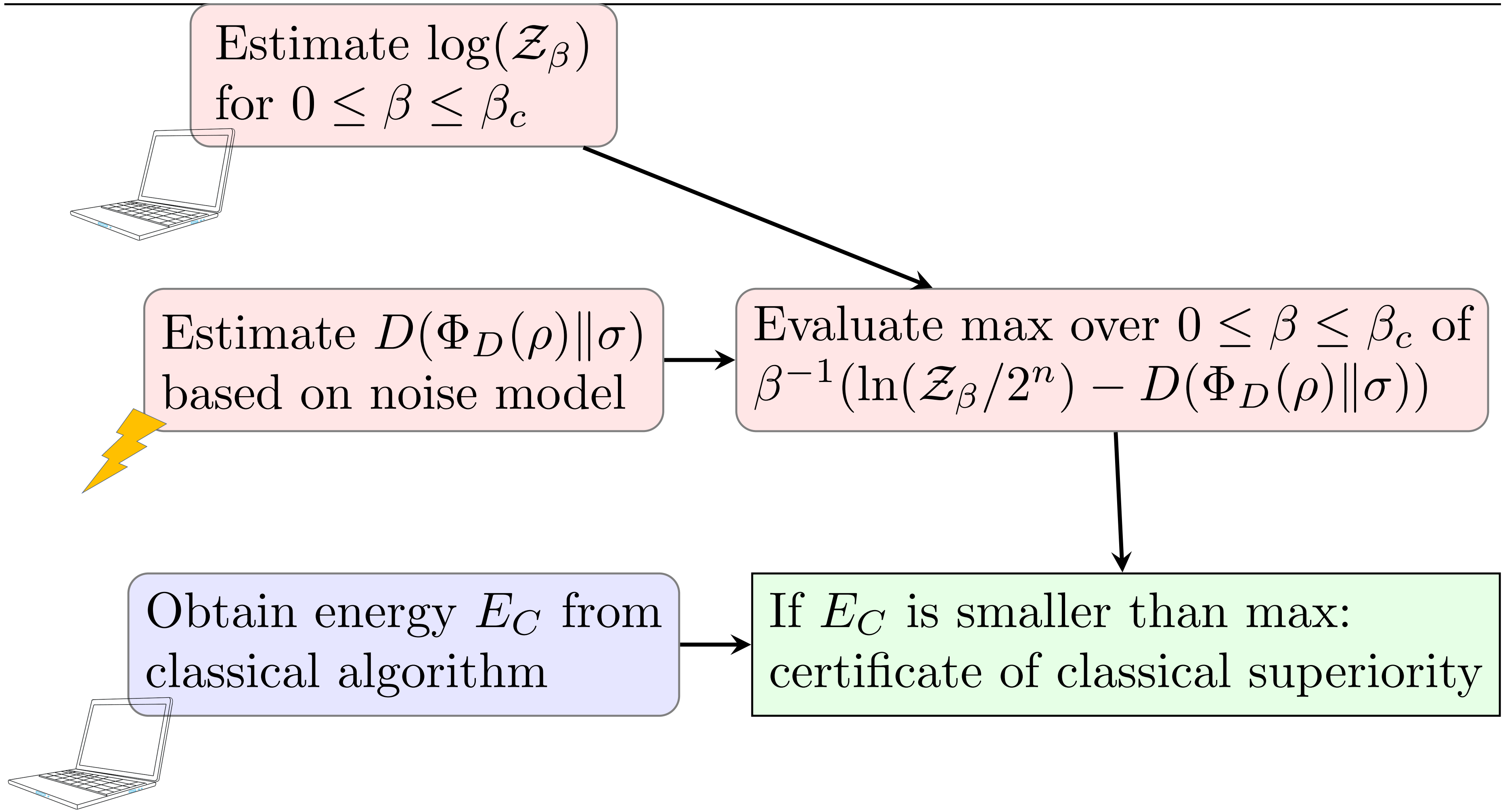
- **Step 2.B:** relate the expected output's energy to the partition function.

- Let Φ_D be a noisy quantum circuit of depth D. Then for all initial states ρ:

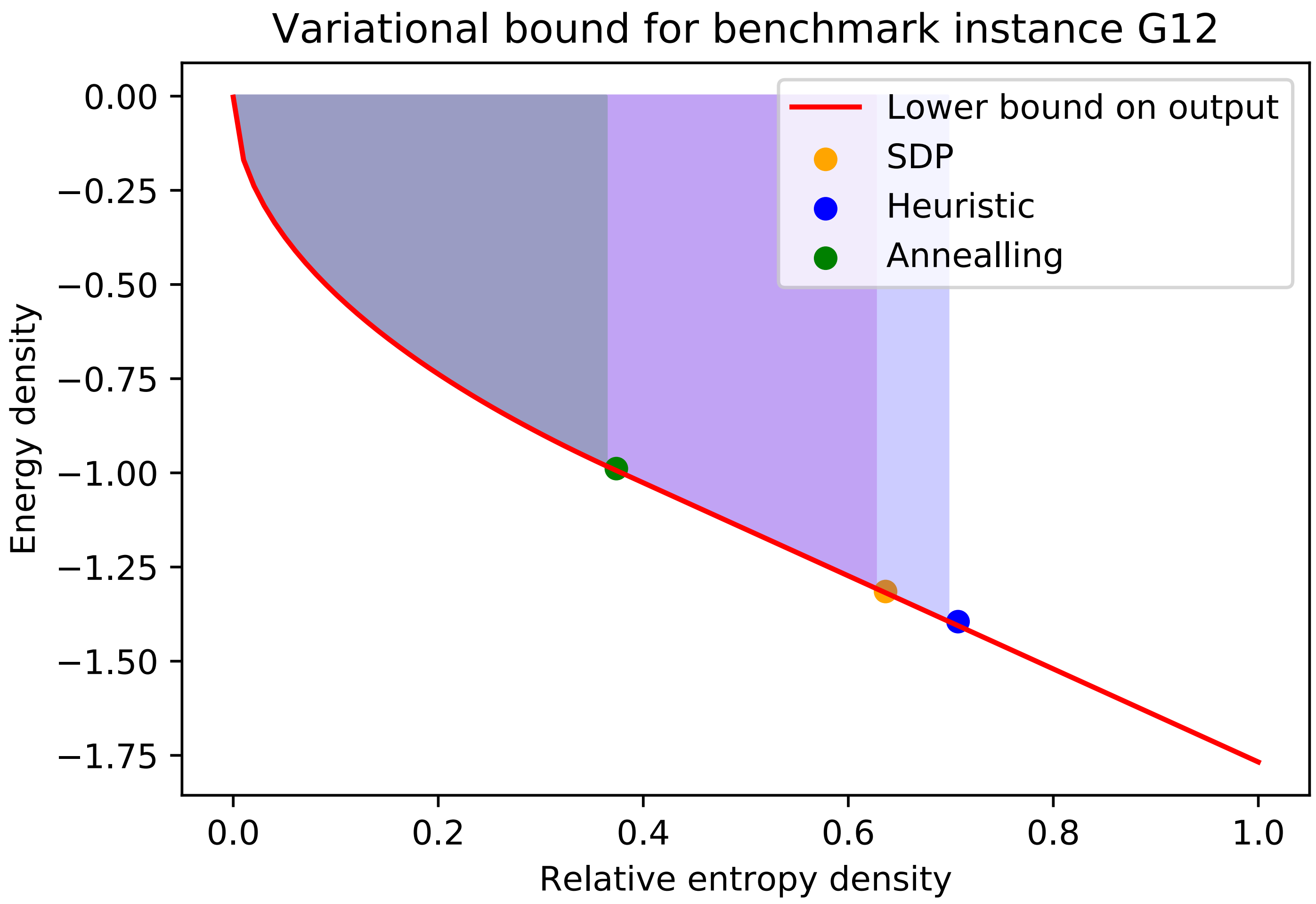
$$\operatorname{tr}\left[\Phi_D(\rho)H\right] \geq \sup_{\beta>0} \beta^{-1}\left(-\ln(\operatorname{tr}\left[e^{-\beta H}\right]/2^n) - D(\Phi_D(\rho)||\sigma_0)\right).$$

- **Step 3.B:** evaluate the partition function for range $\beta \leq \beta_c$ and obtain lower bounds.

Flowchart of verification



Example: Ising model (Gset instance 12)



800 qubits.
Predicts advantage will be lost already at depth $0.18 \times p^{-1}$ against algorithms that run in less than a second on a laptop.

Conclusion and open questions

- **Take home message 1:** If the problem's topology does not match the device's, a quantum advantage is unlikely for variational algorithms in the NISQ era.
- **Take home message 2:** have a framework to assess the opportunity window for quantum speedups under noise against classical algorithms.
- Conclusions are less stringent for quantum problems. Larger opportunity window.

Conclusion and open questions

- Impact of error mitigation/primitive error correction?
- Analysis beyond first order?
- Larger scale numerics/more specialized analysis?

Thanks! Questions?