RL-Driven Quantum Computation

POOYA RONAGH

1QB Information Technologies (1QBit)
Head of Hardware Innovation Lab

Institute for Quantum Computing (IQC), University of Waterloo
Postdoctoral Research Fellow

Perimeter Institute Quantum Intelligence Lab (PIQuIL)
Scientific Lead

Kyle Mills
Arthur Pesah
Background

Quantum simulation

**RESEARCH ARTICLES**

Universal Quantum Simulators

Seth Lloyd
+ See all authors and affiliations

*Science* 23 Aug 1996:
Vol. 273, Issue 5278, pp. 1073-1078
DOI: 10.1126/science.273.5278.1073

**REPORT**

Simulated Quantum Computation of Molecular Energies

Alán Aspuru-Guzik¹,², Anthony D. Dutoi¹,², Peter J. Love², Martin Head-Gordon¹,³

←¹ These authors contributed equally to this work.
+ See all authors and affiliations

*Science* 09 Sep 2005:
Vol. 309, Issue 5741, pp. 1704-1707
DOI: 10.1126/science.1113479
Background

Limitations of stoquastic annealers

- AM
- QMA
- StoqMA
- MA
- NP

k-local stoquastic Hamiltonian
[Bravyi, DiVincenzo, Oliveira, Terhal, 08]

k-SAT
Background

Limitations of stoquastic annealers

- k-local Hamiltonian
  [Kempe, Kitaev, Regev 06]

- k-local stoquastic Hamiltonian
  [Bravyi, DiVincenzo, Oliveira, Terhal, 08]

Diagram:

- AM
- QMA
- StoqMA
- MA
- NP
- k-SAT
Background

Limitations of stoquastic annealers

AM

QMA

k-local Hamiltonian
[Kempe, Kitaev, Regev 06]

StoqMA

k-local stoquastic Hamiltonian
[Bravyi, DiVincenzo, Oliveira, Terhal, 08]

MA

NP

k-SAT

k-sat
**Background**

*Limitations of stoquastic annealers*

- **AM**
- **QMA**
- **StoqMA**
- **MA**
- **NP**

**k-local Hamiltonian**
Kempe, Kitaev, Regev 06

**k-local stoquastic Hamiltonian**
Bravyi, DiVincenzo, Oliveira, Terhal, 08

**[Denchev et al. 2016]**

Pooya Ronagh | 1QBit | IQC | UW
Background

Good news: more control. Bad news: more control

Demonstration of nonstoquastic Hamiltonian in coupled superconducting flux qubits

I. Ozfidan,1,∗ C. Deng,1,∗ A. Y. Smirnov,1 T. Lauting,1 R. Harris,1 L. Swenson,1 J. Whittaker,1 F. Altomare,1 M. Babcock,1 C. Baron,1 A.J. Berkley,1 K. Boothby,1 H. Christiani,1 P. Bunyk,1 C. Enderud,1 B. Evert,1 M. Hager,1 A. Hajda,1 J. Hilton,1 S. Huang,1 E. Hoskinson,1 M.W. Johnson,1 K. Jooya,1 E. Ladizinsky,1 N. Ladizinsky,1 R. Li,1 A. MacDonald,1 D. Marsden,1 G. Marsden,1 T. Medina,1 R. Molavi,1 R. Neufeld,1 M. Nissen,1 M. Norouzpour,1 T. Oh,1 I. Pavlov,1 I. Perminov,1 G. Poulin-Lamarre,1 M. Reis,1 T. Prescott,1 C. Rich,1 Y. Sato,1 G. Sterling,1 N. Tsai,1 M. Volkman,1 W. Wilkinson,1 J. Yao,1 and M. H. Amin1,2

1D-Wave Systems Inc., 3023 Beta Avenue, Burnaby BC, Canada, V5G 4M9
2Department of Physics, Simon Fraser University, Burnaby BC, Canada, V5A 1S6
(Dated: March 19, 2019)

Exploring More-Coherent Quantum Annealing

Sergey Novikov1, Robert Hinkey1, Steven Dsseler1, James I. Basham1, Tameem Albash1, Andrew Risinger1, David Ferguson1, Damiad Ladar1,∗, Kennketh M. Zick1
1Northrop Grumman Corporation, Linthicum, MD 21090 USA
sergey.novikov@nngc.com
2Information Sciences Institute, University of Southern California, Marina del Rey, CA 90292 USA
3Department of Physics and Astronomy, University of Southern California, Los Angeles, California 90089 USA
4Center for Quantum Information Science & Technology, University of Southern California, Los Angeles, California 90089 USA
5Department of Electrical Engineering, University of Southern California, Los Angeles, California 90089, USA
6Department of Chemistry, University of Southern California, Los Angeles, California 90089, USA
[albash, ladar]@usc.edu

Pooya Ronagh | 1QBit | IQC | UW
Background

Good news: more control. Bad news: more control

\[ H(s, \lambda) = -s \lambda N \left( \frac{1}{N} \sum_{i=1}^{N} \sigma_i^z \right)^p + s(1 - \lambda) N \left( \frac{1}{N} \sum_{i=1}^{N} \sigma_i^x \right)^k - (1 - s) \sum_{i=1}^{N} \sigma_i^x \]

\[ H_0 = - \sum_{i_1, \ldots, i_p=1}^{N} J_{i_1 \ldots i_p} \sigma_{i_1}^z \cdots \sigma_{i_p}^z \]

Background

Good news: more control. Bad news: more control

\[ H(\tau) = (1 - \tau)H_B + \lambda \tau (1 - \tau)H_I + \tau H_P \]

\[ H_P = \sum_{i<j=1}^{N} J_{ij} \sigma_i^z \sigma_j^z + \sum_{i=1}^{N} h_i \sigma_i^z \quad H_B = \sum_{i=1}^{N} \sigma_i^x \]

\[ H_I^F = - \sum_{i<j=1}^{N} \sigma_i^x \sigma_j^x \quad \text{stoquastic} \]

\[ \begin{align*}
H_I^A &= + \sum_{i<j=1}^{N} \sigma_i^x \sigma_j^x, \\
H_I^M &= \sum_{i<j=1}^{N} r_{ij} \sigma_i^x \sigma_j^x.
\end{align*} \quad \text{non-stoquastic} \]

Background

Good news: more control. Bad news: more control

\begin{align*}
H(\tau) &= (1 - \tau)H_B + \lambda \tau (1 - \tau) H_I + \tau H_P \\
H_P &= \sum_{i<j=1}^{N} J_{ij} \sigma^z_i \sigma^z_j + \sum_{i=1}^{N} h_i \sigma^z_i \\
H_B &= \sum_{i=1}^{N} \sigma^x_i \\
H_I^F &= - \sum_{i<j=1}^{N} \sigma^x_i \sigma^x_j \quad \text{stoquastic} \\
H_I^A &= + \sum_{i<j=1}^{N} \sigma^x_i \sigma^x_j, \quad \text{non-stoquastic} \\
H_I^M &= \sum_{i<j=1}^{N} r_{ij} \sigma^x_i \sigma^x_j.
\end{align*}

Stoquastic:
\begin{itemize}
\item Fraction of the affected instances is large and increases with N.
\item Majority of improved instances are easy for vanilla QA and only marginal improvement in success probabilities.
\item Minimum gaps are large and they further increase once HF is applied.
\item Numbers of anticrossings decrease.
\end{itemize}

Nonstoquastic:
\begin{itemize}
\item Fractions of affected instances is small and remain relatively constant with N.
\item Majority of affected instances are hard and significant improvements to the initial success probabilities.
\end{itemize}

The Goal

Use optimal control (RL) to solve this optimal control problem!

What we want to do
Treat control of a quantum system as a temporal optimal control problem and have a controller decide/change the course of QC as it happens.
Applications:
Quantum annealing
Non-stoquastic QA
Classical Monte-Carlo
NISQ: VAE, QAQA, etc.
Error mitigation/correction
RL-driven non-stoquastic path

The p-spin model in finite size
RL-driven non-stoquastic path

*The p-spin model in finite size*
Weak-Strong Cluster Model

The power of collective tunneling
Weak-Strong Cluster Model

The power of collective tunneling

Mandra et al. 2016
Weak-Strong Cluster Model

The power of collective tunneling

Mandra et al. 2016

16-spin distribution of energies:

-2.53 (global min)
-2.47 (local min)
1.97
-1.03

2.25 (global max, degenerate)
Weak-Strong Cluster Model

The power of collective tunneling

\[ o \in \{-1, 1\}^{N_{\text{spin}} \times N_{\text{reads}}} \]

<table>
<thead>
<tr>
<th>( N_{\text{spin}} )</th>
<th>Number of nodes (spins) in the graph</th>
</tr>
</thead>
</table>

1 sweep

\( N_{\text{spin}} \) random spin flips

1 step

One opportunity to interact with the environment (receive state, suggest action)

\( N_{\text{steps}} \) Number of steps per episode

\( N_{\text{sweeps}} \) Total number of sweeps per episode, thus \( N_{\text{sweeps}} / N_{\text{steps}} \) sweeps are performed per step.

\( N_{\text{reads or rep}} \)

Number of simultaneous anneals

\( N_{\text{buffer}} \)

Size of the buffer, i.e. how many steps to take between updates to the policy network
RL on Weak-Strong Cluster Model

The model

- Can we use reinforcement learning (RL) to learn a dynamic temperature schedule to consistently find the ground state of the weak-strong clusters model.
RL on Weak-Strong Cluster Model

Simulations

![Graphs showing simulations results]
Reward structure

Numerical results

- Sparse reward
  - reward the agent with the average (over reads) final energy of an episode
- Dense reward
  - reward the agent every step with the average energy difference between the previous state and the new state.
    - Overall, better policies will have better more decreases in energy than increases.
- Engineering rewards from syndrome measurements
- Might want a nonlinear reward transformation, e.g. \( R = \exp(-\sum(dE)) \)
Sparse reward, $N_{\text{steps}}=10$, $N_{\text{buffer}}=8192$
Random initial temperature

Avoid any hyperparameter tuning

4000 sweeps
Destructive measurements

Collapse of the wave-function
Destructive measurements

Collapse of the wave-function
Destructive measurements

Collapse of the wave-function

4000 sweeps
Generalization

Turning this game into an efficient classical solver outperforming Monte-Carlo simulations

Question: Can we use RL to learn an adaptive temperature schedule to consistently find the ground state?
Classes of problems: 2D toroidal models with no local field biases and using normally-distributed couplers selected from $N(0, 0.5, [-1, 1])$. 

Pooya Ronagh | 1QBit | IQC | UW
Generalization

Turning this game into an efficient classical solver outperforming Monte-Carlo simulations
Concluding remarks

Results

Outlook
Concluding remarks

Results

- RL has been previously used for solving quantum control problems; here we introduce RL-driven quantum computation as a method for **guiding the course of quantum computation**.
- Applications:
  - NISQ algorithms
  - Non-stoquastic adiabatic quantum computation
  - Classical Monte-Carlo simulations
  - Hyperparameter scheduling

Outlook

- Experiments on prototype devices
- Large scale classical optimization
- Hyperparameter tuning and scheduling software stack
1QBit by the Numbers

Industry leader in quantum software and solutions

2012  108  45  30  21

Founded  Person team with 45 PhDs  Million CAD raised in Series B  Patents filed with 10 granted  Research papers
Hardware Innovation

**Goals:** Support the production of advanced hardware as back-end of 1QBit's software stacks

**Partners:** Superconducting, Optics and photonics, Classical HPC

**Deliverables:**
- Creating, communicating, and supporting milestones for industrially viable hardware products.
- Creating, communicating, and supporting roadmaps for experimental prototypes.
- Support the design and fabrication of experimental prototypes.
- R&D in the design and analysis of the hardware, as well as applications of them.

We are hiring!
pooya@1qbit.com