Testing importance sampling on a quantum annealer for strong coupling lattice QCD

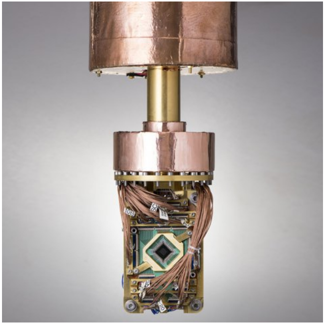
Jangho Kim
IAS-4, Forschungzentrum Jülich

in collaboration with Thomas Luu, Wolfgang Unger

August 2, 2023@Fermilab

$arXiv : 2305.18179$

D-wave Quantum Computer



- The D-Wave quantum annealer consists of an array of metal loops with Josephson junctions.
- The two-state level system of each superconducting loop constitutes a single qubit.
- D-Wave's quantum processing units (QPUs) are composed of qubits placed in arrays and coupled in pairs.

Annealing process

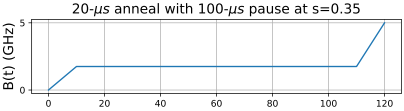
- The array of qubits can be described as an form of Ising spin glass,

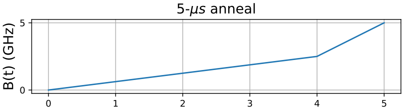
$$H_{QUBO} = -\sum_{i<j} Q_{ij}\sigma_z^i\sigma_z^j + \sum_i Q_i\sigma_z^i \quad (1)$$

- Q is a upper-triangle or a symmetric matrix.
- By applying an external magnetic field, a non-commuting transverse field σ_x is introduced at each site i.
- The full Hamiltonian is expressed by the transverse field and the QUBO Hamiltonian with time-dependent coefficients $A(t)$ and $B(t)$,

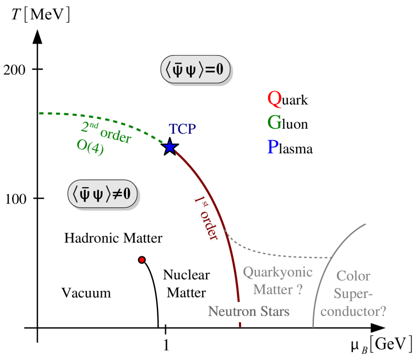
$$H(s) = -A(t)\sum_i \sigma_x^i + B(t)H_{QUBO} \quad (2)$$

- Starting from $A/B \gg 1$ at $t=0$, it reacheds $A/B \approx 0$ after 'anneal time' t_f.
- We use two customized annealing profiles. The behavior of $A(t)$ is roughly inversely proportional to $B(t)$.

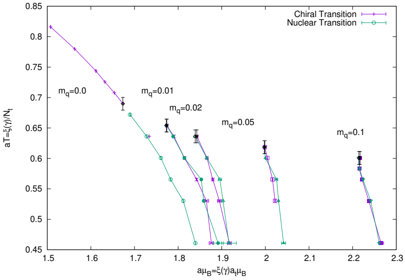




Strong Coupling Lattice QCD



(1) Speculated QCD phase diagram



(2) Simulation at $\beta = 0$

- Strong Coupling Lattice QCD is an effective theory of QCD at the zero limit of inverse coupling $\beta = 2N_c/g^2 = 0$.
- SCQCD shares important features with QCD, confinement, chiral symmetry breaking and restoration at the chiral transition temperature and nuclear liquid gas transition.
- It is extendable to finite inverse coupling β with gauge corrections.

Strong Coupling Lattice QCD Dual representation

- Partition function is

$$Z = \sum_{\{k,n,\ell\}}^{GC} \underbrace{\prod_{b=(x,\hat{\mu})} \frac{(N_c - k_b)!}{N_c! k_b!} \gamma^{2k_b \delta_{\hat{0},\hat{\mu}}}}_{\text{meson hoppings}} \underbrace{\prod_x \frac{N_c!}{n_x!} (2am_q)^{n_x}}_{\text{chiral condensate}} \underbrace{\prod_\ell w(\ell, \mu)}_{\text{baryon hoppings}} \quad (3)$$

$$w(\ell) = \frac{1}{\prod_{x \in \ell} N_c!} \sigma(\ell) \gamma^{N_c N_{\hat{0}}} \exp\left(N_c N_t r_\ell a_t \mu\right) \quad (4)$$

- k_b : bond occupation number $(0 \sim N_c)$
- n_x : site occupation number $(0 \sim N_c)$
- γ : anisotropy (a/a_t) (Changing temperature continuously)
- am_q : quark mass
- $\sigma(\ell)$: sign factor (± 1)
- Grassmann Constraint(GC)

$$n_x + \sum_{\pm\hat{\mu}} \left(k_{x,\hat{\mu}} + \frac{N}{2} |\ell_{x,\hat{\mu}}| \right) = N_c \quad (5)$$

$U(1)$ gauge theory ($N_c = 1$)

- For the first, we choose the simplest gauge group $U(1)$.
- Partition function

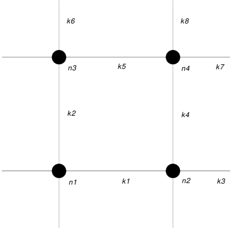
$$Z = \sum_{\{conf\}} e^{-S} = \sum_{\{k,n,\ell\}}^{GC} \prod_{b=(x,\hat{\mu})} \gamma^{2k_b\delta_{0,\hat{\mu}}} \prod_{x} (2am_q)^{n_x} \tag{6}$$

$$S = -\sum_{b=(x,\hat{\mu})} 2k_b\delta_{0,\hat{\mu}} \log(\gamma) - \sum_{x} n_x \log(2am_q) \tag{7}$$

- $k_b \in \{0,1\}$ and $n_x \in \{0,1\}$
- $\vec{k}_b^T = (k_1, k_2, ..., k_E)$, $\vec{n}_x^T = (n_1, n_2, ..., n_V)$
- Property of binary number $k_i^2 = k_i$, $n_i^2 = n_i$
- The action is written in diagonal weight matrix form with binary vector x.

$$S = x^T W x = (\vec{k}_b^T, \vec{n}_x^T) \begin{pmatrix} -2\delta_{0,\hat{\mu}} \log(\gamma) \mathbb{1}_{E\times E} & \mathbb{0}_{V\times E} \\ \mathbb{0}_{E\times V} & -\log(2am_q)\mathbb{1}_{V\times V} \end{pmatrix} \begin{pmatrix} \vec{k}_b \\ \vec{n}_x \end{pmatrix} \tag{8}$$

$U(1)$ on 2×2 Lattice



- Grassmann constraint

$$\sum_{\mu=\pm 0,\cdots,\pm d} k_\mu(x) + n_x = 1, \tag{9}$$

The matrix form of this constraint is:

$$A \cdot x + b = \left(\begin{array}{cccccccc|cccc} 1 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 1 \end{array}\right) \left(\begin{array}{c} k_1 \\ k_2 \\ k_3 \\ k_4 \\ k_5 \\ k_6 \\ k_7 \\ k_8 \\ \hline n_1 \\ n_2 \\ n_3 \\ n_4 \end{array}\right) + \left(\begin{array}{c} -1 \\ -1 \\ -1 \\ -1 \end{array}\right) = 0 \tag{10}$$

Constructing the QUBO matrix

- Combining the action and Grassmann constraint.

$$\chi^2 = x^T W x + p||Ax + b||^2 \tag{11}$$

- p: penalty factor which controls the balance between action and constraint.
- The aim is to find the solution vector x which minimizes χ^2.
- The matrix formulation required by the D-wave API is

$$\chi^2 = x^T Q x + C \tag{12}$$

- The QUBO matrix Q and the constant C is

$$Q = W + p\left(A^T A + diag(2b^T A)\right), \quad C = pb^T b \tag{13}$$

$U(3)$ Theory

- Partition function

$$Z = \sum_{\{k,n,\ell\}}^{GC} \prod_{b=(x,\hat{\mu})} \frac{(3-k_b)!}{3!k_b!} \gamma^{2k_b\delta_{0,\hat{\mu}}} \prod_{x} \frac{3!}{n_x!} (2am_q)^{n_x} \tag{14}$$

- $k_b, n_x \in \{0,1,2,3\}$ and it can be expressed by combining two binary numbers.

$$0 \mapsto (0,0), \qquad 1 \mapsto (0,1), \qquad 2 \mapsto (1,0) \qquad 3 \mapsto (1,1) \tag{15}$$

$$S = x^T W x = (\vec{k}_b^T, \vec{n}_x^T) \begin{pmatrix} (D_{2\times 2})\mathbb{1}_{E\times E} & \mathbb{0}_{2V\times 2E} \\ \mathbb{0}_{2E\times 2V} & (M_{2\times 2})\mathbb{1}_{V\times V} \end{pmatrix} \begin{pmatrix} \vec{k}_b \\ \vec{n}_x \end{pmatrix} \tag{16}$$

$$D_{2\times 2} = \begin{pmatrix} \log(12) - 4\delta_{0,\hat{\mu}} \log(\gamma) & 0 \\ 0 & \log(3) - 2\delta_{0,\hat{\mu}} \log(\gamma) \end{pmatrix} \tag{17}$$

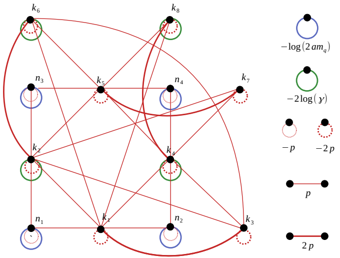
$$M_{2\times 2} = \begin{pmatrix} -2\log(2am_q) + \log(2) & \log(3) \\ 0 & -\log(2am_q) \end{pmatrix} \tag{18}$$

- Grassmann constraint

$$\sum_{\mu=\pm 0,\cdots,\pm d} k_\mu(x) + n_x = 3, \tag{19}$$

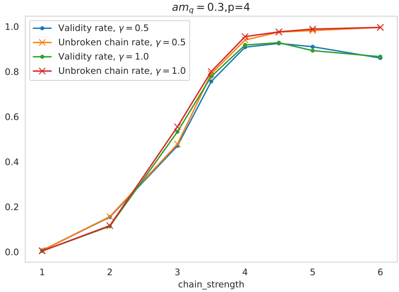
Parameter tuning

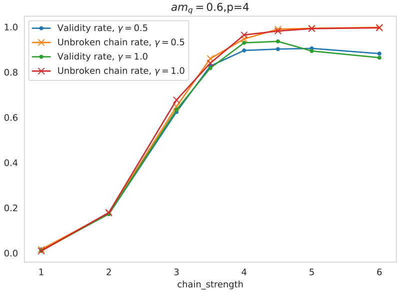
- One of the free parameters is the **chain_strength**.
- It controls the strength of chains used to build physical qubits into logical qubits, ensuring that physical qubits act in unison.
- Our problem is not the same topology as the QPU, so we can't find a one-to-one embedding, so non-trivial **chain_strength** is required to retain logical qubits.



Optimal **chain_strength**

- Validity rate is the number of valid solution vectors which satisfy the constraint over the total number of samples.
- Validity rate is identical to the **unbroken_chain_rate** for small **chain_strength**.
- Validity rate starts deviating from **unbroken_chain_rate** after some specific value of **chain_strength**.





Penalty factor p

- Penalty factor controls the balance between the action and constraint.

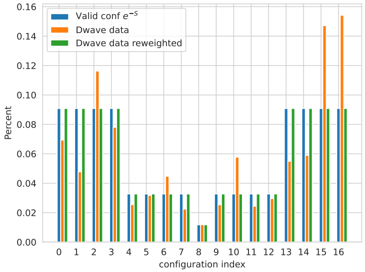
$$Q = W + p\left(A^T A + diag(2b^T A)\right) \tag{20}$$

- if p is small, finding valid solution will be hard.

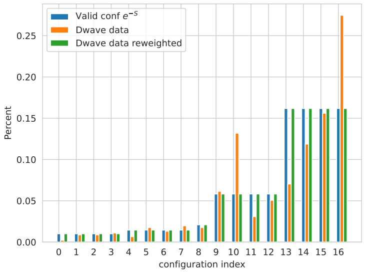
group	lattice	configurations	binary vectors	percentage(%)
U(1)	2×2	17	2^{12}	0.4
U(1)	4×4	41025	2^{48}	1.5×10^{-8}
U(1)	6×6	23079663560	2^{108}	7×10^{-21}
U(1)	$2 \times 2 \times 2$	689	2^{32}	0.00002
U(1)	$2 \times 2 \times 2 \times 2$	1898625	2^{80}	1.6×10^{-16}
U(2)	2×2	135	2^{24}	0.0008
U(3)	2×2	695	2^{24}	0.004
U(3)	$2 \times 2 \times 2$	8750060	2^{64}	4.7×10^{-11}

- If p is too big, the action part of QUBO matrix will be ignored.
- Finding an optimized p is important.

Distribution of valid configurations of $U(1)$ theory



(3) $am_q = 0.3, \gamma = 1$



(4) $am_q = 0.3, \gamma = 0.5$

- D-wave distribution is not exactly same with the ideal distribution.
- D-wave find the important configurations more often. (Importance sampling)
- We can always compute e^{-S} for given configuration.

Observables

- We measure two independent observables the number of monomers $\langle M \rangle$ and the number of temporal dimers $\langle D_t \rangle$.

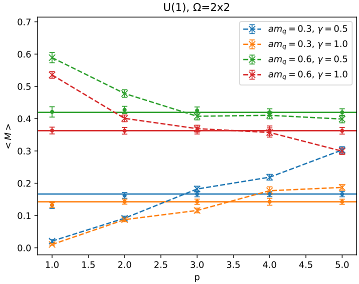
$$M = \sum_{x \in \Omega} n_x, \qquad D_t = \sum_{x \in \Omega} k_0(x) \tag{21}$$

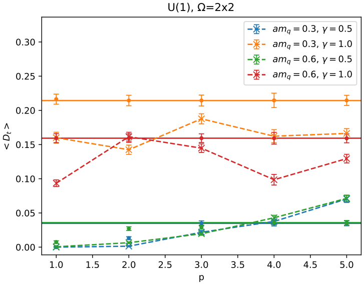
- They are related to chiral condensate and energy density.

$$\begin{aligned} a^{d-1} \langle \bar{\psi}\psi \rangle &= a^{d-1} \frac{T}{V} \frac{\partial \log Z}{\partial m_q} = \frac{1}{\Omega} \frac{1}{2am_q} \langle M \rangle \\ a^d \langle \epsilon \rangle &= -\frac{a^d}{V} \frac{\partial \log Z}{\partial T^{-1}} = \frac{1}{\Omega} \left(\frac{\xi}{\gamma} \frac{d\gamma}{d\xi} \langle 2D_t \rangle - \langle M \rangle \right) = \frac{1}{\Omega} (\langle D_t \rangle - \langle M \rangle) \end{aligned} \tag{22}$$

- Ω: spatial volume
- γ: anisotropic
- $aT = \xi(\gamma)/N_t$, $\xi(\gamma) = \kappa \gamma^2$ at strong coupling

$U(1)$ gauge group on 2×2

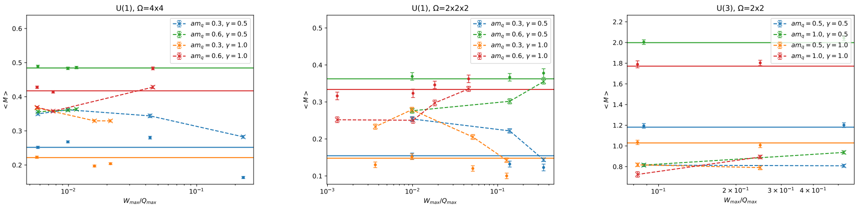




- The data points connected by dashed line are the D-wave raw data.
- For large enough p, D-wave finds all 17 confs, and reweighting method produces the correct distribution.
- if p is very small, the action in the QUBO matrix is emphasized. Hence, D-wave samples the distribution very near to the global minimum.

The number of monomers (Chiral condensate)

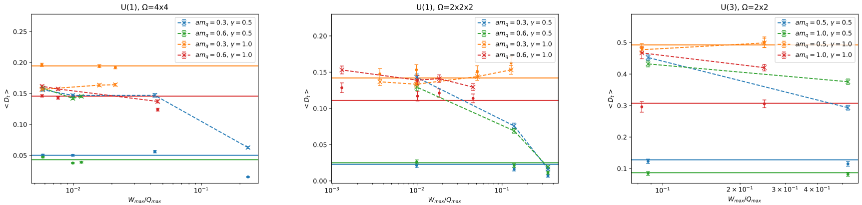
$$\langle\bar{\psi}\psi\rangle = \frac{1}{\Omega}\frac{1}{2m_q}\langle M\rangle \tag{23}$$



- In our choice of physical parameter, $Q_{max} = 2p$. So we use the ratio of W_{max}/Q_{max} for the tuning parameter.
- Where $W_{max}/Q_{max} \approx 0.01$, reweighted results agree with the exact solutions.
- In the case of $U(1)$ on 4×4 lattice, D-wave finds about $(700 - 1800)$ valid configurations in 41025 total which is $1.7 - 4.3\%$.

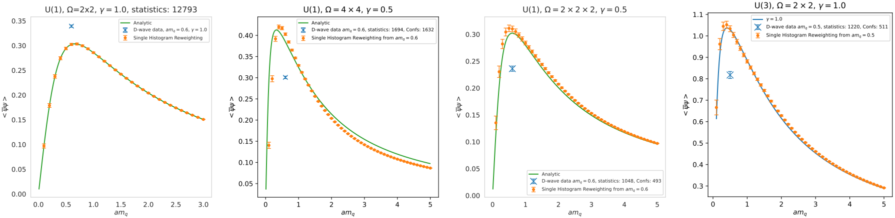
The number of temporal dimers

$$\langle \epsilon \rangle = \frac{1}{\Omega}(\langle D_t \rangle - \langle M \rangle) \quad (24)$$



- Another observable also shows good agreement with the analytic solutions where $W_{max}/Q_{max} \approx 0.01$.

Reweighting to other physical parameters



- $U(1)$ on 2×2: Since we have 17 all confs, no restriction for reweighting range.
- $U(1)$ on 4×4: $1.7 - 4.3\%$ of valid confs, but 16 monomer conf is missing. Reweighting does not describe well the large quark mass region.
- $U(1)$ on $2 \times 2 \times 2$ and $U(3)$ on 2×2: we have about 70% of valid confs. reweighting works for a much longer range.
- The errorbars are purely statistical.

Conclusion

- We have demonstrated that lattice gauge theory in the strong coupling limit on D-wave quantum annealer.
- As a proof of principle, $U(1)$ and $U(3)$ on various small volumes are successfully simulated by the D-Wave quantum annealer.
- In particular, we have demonstrated that importance sampling is feasible on the quantum annealer.
- The accuracy is greatly enhanced by the histogram reweighting method.
- In that case, the tuning of D-Wave parameters is less crucial.
- As introduce the static baryon, $SU(N_c)$ gauge group can be simulated on D-wave.
- For larger volume, we propose an iterative scheme by decomposing local updates on even and odd sites to deal with a more realistic. It need hybrid classical/quantum computing.
- We expect the quantum advantage over the worm algorithm by comparing how fast D-wave can reach equilibrium and how shorter the autocorrelation is on large volumes and low temperatures