QMC of everything: A universal algorithm for simulating arbitrary quantum many-body systems

Itay Hen

Information Sciences Institute, USC

Based partly on: arXiv:2307.06503, arXiv:2006.02539

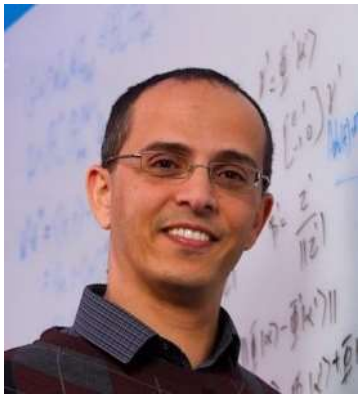
IPAM workshop: Many-body Quantum Systems via Classical and Quantum Computation

November 7, 2023

Main collaborators and funding

Lev Barash

Arman Babakhani



Amir Kalev

- Funding:

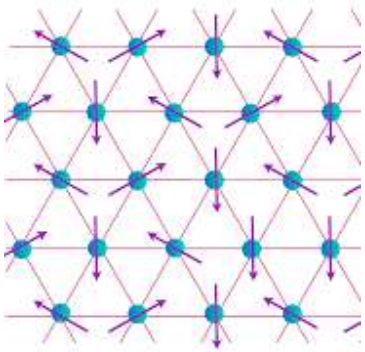
Outline

- A brief intro to quantum Monte Carlo (QMC) simulations
- Permutation Matrix Representation QMC
- QMC of everything: Provably ergodic Markov chains for arbitrary spin ½ systems
- Generalization to higher-spin particles, fermions and bosons
- PMR in quantum computing
- Summary and conclusions

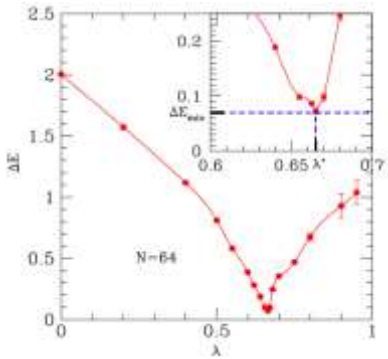
A brief intro to QMC simulations

Motivation

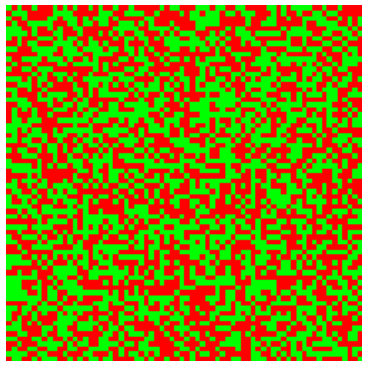
- Generally, large-scale quantum many-body systems cannot be studied analytically. Too many degrees of freedom.
- Quantum Monte Carlo (QMC) sampling is often the only approach available to study these without approximations.
- Still, QMC is known to be inefficient in very many cases.



sign problem
(negative weights)



small energy gaps
(phase transitions)



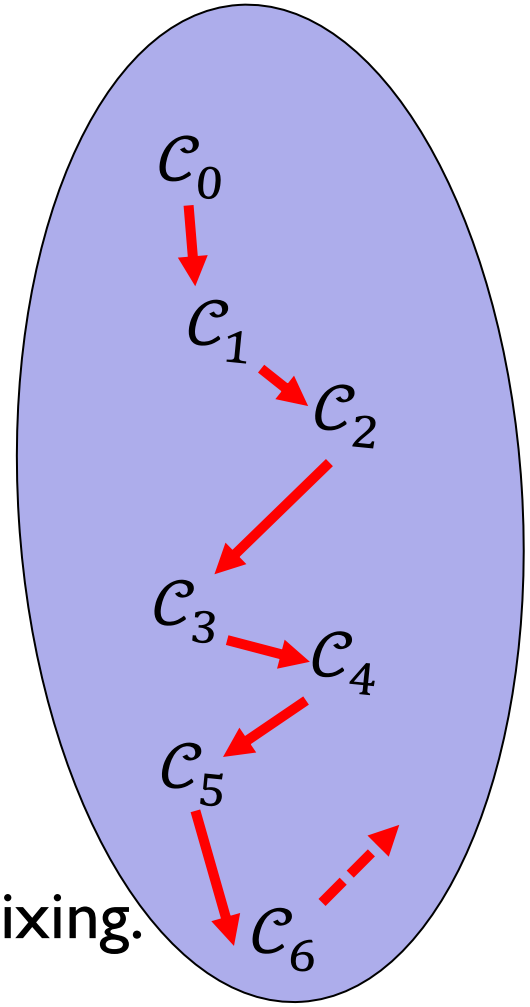
classical glassiness

- Also: for each model, one usually needs a specially tailored algorithm.

What is QMC?

- Goal is to write $Z = Tr\left[e^{-\beta H}\right]$ as a sum of *non-negative* weights $W_{\mathcal{C}}$ for configurations $\mathcal{C}$: $Z = \sum_{\mathcal{C}} W_{\mathcal{C}}$.
- Similarly write observables as $\langle A \rangle = \frac{1}{Z} Tr\left[A\, e^{-\beta H}\right]$ as $\frac{1}{Z}\sum_{\mathcal{C}} A_{\mathcal{C}} W_{\mathcal{C}}$.
- Use Markov chain to sample each $\mathcal{C}$ with probability $p_{\mathcal{C}} = \frac{W_{\mathcal{C}}}{Z}$.
- Collect statistics about A by measuring $A_{\mathcal{C}}$. Visit important configurations more often.
- Hope for (I) no sign problem and (II) rapid mixing.

Markov Chain Monte Carlo (MCMC) in QMC configuration space

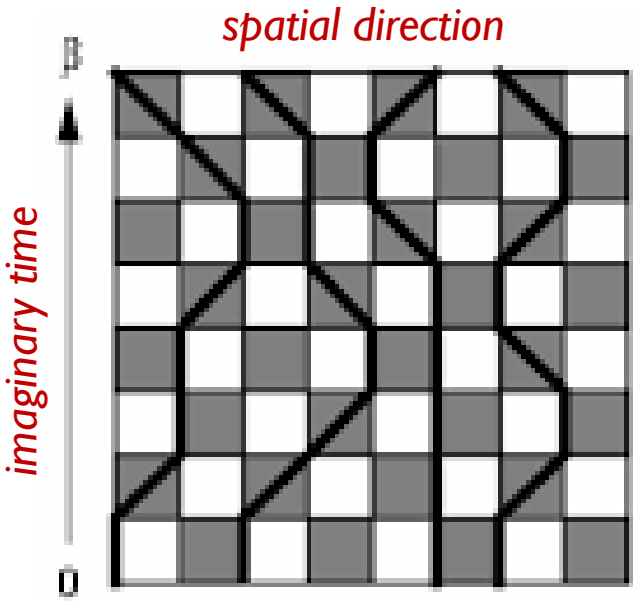


Standard QMC approaches

- Path integral Monte Carlo methods are prone to Trotterization errors.

$$e^{-\beta H} = \left(e^{-\Delta t H}\right)^L \approx \left(e^{-\Delta t A} e^{-\Delta t B}\right)^L \quad \text{but} \quad e^{\Delta t\, A} e^{\Delta t\, B} \neq e^{\Delta t\,(A+B)}$$

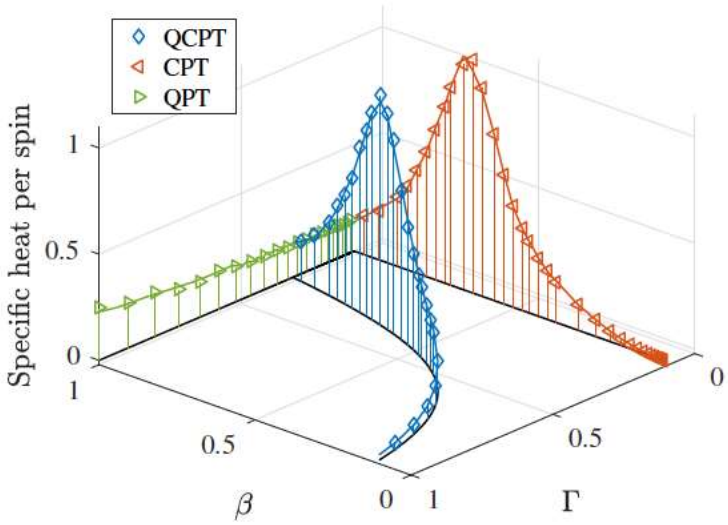
- Other schemes which are immune to Trotterization errors:
  - Continuous-time path integral Monte Carlo [pioneered by Prokof'ev et al].
  - Stochastic series expansion (SSE) [pioneered by Sandvik].
- These have their own "issues".
- Common annoyance: need to design update rules on a model-by-model basis.



Permutation Matrix Representation Quantum Monte Carlo

Permutation matrix representation

- PMR-QMC is a novel series expansion of the quantum partition function (T. Albash, G. Wagenbreth, IH, PRE, 2017).
- It is *parameter-free* and *Trotter-error free*.
- It is *universal* and *abstract*: it applies essentially to any model and does not require model-specific adjustments.
- It does *not* solve the sign problem but can be used to solve or mitigate it in certain cases.
- Becomes a thermal classical simulation in the classical limit.

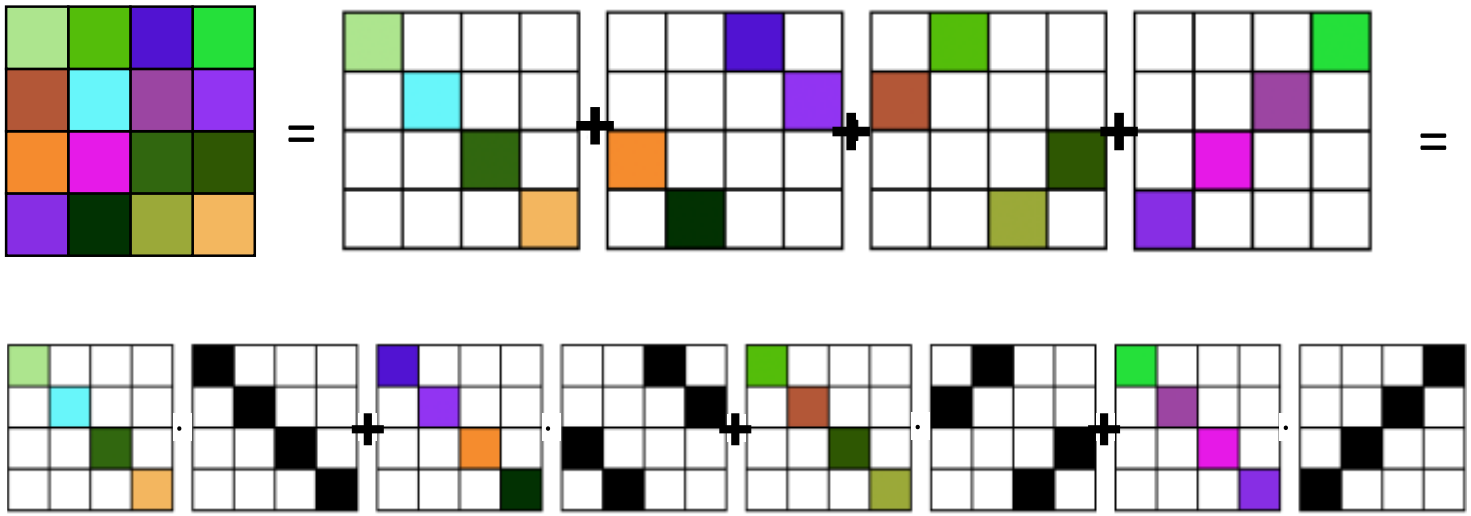


Permutation matrix representation

- Choose a basis $\{|z\rangle\}$ (call it the "computational" basis).
- Cast Hamiltonian as a sum of M monomial (generalized permutation) operators $\tilde{P}_j = D_j P_j$ (always possible and easy to do):

$$H = \sum_{j=0}^{M} \tilde{P}_j = \sum_{j=0}^{M} D_j P_j = D_0 + \sum_{j=1}^{M} D_j P_j$$

Permutation matrix representation of the Hamiltonian



Permutation matrix representation

$$H = \sum_{j=0}^{M} \tilde{P}_j = \sum_{j=0}^{M} D_j P_j = D_0 + \sum_{j=1}^{M} D_j P_j$$

- Sum of products of diagonal (non-local) operators D_j and permutation operators P_j, where $P_j|z\rangle = |z'\rangle (\neq |z\rangle)$.
- $P_0 = \mathbb{I}$ making D_0, the "classical"/diagonal Hamiltonian.
- Rest of permutations P_j have *no fixed points*.
- E.g., for spin models, P_j's would be $\mathbb{I}, X_1, X_2, \ldots, X_1X_2, \ldots$
- The D_j's are generally non-local diagonal operators.

The off-diagonal series expansion

- Expand the partition function $\mathrm{Z} = Tr\left[e^{-\beta H}\right]$ in a Taylor series substituting $Tr[\cdot] = \sum\langle z|\cdot|z\rangle$:

$$\mathrm{Z} = \sum \left\langle z \middle| \sum \frac{(-\beta)^n}{n!} \left(D_0 + \sum D_j P_j\right)^n \middle| z \right\rangle.$$

- Expanding, we get

$$\mathrm{Z} = \sum \sum \frac{(-\beta)^n}{n!} \left\langle z \middle| \begin{array}{c} sum\ of\ all\ possible\ products \\ of\ length\ n\ of\ D_0\ and\ D_j P_j \end{array} \middle| z \right\rangle.$$

T. Albash, G. Wagenbreth and I. Hen, "Off-Diagonal Expansion Quantum Monte Carlo", Phys. Rev. E **96**, 063309 (2017).

The off-diagonal series expansion

- Take one such product for example:

$$\langle z|D_0 \cdot D_0 \cdot D_1P_1 \cdot D_0 \cdot D_1P_1 \cdot D_0 \cdot D_0 \cdot D_2P_2 \cdot D_0|z\rangle$$

$$= \quad \downarrow |z\rangle$$

$$\langle z|D_0 \cdot D_0 \cdot D_1P_1 \cdot D_0 \cdot D_1P_1 \cdot D_0 \cdot D_0 \cdot D_2P_2|z\rangle \cdot E(z)$$

$$= \quad \downarrow |z'\rangle$$

$$\langle z|D_0 \cdot D_0 \cdot D_1P_1 \cdot D_0 \cdot D_1P_1 \cdot D_0 \cdot D_0\, P_2|z\rangle \cdot d_2(z')E(z)$$

$$\downarrow |z'\rangle$$

$$\langle z|D_0 \cdot D_0 \cdot D_1P_1 \cdot H_c \cdot D_1P_1P_2|z\rangle \cdot E^2(z')d_2(z')E(z)$$

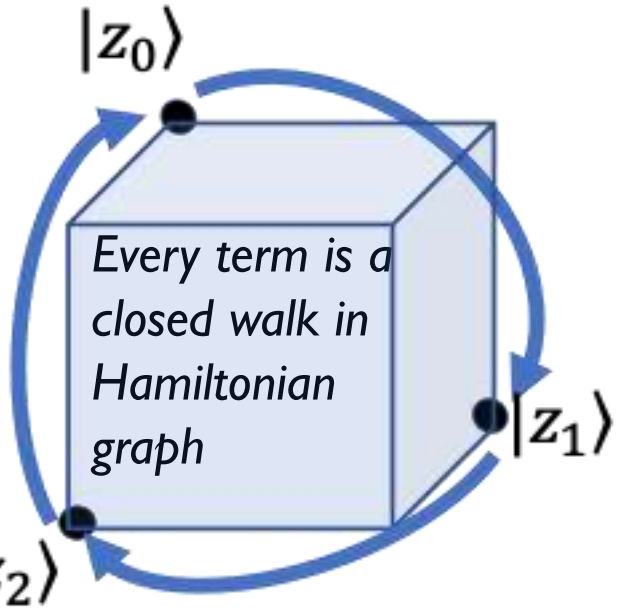
$$\downarrow |z'''\rangle \qquad \vdots$$

$$\underbrace{\langle z|P_1P_1P_2|z\rangle}_{\text{This is the "off-diagonal core" consisting of a unique product of permutation matrices. Evaluates either to } 0 \text{ or to } 1.} \cdot E^2(z''')d_1(z''')E(z'')d_1(z'')E^2(z')d_2(z')E(z)$$

The off-diagonal series expansion

- There is an infinite number of terms that share the same off-diagonal core $\left\langle z \middle| P_{i_q} \dots P_{i_2} P_{i_1} \middle| z \right\rangle$. We group them together.
- Each pair $\left\{ |z\rangle, S_{\vec{i}_q} = P_{i_q} \dots P_{i_2} P_{i_1} \right\}$ defines a closed walk on the Hamiltonian graph (where basis states are nodes, off-diagonal elements are edges):

$$|z\rangle = |z_0\rangle \xrightarrow{P_{i_1}} |z_1\rangle \xrightarrow{P_{i_2}} \cdots \xrightarrow{P_{i_q}} |z_q\rangle = |z\rangle$$



- This path "induces" a sequence of classical energies $D_0|z\rangle = E_z|z\rangle$: $E_{z_0}, E_{z_1}, \dots, E_{z_q}$.
- And accompanying $d_{i_1}(z_1), d_{i_2}(z_2), \dots$

Grouping with divided differences

- Summing the contributions of all terms that share the same off-diagonal core, we get the *infinite sum*

$$\left(\sum_{n=q}^{\infty}\frac{\beta^n(-1)^{n-q}}{n!}\sum_{\sum k_i=n-q}E_c^{k_0}(z_0)\cdot\ldots\cdot E_c^{k_q}(z_q)\right)$$

- As it turns out, this sum can be neatly evaluated to the easily computable

$$e^{-\beta\left[E_{z_0},E_{z_1},\ldots,E_{z_q}\right]}$$

This is the *divided difference* of the Boltzmann factors of *sequences* (multi-sets) of intermediate classical energies along the imaginary time direction.

Digression: divided differences

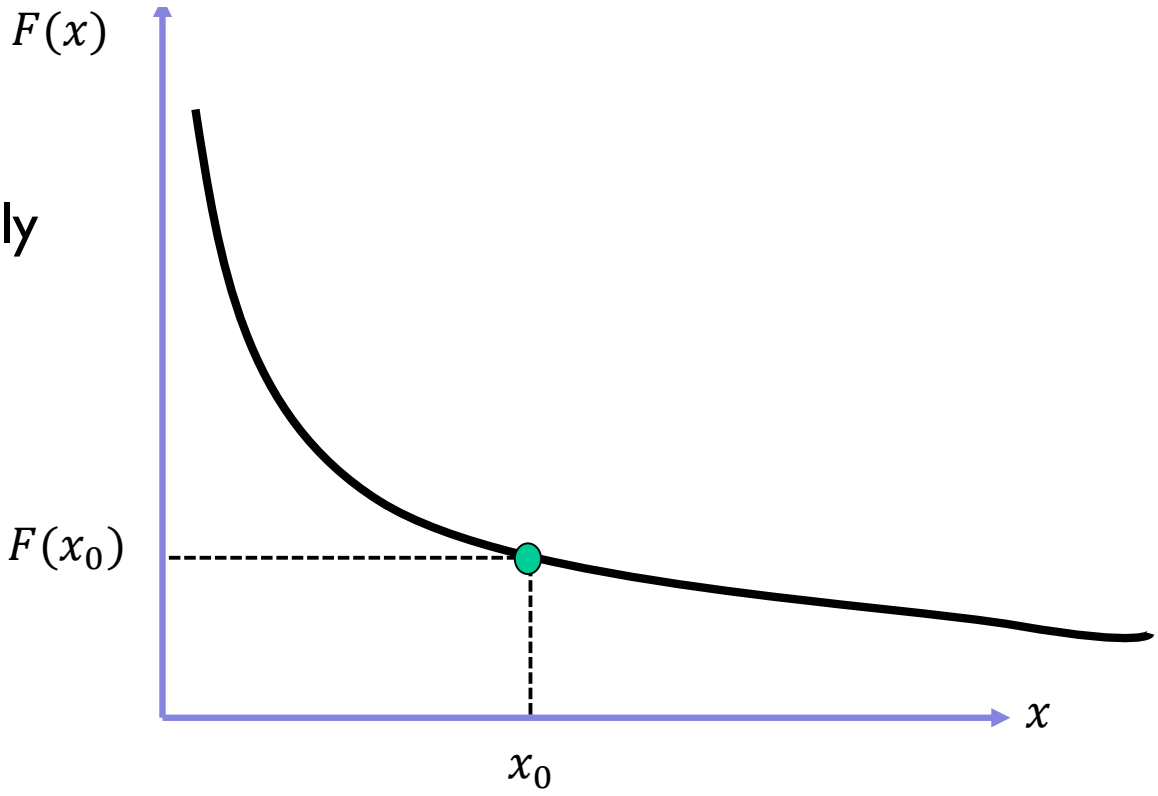
Digression: divided differences

- The divided differences of a function $F(x)$ with respect to the input multi-set $[x_0, \ldots x_q]$ is given by:

$$F([x_0, \ldots, x_q]) \equiv \sum_{j=0}^{q} \frac{F(x_j)}{\prod_{k \neq j}(x_j - x_k)}$$

- The divided difference of a function with an input multi-set of size one, is simply

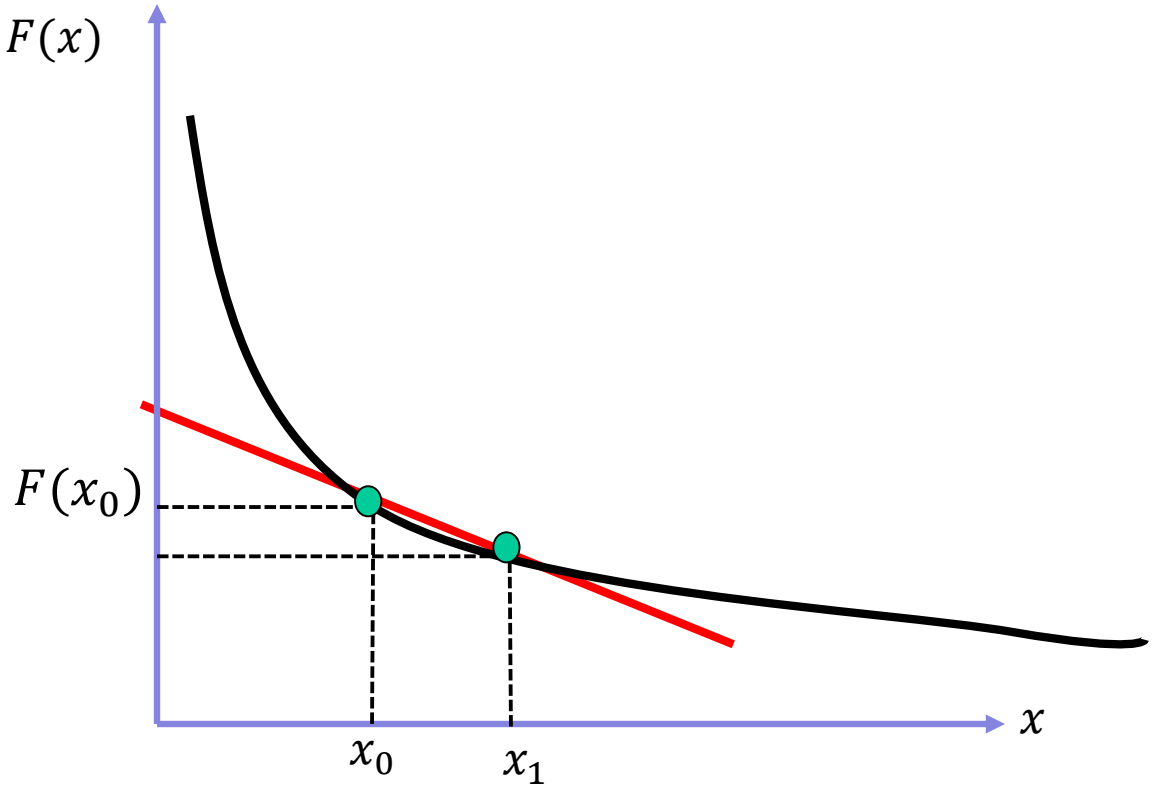
$$F[x_0] = F(x_0)$$



Digression: divided differences

- The divided differences of a function taking as input a multi set with two elements is:

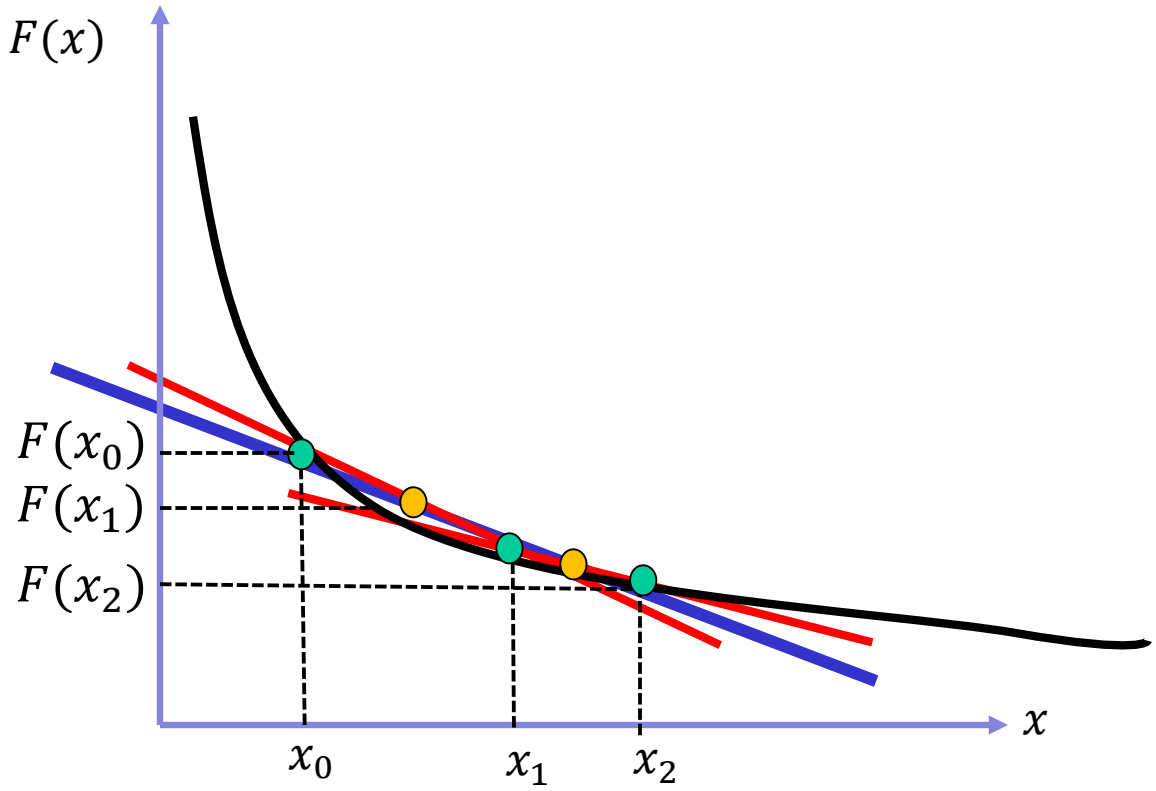
$$F[x_0, x_1] = \frac{F(x_1) - F(x_0)}{x_1 - x_0} \approx F'(\xi)$$



Digression: divided differences

- The divided differences of a function taking as input a multi set with three elements is:

$$F[x_0, x_1, x_2] = \frac{F[x_0, x_1] - F[x_1, x_2]}{x_0 - x_2} \approx \frac{1}{2} F''(\xi)$$



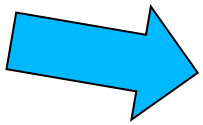
Digression: divided differences

- In the general case, the evaluation of the divided differences of a function with $q+1$ inputs

$$F([x_0,\ldots,x_q]) \equiv \sum_{j=0}^{q} \frac{F(x_j)}{\prod_{k\neq j}(x_j - x_k)}$$

is done via the recursion relation:

$$F[x_0,\ldots,x_q] = \frac{F[x_0,\ldots,x_{q-1}] - F[x_1,\ldots,x_q]}{x_0 - x_q}$$



- Also: $$F[x_0,\ldots,x_q] = \frac{F^{(n)}(\xi)}{n!}$$

- The computational cost of calculating divided differences scales as q^2 in the worst case.

The partition function

- Grouping all terms together, we end up with the final form

$$\mathrm{Z} = \sum_{z, S_{\vec{i}_q} = \mathbb{I}} D_{(z, S_{\vec{i}_q})} \, e^{-\beta \left[E_{z_0}, E_{z_1}, \ldots, E_{z_q} \right]}$$

Z is a sum over all basis states $\{|z\rangle\}$ and all the combinations of products of P_j operators

$D_{(z, S_{\vec{i}_q})}$ is a product of q off−diagonal matrix elements

This is the divided-difference exponential whose inputs are the classical energies along the walk. Its sign is $(-1)^q$.

$S_{\vec{i}_q} = P_{i_q} \ldots P_{i_2} P_{i_1}$ (with q the size of the sequence) that evaluate to $S_{\vec{i}_q} = \mathbb{I}$.

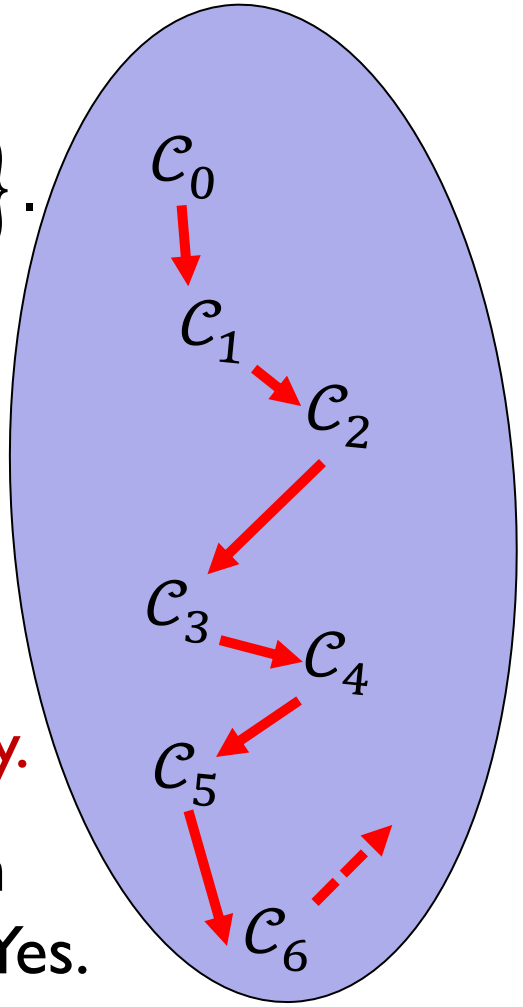
- Each summand is a "generalized Boltzmann weight" $W = D_{(z, S_q)} e^{-\beta \left[E_{z_0}, E_{z_1}, \ldots, E_{z_q} \right]}$.

Provably ergodic QMC for arbitrary spin ½ systems

Guaranteeing ergodicity

- Recall we have a Markov chain to sample each $\mathcal{C}$ with probability $p_{\mathcal{C}} = \frac{W_{\mathcal{C}}}{Z}$.
- Here, $\mathcal{C}$ is a pair $\left\{|z\rangle, S_{\vec{i}_q} = P_{i_q} \ldots P_{i_2} P_{i_1} (= \mathbb{I})\right\}$.
- Covering the $|z\rangle$ space is easy – "classical moves", e.g., spin flips.
- We must also make sure the process is ergodic for any set of $\{P_i\}$ appearing in H.
- This becomes an interesting problem in theory.
- Can we produce a finite set of moves that can generate all sequences $S_{\vec{i}_q}$ that evaluate to $\mathbb{I}$? Yes.

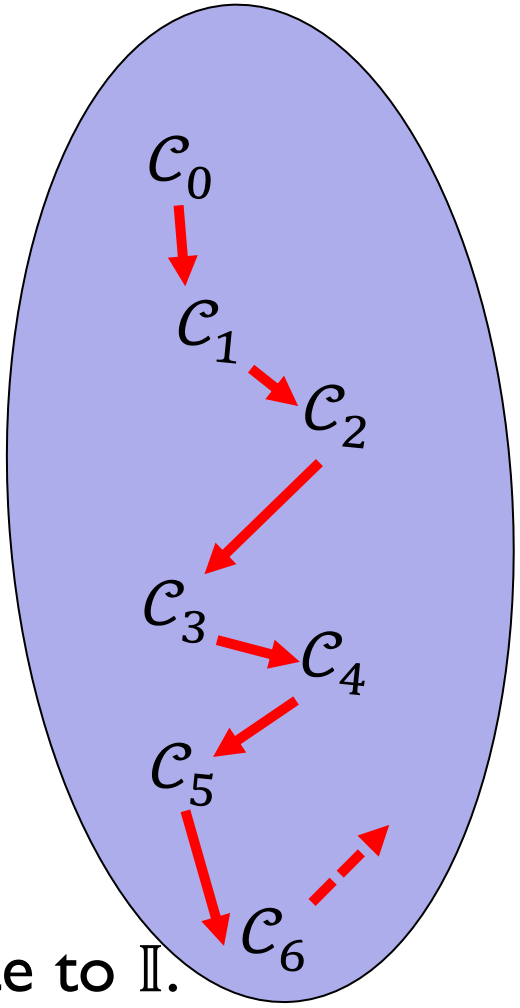
Markov Chain Monte Carlo (MCMC) in QMC configuration space



Guaranteeing ergodicity

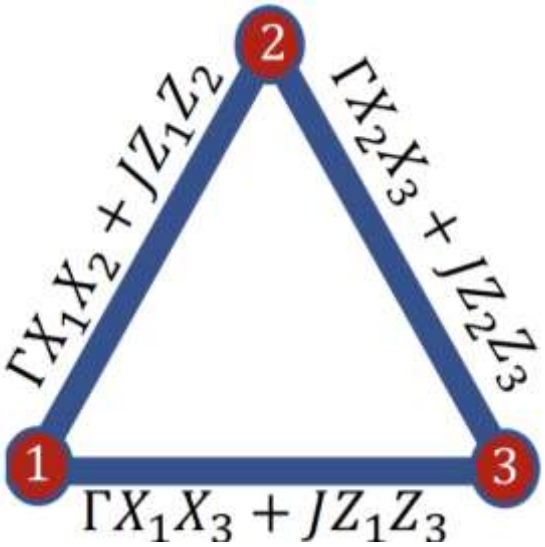
- Take for example, spin ½ systems. Here the $\{P_i\}$ are Pauli-X strings.
- As group elements, they obey $P_i = P_i^{-1}, [P_i, P_j] = 0$.
- We were able to prove that all $S_{\vec{i}_q}$ can be generated with:
  - Operator swapping, e.g., $P_3P_1P_2P_3 \rightarrow P_3P_2P_1P_3$
  - Pair insertion/deletion, e.g., $P_2P_1 \rightarrow P_2P_3P_3P_1$
  - Insertion and deletion of the "fundamental cycles" of the model, e.g., $P_1P_2 \rightarrow P_2P_4P_5P_6P_1$
- These are products of distinct P_i's that evaluate to $\mathbb{I}$.

Markov Chain Monte Carlo (MCMC) in QMC configuration space



Guaranteeing ergodicity

- On a triangular lattice, we may have:
  $P_1 = X_2X_3,\ P_2 = X_1X_3$ and $P_3 = X_1X_2$.
- Here a fundamental cycle is $P_1P_2P_3 = \mathbb{I}$.
- Why? Denote:
  $P_1 \equiv [0,1,1], P_2 \equiv [1,0,1], P_3 \equiv [1,1,0]$.
- Products of permutations correspond to mod-2 addition of these vectors.

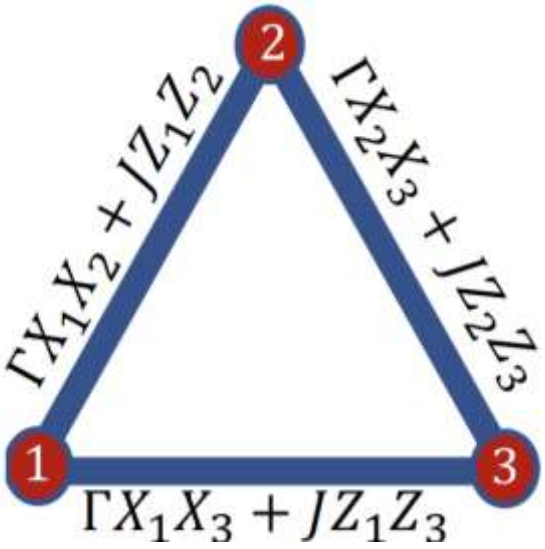


XY model on the triangular lattice.

- The fundamental cycles corresponds to the null-space vectors of the P matrices: $[0,1,1] \oplus [1,0,1] \oplus [1,1,0] = [0,0,0]$.
- All fundamental cycles may be found automatically and efficiently!

Guaranteeing ergodicity

- On a triangular lattice, we may have:
  $P_1 = X_2X_3,\ P_2 = X_1X_3$ and $P_3 = X_1X_2$.
- Here a fundamental cycle is $P_1P_2P_3 = \mathbb{I}$.
- E.g., the sequence $P_2P_1P_2P_1P_3P_1P_2$.
- Via local swaps: $P_1P_1P_1P_2P_2P_2P_3$.
- Pair deletions: $P_1P_2P_3$.
- Fundamental cycle deletion: $\mathbb{I}$.
- This leads to update moves with verified ergodicity for any and all spin ½ systems!

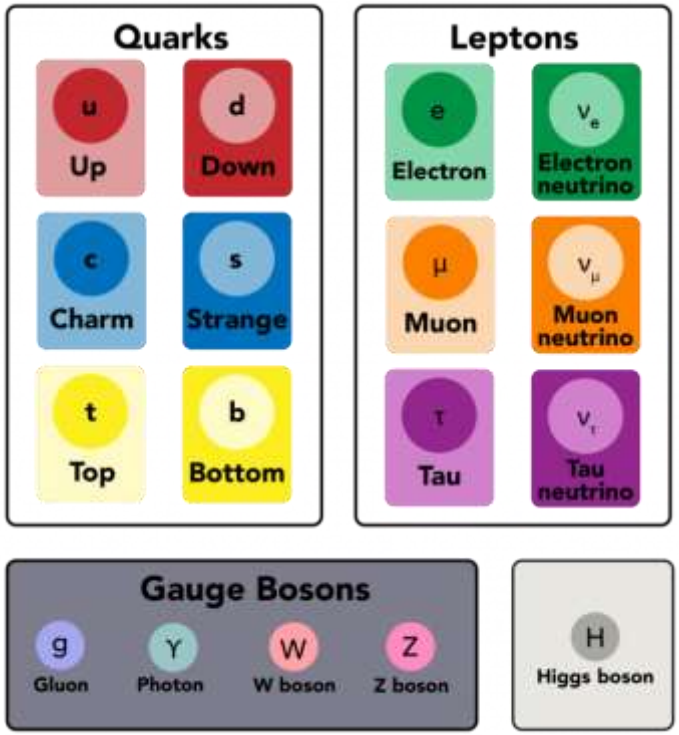


XY model on the triangular lattice.

Generalization to higher-spin particles, fermions and bosons

Other types of systems?

- What about fermions?
  Can be mapped via a Jordan-Wigner transformation to spin ½ particles.
- What about higher spin particles?
  Slightly less trivial group structure: $P_i = P_i^{-1} \rightarrow P_i^{2s} = P_i^{-1}$ but just as doable. mod $2 \rightarrow$ mod $2s + 1$.
- What about bosons?
  Can be treated as infinitely-high-spin particles.



Particles of the standard model

- We thus have a method that automatically simulates essentially all condensed matter models.

PMR in quantum computing

Same expansion but for quantum circuits

- In the context of quantum Hamiltonian simulation algorithms:

$$e^{-iHt}|z\rangle = \sum_{q=0}^{\infty} \sum_{\mathbf{i}_q} e^{-it[E_z,\ldots,E_{z_q}]} D_{\mathbf{i}_q} P_{\mathbf{i}_q} |z\rangle$$

Sum over all the combinations of products of P_j operators

$P_{\boldsymbol{i}_q} = P_{i_q} \ldots P_{i_2} P_{i_1}$ (with q the size of the sequence)

This is the divided-difference exponential whose inputs are the classical energies along the path.

$D_{\boldsymbol{i}_q}$ is the product of all $d_{i_j}(z_j)$ along path. One for each permutation operator.

- $\boldsymbol{i}_q = (i_1, i_2, \ldots, i_q)$ is a multi-index. Each index $i_j = 1 \ldots M$ picks an off-diagonal permutation.

Off-diagonal time evolution

- Consider from now on the short-time evolution operator

$$U = e^{-iH\Delta t}$$

- Using the off-diagonal series expansion we can write U as:

$$U = U\sum_z |z\rangle\langle z| = \sum_z U|z\rangle\langle z| = \sum_z V_z|z\rangle\langle z|$$

- Explicitly:

$$U = \sum_z e^{-i\Delta t E_z} \sum_{q=0}^{\infty}\sum_{\mathbf{i}_q} e^{-i\Delta t[\Delta_z,\ldots,\Delta_{z_q}]} D_{\mathbf{i}_q} P_{\mathbf{i}_q}|z\rangle\langle z|$$

$$= \underbrace{\left(\sum_z \sum_{q=0}^{\infty}\sum_{\mathbf{i}_q} e^{-i\Delta t[\Delta_z,\ldots,\Delta_{z_q}]} D_{\mathbf{i}_q} P_{\mathbf{i}_q}|z\rangle\langle z|\right)}_{\text{The off-diagonal evolution } U_{OD}} e^{-i\Delta t D_0}$$

The circuit

- Write the time-evolution operator as a product of short-time evolution operators.

$$e^{-iHt} = \underbrace{e^{-iH\Delta t} e^{-iH\Delta t} \dots e^{-iH\Delta t}}_{r \text{ times}}$$

- Time step is set to $\Delta t = ln2 / \sum_{i=1} \Gamma_i$.
- Important: $r = t/\Delta t$, number of repetitions is proportional to the "dimensionless time" $T = t \sum_{i=1} \Gamma_i$ where $\Gamma_i = |D_i|$ are the norms of the non-local off-diagonal operators.
- Diagonal part has been "integrated out".

Comparison with Taylor-based LCU

- Comparison with truncated Taylor series method is warranted.
- Gate cost of both algorithms scales linearly with their respective "dimensionless times" T. Qubit cost scales with their respective expansion orders $Q \sim \log T$ and respective number of operators M to which the Hamiltonian is decomposed.
- These numbers are very different.
- Examples:

Hamiltonian	$H = \sum_{ij} J_{ij} Z_i Z_j$	
Method	this paper	Taylor series LCU [1]
No. of LCU unitaries	0	N^2
Dimensionless time (T)	0	$t \sum_{ij} \lvert J_{ij} \rvert$
Comments	H is diagonal	-

Hamiltonian	$H = \sum_{ij} J_{ij} Z_i Z_j + \sum_{ij} \tilde{J}_{ij} Z_i X_j$	
Method	this paper	Taylor series LCU [1]
No. of LCU unitaries	$N + 1$	$2N^2$
Dimensionless time (T)	$t \sum_j \lvert \sum_i \tilde{J}_{ij} \rvert$	$t \sum_{ij} (\lvert J_{ij} \rvert + \lvert \tilde{J}_{ij} \rvert)$
Comments	$D_0 = \sum_{ij} J_{ij} Z_i Z_j$ $D_j = \sum_i J_{ij} Z_i$	-

Summary and conclusions

Summary and conclusions

- We developed a very powerful classical simulation tool: one algorithm to fit them all…
- We have not solved the sign problem (but are working on it…).
- Haven't discussed measurements: We can show (paper in preparation) that essentially any conceivable physical operator may be measured, including integrated observables.
- Method can be successfully ported to quantum computers.

QMC of everything: A universal algorithm for simulating arbitrary quantum many-body systems

Itay Hen

Information Sciences Institute, USC

Based partly on: arXiv:2307.06503, arXiv:2006.02539

IPAM workshop: Many-body Quantum Systems via Classical and Quantum Computation

November 7, 2023