Classical Computer Science and Quantum Computing: High Performance Computing and Quantum Simulation

Edgar Solomonik

Department of Computer Science, University of Illinois at Urbana-Champaign

L P N A @CS@Illinois

NSF Workshop: Enabling Quantum Leap

Quantum algorithms for quantum chemistry and materials

Alexandria, VA

January 23, 2019

Outline

1. Quantum Circuit Emulation
2. Quantum Chemistry with HPC
3. Tensor Networks
4. Broadening Computer Science Participation and Education in Quantum

Applications of Quantum Circuit Emulation

- The best techniques for classical emulation of a general quantum circuit have exponential cost in the number of qubits
- However, HPC resources enable emulation of general quantum circuits with roughly 50-qubits, comparable to NISQ devices
- Quantum circuit emulation is useful in the near term as it enables
  - small-scale quantum algorithm testing
  - modelling effects of noise in quantum circuits
  - verification of NISQ-scale quantum circuits (e.g. for Google's random circuit sampling scheme[1])
  - development of methods for efficient approximation of specific quantum circuits (on specific sets of inputs)
  - tuning of hybrid quantum-classical algorithms (e.g. variational quantum eigensolver for quantum chemistry[2])

[1] Bouland, Fefferman, Nirkhe, and Vazirani arXiv:1803.04402

[2] Colless et al. Physical Review X 8.1 (2018).

Emulation of Quantum Gates and Quantum Circuits

- Consider an n-qubit quantum state

$$|\psi\rangle = \sum_{\boldsymbol{i}\in\{0,1\}^n} t^{\psi}_{\boldsymbol{i}} |i_1 \cdots i_n\rangle \text{ with } t_{\boldsymbol{i}} \in \mathbb{C}$$

- Quantum circuits generally consist of 1-qubit and 2-qubit gates

$$|\phi\rangle = \boldsymbol{U}^{(s)}|\psi\rangle \Rightarrow t^{\phi}_{i_1\cdots i_n} = \sum_{j_s=0}^{1} u^{(s)}_{i_s j_s} t^{\psi}_{i_1\cdots i_{s-1} j_s i_{s+1}\cdots i_n}$$

$$|\phi\rangle = \boldsymbol{U}^{(s,t)}|\psi\rangle \Rightarrow t^{\phi}_{i_1\cdots i_n} = \sum_{j_s=0}^{1} \sum_{j_t=0}^{1} u^{(s)}_{i_s i_t j_s j_t} t^{\psi}_{i_1\cdots i_{s-1} j_s i_{s+1}\cdots i_{t-1} j_t i_{t+1}\cdots i_n}$$

- A quantum gate can be emulated as an $O(2^n)$-cost tensor contraction
- An n-qubit quantum circuit with depth D and $O(nD)$ gates can be simulated classically with $O(nD2^n)$ cost and $O(2^n)$ storage

- Can improve cost, memory footprint, and parallelizability of emulation
- Subsets of the circuit that work on independent sets of qubits commute with one another
- Improving storage overhead is possible via tensor slicing
  - For example, if part of a circuit $\boldsymbol{U}$ does not operate on the first qubit, we can compute $|\phi\rangle = \boldsymbol{U}|\psi\rangle$ by computing in sequence

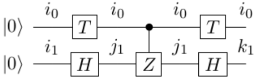
$$t^{\phi}_{0i_2\cdots i_n} = \sum_{j_2\cdots j_n=0}^{1} u_{i_2\cdots i_n j_2\cdots j_n} t^{\psi}_{0j_2\cdots j_n}$$

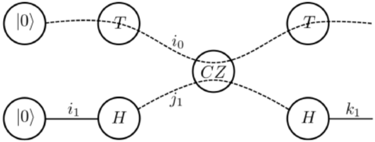
$$t^{\phi}_{1i_2\cdots i_n} = \sum_{j_2\cdots j_n=0}^{1} u_{i_2\cdots i_n j_2\cdots j_n} t^{\psi}_{1j_2\cdots j_n}$$

expanding $u_{i_2\cdots i_n j_2\cdots j_n}$ appropriately in terms of gates, e.g. if it consists only of $n-1$ single qubit gates

$$u_{i_2\cdots i_n j_2\cdots j_n} = \prod_{k=2}^{n} u^{(k)}_{i_k j_k}$$

- Naive quantum circuit emulation has a low arithmetic intensity (flop-byte ratio), requiring $O(2^n)$ memory traffic for $O(2^n)$ floating point operations
- Tensor slicing can improve this by applying multiple gates to a subtensor that fits into fast/local memory
- Alternatively, can use gate aggregation, combining gates to perform fewer reads/writes of the amplitudes $\boldsymbol{T}^{\psi}$ and $\boldsymbol{T}^{\phi}$ but increasing cost
- These and other optimizations can be expressed via a tensor network representation of a quantum circuit[3,4]





[3] Markov and Shi SIAM JC 2007
[4] Pednault et al. arXiv:1710.05867

Specialized Quantum Circuit Emulation

- If only a single amplitude (element of $\boldsymbol{T}^{\psi}$ of $|\psi\rangle = \boldsymbol{U}|0\rangle$) is desired, the Feynman algorithm can be used with $O(nD)$ space and $O(D4^n)$ cost.
- Further, for any $k \leq n$, computation of a single amplitude of $\boldsymbol{U}|0\rangle$ is possible with cost $O(n2^nD^k)$ and $O(2^{n-k}\log D)$ memory[5]
- Explicit calculation of amplitudes is strong circuit simulation
  - aforementioned methods can be classified as monotone simulators
  - approximate monotone simulation of general circuits has $\Omega(2^n)$ cost[6]
- Weak simulation, which samples the output distribution $\boldsymbol{U}|0\rangle$ of a circuit $\boldsymbol{U}$, can be more efficient
- Any quantum algorithm can be expressed with Clifford gates and t phase-shift 1-qubit gates, with Clifford gates being cheap to emulate
  - For $t = 0$, strong simulation has polynomial cost Gottesman-Knill theorem
  - Strong simulation has $O(2^{t/2})$ cost and weak has $O(2^{0.23t})$ cost[7]

[5] Aaronson and Chen arXiv:1612.05903

[6] Huang, Newman, and Szegedy arXiv:1804.01368

[7] Bravyi and Gosset arXiv:1601.07601

Simulation of Noisy Circuits

- NISQ devices are expected to suffer from the effect of noise, especially with increasing circuit depth
- Quantum error correction can be used to protect from noise, but requires many physical qubits per logical qubit[8]
- The presence of noise provides possibility of more efficient simulation (e.g. under a uniform noise rate, generic random quantum circuits can be efficiently simulated classically by tensor networks[9])
- One alternative avenue to error correction is the development of noise resilient algorithms, which may be viable for physical simulation[10]

[8] Terhal arXiv:1302.3428

[9] Gao and Duan arXiv:1810.03176

[10] Isaac Kim arXiv:1703.0003

Tensor Networks for Quantum Chemistry

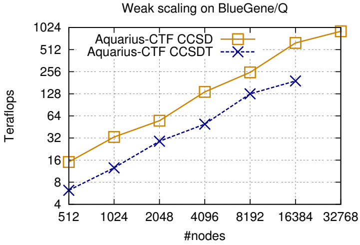
- Classical simulation of quantum chemistry involves many of the same challenges as quantum circuit emulation
- High-accuracy quantum chemistry requires approximation of quantum states/wavefunctions
- Memory footprint is of primary importance, leading to use of similar techniques
  - CCSD(T) and CCSDT(Q) methods rely on tensor slicing
- Many wavefunction methods require ‘factorization’ of tensor equations, which have some similarity to optimization of contractions arising from tensor networks of quantum circuits

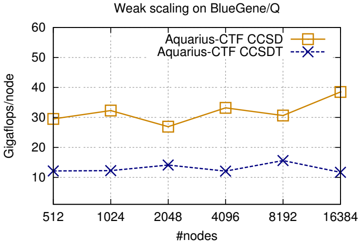
Tensors in quantum chemistry

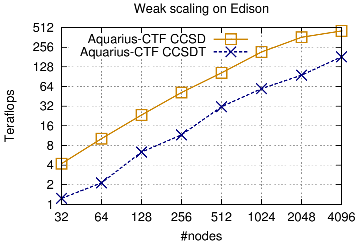
Tensor contractions dominate cost of many wavefunction methods

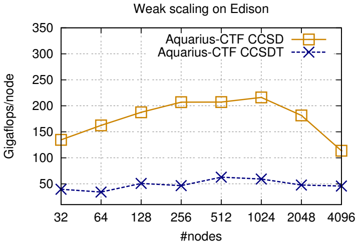
- Orbital transformations (tensor times matrix)
- Dense tensor contractions in Post-Hartree-Fock methods
  - Møller-Plesset perturbation
  - configuration interaction
  - coupled cluster
- Sparse tensors
  - localized orbitals (basis functions with compact support)
  - screening of elements
- Tensor decomposition/factorization
  - density fitting
  - resolution of identity

CCSD up to 55 (50) water molecules with cc-pVDZ
CCSDT up to 10 water molecules with cc-pVDZ









Tensor Networks Problems and Software

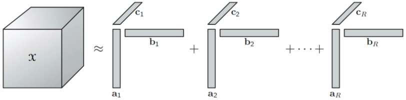
- Tensor networks bridge together quantum circuits, simulation of physical quantum systems, and numerical optimization algorithms
- They also provide a unified software base, with the main kernels being tensor contractions and numerical matrix factorizations
  - Some examples of high-performance productive libraries for tensor algebra in quantum chemistry (QC) and quantum information science (QIS) are TCE (QC), ITensor (QIS), TiledArray (QC), Cyclops (QC+QIS)

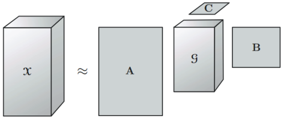
```
Z["abij"]    += V["ijab"];                    // C++
Z.i("abij") << V.i("ijab")                    // Python
W["mnij"]    += 0.5*W["mnef"]*T["efij"];      // C++
W.i("mnij") << 0.5*W.i("mnef")*T.i("efij") // Python
einsum("mnef,efij->mnij",W,T)  // numpy-style Python
```

  - These libraries support tensor transposition and contraction, (block) sparsity, optimization of contraction order, and tensor decomposition
- Classical problems on tensor networks: contraction and optimization

Tensor Network Types

- Different tensor networks arise within different problem domains
  - classical tensor decompositions: CP and Tucker[11]





  - 1D/2D lattices for quantum systems: MPS (tensor train), PEPS, MERA

$$t_{i_1 \cdots i_n} \approx \boldsymbol{w}^{(1)}_{(i_1)} \boldsymbol{W}^{(2)}_{(i_2)} \cdots \boldsymbol{W}^{(n-1)}_{(i_{n-1})} \boldsymbol{w}^{(n)}_{(i_n)}$$

  - quantum chemistry: wavefunction ansatz dependent, e.g. tensor hypercontraction[12]

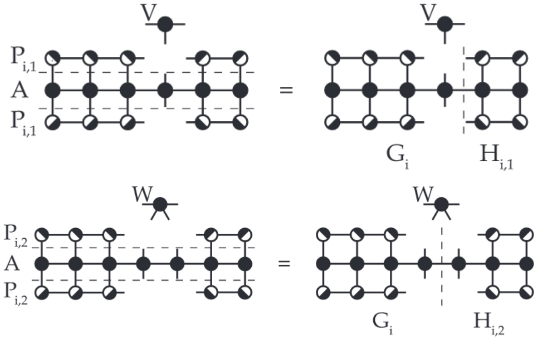
$$v^{\mu\nu}_{\lambda\sigma} \approx \sum_{p,q} x_{p\mu} x_{q\nu} z_{pq} x_{p\lambda} x_{q\sigma}$$

[11] Kolda and Bader SIAM Review 2009

[12] Hohenstein, Kokkila, Parrish, and Martinez JCP 2013

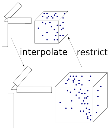
Tensor Decomposition Algorithms

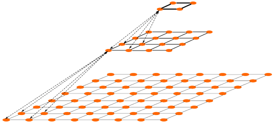
- Approximate contraction, as well as eigenvalue and fitting problems with tensor networks can be cast as optimization algorithms
- Most algorithms perform either a variant of gradient descent or alternating least squares (ALS)
- ALS (for MPS/PEPS → DMRG) is most effective for tensor networks
  - update each site/factor in network individually by quadratic optimization[13]



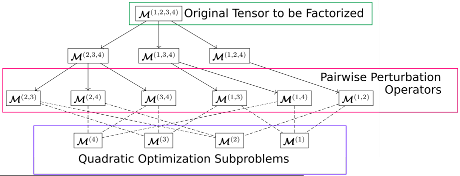
[13] Holtz, Rohwedder, and Schneider SISC 2012

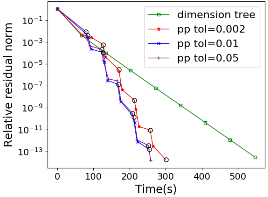
- Dimension trees amortize cost across quadratic subproblems
- Randomization/sampling can reduce cost of SVD and contractions[14]
- Multigrid/multilevel optimization employs hierarchy of networks[15]





- Pairwise perturbation approximates ALS with less asymptotic cost[16]





[14] Battaglino, Ballard, and Kolda SIMAX 2018
[15] De Sterck and Miller SISC 2013
[16] Ma and S. arXiv:1811.10573

Matrix Multiplication Algorithms from CP Decomposition

Fast algorithms for matrix multiplication correspond to CP decompositions

$$
\begin{aligned}
c_i &= \sum_{r=1}^{r} f_{ir}^{(C)} \Big(\sum_j f_{jr}^{(A)} a_j \Big) \Big(\sum_k f_{kr}^{(B)} b_k \Big) \\
&= \sum_j \sum_k \Big(\sum_{r=1}^{r} f_{ir}^{(C)} f_{jr}^{(A)} f_{kr}^{(B)} \Big) a_j b_k \\
&= \sum_j \sum_k t_{ijk} a_j b_k \quad \text{where} \quad t_{ijk} = \sum_{r=1}^{r} f_{ir}^{(C)} f_{jr}^{(A)} f_{kr}^{(B)}
\end{aligned}
$$

For multiplication of $n \times n$ matrices $\boldsymbol{C} = \boldsymbol{AB}$,

- $\boldsymbol{T}$ is $n^2 \times n^2 \times n^2$, (in/out)puts are $\boldsymbol{a}$=vec($\boldsymbol{A}$), $\boldsymbol{b}$=vec($\boldsymbol{B}$), $\boldsymbol{c}$=vec($\boldsymbol{C}$)
- Classical algorithm has rank $r = n^3$
- Strassen's algorithm has rank $r \approx n^{\log_2(7)}$
- For $n = 2$, CP rank is 7, for $n = 3$, optimal rank is open, $r \in [19, 23]$
- Tiny size of problem may make it a candidate for quantum acceleration

Broadening Participation in QIS and QC

- Quantum information science is a young and growing field with a time-critical need for broadening participation
- Place emphasis on learnability of quantum information
  - quantum carries and sometimes prides itself in difficulty/sophistication, which can inhibit confidence in students
  - students, especially underrepresented minorities and women, take more positively and are more likely to pursue learning/challenges if the carry the belief that knowledge and capability of understanding is not innate[17]
- Place emphasis on applications of quantum computing (this focus has been successful in broadening participation in CS[18])
- Create and foster an inclusive sense of community, make use codes of conduct and supervision/mentorship to prevent exclusionary culture at community events such as hackathons[19]

[17] Hoskins, Lopatto, and Stevens LSE 2011

[18] Eney, Lazowska, Martin, and Reges IEEE Computer, 2013

[19] Warner and Guo ICER 2017

Computer Science Education in Quantum

- The field of quantum computing and quantum information is dominated by physicists and theoretical computer scientists
- The frontier of quantum computing research has a growing need for more practical software development and applied mathematics
- Computer science students (at UIUC and similar departments) who pursue quantum information primarily come through two pipelines
  - theoretical computer science PhD students who become interested in QIS
  - undergraduate double majors in physics and CS
- Early education in quantum mechanics as part of CS core programs is valuable (but trend seems to be going in the opposite direction with CS+X programs with looser core requirements)
- Programs often lack undergraduate-level pure QI courses and courses in QI+physical simulation
- Michael Nielsen and Isaac Chuang's QC and QI textbook is more accessible than David Griffiths' QM to advanced CS undergraduates

Overview and Conclusion

- Quantum circuit emulation is valuable for quantum algorithm development and quantum computer verification
- Specialized quantum circuit simulation (leveraging specific gates and noise) has interesting theoretical and practical potential
- Capability of HPC systems in high-accuracy quantum chemistry simulation should play part in hybrid quantum-classical algorithms
- Tensor software provides a unified toolbox for quantum simulation
- Faster numerical algorithms for tensor network optimization is an active and promising area of research
- Fast matrix multiplication is a fundamental problem that could benefit from new quantum computing capabilities and tensor network methods
- Broadening participation and rethinking education in QIS is necessary to diversify and strengthen expertise in quantum computing

Backup slides