Optimal scaling quantum linear systems solver via discrete adiabatic theorem

PEDRO C.S. COSTA
DOMINIC W. BERRY
YUVAL R. SANDERS
DONG AN
YUAN SU
RYAN BABBUSH

2111.08152
Why do we care?

- Many problems can be posed as systems of linear equations.
  - Partial differential equations
  - EM scattering cross-section
  - Machine learning

- Many problems can be solved by adiabatic algorithms.
  - Ground states
  - Discrete optimisation
Quantum linear systems problem

- Need to solve
  \[ A\vec{x} = \vec{b} \]
- Initially prepare state
  \[ |b\rangle = \sum_{j=1}^{N} b_j |j\rangle \]
- Apply operation \( A^{-1} \)
  \[ |x\rangle = A^{-1} |b\rangle \]
- Obtain global properties from sampling \( |x\rangle \) (not explicit listing).
- Exponential speedup in \( N \).

### Complexity scaling

- **Two parameters:**
  1. \( \kappa \) – condition number (ratio of maximum & minimum singular values)
  2. \( \epsilon \) – allowable error in solution

<table>
<thead>
<tr>
<th>Year</th>
<th>Reference</th>
<th>Primary innovation</th>
<th>Complexity</th>
</tr>
</thead>
<tbody>
<tr>
<td>2008</td>
<td>Harrow, Hassidim, Lloyd</td>
<td>Phase estimation</td>
<td>( O(\kappa^2/\epsilon) )</td>
</tr>
<tr>
<td>2012</td>
<td>Ambainis (1010.4458)</td>
<td>Variable-time amplitude amplification</td>
<td>( O(\kappa (\log(\kappa)/\epsilon)^3) )</td>
</tr>
<tr>
<td>2017</td>
<td>Childs, Kothari, Somma</td>
<td>Fourier/Chebyshev fitting using LCU</td>
<td>( O(\kappa \text{polylog}(\kappa/\epsilon)) )</td>
</tr>
<tr>
<td>2018</td>
<td>Subasi, Somma, Orsucci</td>
<td>Adiabatic randomisation method</td>
<td>( O((\kappa \log \kappa)/\epsilon) )</td>
</tr>
<tr>
<td>2019</td>
<td>An, Lin (1909.05500)</td>
<td>Time-optimal adiabatic method</td>
<td>( O(\kappa \text{polylog}(\kappa/\epsilon)) )</td>
</tr>
<tr>
<td>2019</td>
<td>Lin, Tong (Quantum 4, 361)</td>
<td>Adiabatic plus eigenstate filtering</td>
<td>( O(\kappa \log(\kappa/\epsilon)) )</td>
</tr>
<tr>
<td>2021</td>
<td>This work (2111.08152)</td>
<td>Discrete adiabatic theorem</td>
<td>( O(\kappa \log(1/\epsilon)) )</td>
</tr>
</tbody>
</table>
Continuous adiabatic algorithm

- For continuous adiabatic algorithm, simulate time-dependent Hamiltonian
  \[ i \frac{d}{dt} |\psi(t)\rangle = H(t)|\psi(t)\rangle \]

- For short time, the time-evolution operator is given by Dyson series
  \[ U(t, 0) \approx \sum_{k=0}^{K} (-1)^k \int_0^t dt_1 \int_{t_1}^t dt_2 \cdots \int_{t_{k-1}}^t dt_k H(t_k) \cdots H(t_2)H(t_1) \]

- Gives \( \log(\kappa/\epsilon) \) overhead for adiabatic solution of QLSP.

M. Kieferová, A. Scherer, and D. W. Berry, PRA 99, 042314 (2019); 1805.00582.
Linear combinations of unitaries

\[ A = \sum_{\ell} w_{\ell} U_{\ell} \]

Cost is typically

\[ \lambda = \sum_{\ell} |w_{\ell}| \]
Block encoding

\[ A = \lambda \langle 0 | U | 0 \rangle \]

Qubitisation

- Construct quantum walk using reflection:

\[
 W = U R
\equiv e^{i \arcsin A/\lambda}
\]

- Eigenvalues are related to those of original matrix.

Adiabatic approach to QLSP

- Easiest for $A$ positive-definite and Hermitian.
- Initial and final Hamiltonians with $Q_b = I_N - |b\rangle\langle b|$
  $$H_0 = \begin{pmatrix} 0 & Q_b \\ Q_b & 0 \end{pmatrix}, \quad H_1 = \begin{pmatrix} 0 & AQ_b \\ Q_bA & 0 \end{pmatrix}$$
- Final eigenstate is
  $$(A^{-1}b)$$
- For $H(s) = (1 - f(s))H_0 + f(s)H_1$ gap is
  $$\Delta_0(s) = 1 - f(s) + f(s)/\kappa$$
- Schedule is chosen as
  $$\dot{f}(s) = d_p \Delta_0^p(s), \quad d_p = \int_0^1 \Delta_0^{-p}(u) \, du$$
- Solution is
  $$f(s) = \frac{\kappa}{\kappa - 1} \left[ 1 - (1 + s(\kappa^{p-1} - 1))^{\frac{1}{1-p}} \right]$$

Non-symmetric case

- An & Lin expand matrix for general Hermitian case

\[
\begin{align*}
H_0 &= \begin{pmatrix} 0 & (\sigma_z \otimes I)Q_b \\ Q_b(\sigma_z \otimes I) & 0 \end{pmatrix} \\
H_1 &= \begin{pmatrix} 0 & (\sigma_z \otimes A)Q_b \\ Q_b(\sigma_z \otimes A) & 0 \end{pmatrix}
\end{align*}
\]

- Standard (HHL) method for non-Hermitian \( A \) is to replace \( A \) with \( A = \begin{pmatrix} 0 & A \\ A^\dagger & 0 \end{pmatrix} \)

- To simplify, we use

\[
A(f) = \begin{pmatrix} (1-f)I & fA \\ fA^\dagger & (1-f)I \end{pmatrix} \\
H(s) = \begin{pmatrix} 0 & A(f)Q_b \\ Q_bA(f) & 0 \end{pmatrix}
\]

- Gap is changed to

\[
\Delta_0(s) = \sqrt{(1 - f(s))^2 + (f(s)/\kappa)^2} \geq (1 - f(s) + f(s)/\kappa)/\sqrt{2}
\]
Theorem for discrete sequence of operators (with integers $n, T$)

$W_T(n/T), \quad 0 \leq n \leq T - 1$

Overall operation is sequence

$$U_T(s) = W_T(s - 1/T)W_T(s - 2/T) \ldots W_T(1/T)W_T(0) = \prod_{n=0}^{sT-1} W_T(n/T)$$

We call the ideal adiabatic evolution $U^A_T(s)$.

Aim to show error in adiabatic evolution decreases with $T$ as

$$\|U_T(s) - U^A_T(s)\| \leq \frac{\theta}{T}$$
Norm of differences

- For $T$ steps one would want difference of operators to satisfy
  $$W_T(s + 1/T) - W_T(s) = O(T^{-1})$$

- Need to define constant
  $$\|W_T(s + 1/T) - W_T(s)\| \leq \frac{c_1(s)}{T}$$

- Need to define second difference:
  $$DW_T(s) = W_T(s + 1/T) - W_T(s), \quad D^{(2)}W_T(s) = DW_T(s + 1/T) - DW_T(s)$$

- Require bounds on multistep differences:
  $$\|D^{(2)}W_T(s)\| \leq \frac{c_2(s)}{T^2}$$

- Need to consider maximum over neighbouring steps:
  $$\hat{c}_k(s) = \max(c_k(s - 1/T), c_k(s), c_k(s + 1/T))$$
Multistep gap

- Eigenvalues are on circle in complex plane.
- Gaps are needed on both sides, $\Delta_0(s)$. 

$\sigma_P(s)$

$\Delta_0(s)$

$\sigma_Q(s)$
Multistep gap

- Eigenvalues are on circle in complex plane.
- Gaps are needed on both sides, $\Delta_0(s)$.
- Regions must not cross over between steps.
Multistep gap

- Eigenvalues are on circle in complex plane.
- Gaps are needed on both sides, $\Delta_0(s)$.
- Regions must not cross over between steps.
**Multistep gap**

- Eigenvalues are on circle in complex plane.
- Gaps are needed on both sides, $\Delta_0(s)$.
- Regions must not cross over between steps.
- Consider unified regions between three successive steps.
• Eigenvalues are on circle in complex plane.
• Gaps are needed on both sides, $\Delta_0(s)$.
• Regions must not cross over between steps.
• Consider unified regions between three successive steps.
Multistep gap

- Eigenvalues are on circle in complex plane.
- Gaps are needed on both sides, \( \Delta_0(s) \).
- Regions must not cross over between steps.
- Consider unified regions between three successive steps.
Multistep gap

- Eigenvalues are on circle in complex plane.
- Gaps are needed on both sides, $\Delta_0(s)$.
- Regions must not cross over between steps.
- Consider unified regions between three successive steps.
- We call gap between unified regions $\Delta(s)$.
- To simplify expressions need minimum over neighbouring steps too:
  \[ \tilde{\Delta}(s) = \min(\Delta(s - 1/T), \Delta(s), \Delta(s + 1/T)) \]
Discrete adiabatic theorem

The error in the adiabatic evolution is bounded as

\[ \|U_T(s) - U_T^A(s)\| \leq \frac{12\hat{c}_1(0)}{T\Delta(0)^2} + \frac{12\hat{c}_1(s)}{T\Delta(s)^2} + \frac{6\hat{c}_1(s)}{T\Delta(s)} + 305 \sum_{n=1}^{sT-1} \frac{\hat{c}_1(n/T)^2}{T^2\Delta(n/T)^3} + 44 \sum_{n=0}^{sT-1} \frac{\hat{c}_1(n/T)^2}{T^2\Delta(n/T)^2} + 32 \sum_{n=1}^{sT-1} \frac{\hat{c}_2(n/T)}{T^2\Delta(n/T)^2} \]

- \( \hat{c}_1 \) – analogous to first derivative in continuous adiabatic theorem
- \( \hat{c}_2 \) – analogous to second derivative in continuous adiabatic theorem
The error in the adiabatic evolution is bounded as
\[
\| U_T(s) - U_T^A(s) \| \leq \frac{12 \hat{c}_1(0)}{T \Delta(0)^2} + \frac{12 \hat{c}_1(s)}{T \Delta(s)^2} + \frac{6 \hat{c}_1(s)}{T \Delta(s)} + 305 \sum_{n=1}^{sT-1} \frac{\hat{c}_1(n/T)^2}{T^2 \Delta(n/T)^3} + 44 \sum_{n=0}^{sT-1} \frac{\hat{c}_1(n/T)^2}{T^2 \Delta(n/T)^2} + 32 \sum_{n=1}^{sT-1} \frac{\hat{c}_2(n/T)}{T^2 \Delta(n/T)^2}
\]

- \( \hat{c}_1 \) – analogous to first derivative in continuous adiabatic theorem
- \( \hat{c}_2 \) – analogous to second derivative in continuous adiabatic theorem
Method to prove theorem

- “Wave operator” describes difference from ideal adiabatic evolution:
  \[ \Omega_T(s) = U_T^{A+}(s)U_T(s) \]

- Error can be considered entirely in terms of wave operator:
  \[ \|U_T(s) - U_T^A(s)\| = \|\Omega_T(s) - I\| \]

- The “ripple operator” describes the change in the wave operator:
  \[ \Theta_T(s) = \Omega_T(s + 1/T)\Omega_T^+(s) \]

- The “kernel function” describes how close this is to the identity:
  \[ K_T(s) = T(I - \Theta_T(s)) \]

- Write wave operator in terms of these as “Volterra equation”
  \[ \Omega_T(n/T) = I - \frac{1}{T} \sum_{m=0}^{n-1} K_T(m/T)\Omega_T(m/T) \]

Diagonal vs off-diagonal terms

- Resolve identity as $I = P_0 + Q_0$ with initial projections $P_0 = P_T(0)$, $Q_0 = Q_T(0)$.
- Introduce to Volterra equation so

$$
\| \Omega_T(s) - I \| = \frac{1}{T} \left\| \sum_{m=0}^{n-1} (P_0 + Q_0) K_T(m/T)(P_0 + Q_0) \Omega_T(m/T) \right\|
$$
Diagonal vs off-diagonal terms

- Resolve identity as \( I = P_0 + Q_0 \) with initial projections \( P_0 = P_T(0), Q_0 = Q_T(0) \).
- Introduce to Volterra equation so

\[
\| \Omega_T(s) - I \| = \frac{1}{T} \left\| \sum_{m=0}^{n-1} (P_0 + Q_0)K_T(m/T)(P_0 + Q_0)\Omega_T(m/T) \right\|
\]

\[
\leq \frac{1}{T} \left\| \sum_{m=0}^{n-1} P_0K_T(m/T)P_0\Omega_T(m/T) \right\| + \frac{1}{T} \left\| \sum_{m=0}^{n-1} Q_0K_T(m/T)Q_0\Omega_T(m/T) \right\|
\]

\[
+ \frac{1}{T} \left\| \sum_{m=0}^{n-1} P_0K_T(m/T)Q_0\Omega_T(m/T) \right\| + \frac{1}{T} \left\| \sum_{m=0}^{n-1} Q_0K_T(m/T)P_0\Omega_T(m/T) \right\|
\]

“diagonal”

“off-diagonal”
Summation by parts formula

- Off-diagonal term can be rewritten as, with $V_T(s) = W_T^A(s)W_T^\dagger(s)$,
  \[
  \sum_{n=1}^{sT} Q_0 U_T^{A\dagger} \left( \frac{n}{T} \right) \left( I - V_T^\dagger \left( \frac{n-1}{T} \right) \right) U_T^A \left( \frac{n}{T} \right) P_0 \Omega_T \left( \frac{n-1}{T} \right)
  \]

- Summation by parts formula gives
  \[
  \sum_{n=1}^{l} Q_0 U_T^{A\dagger} \left( \frac{n}{T} \right) X \left( \frac{n}{T} \right) U_T^A \left( \frac{n}{T} \right) P_0 Y \left( \frac{n}{T} \right) = B - \frac{1}{T} S
  \]
  
  - $B$ is “boundary term”
  - $S$ is “sum term” – it still involves a sum, but is easier to bound.
  - Requires operator that can only be described by contour integral:
    \[
    \tilde{X}(s) = -\frac{1}{2\pi i} \oint_{\Gamma_T(s)} R_T(s, z)X(s)R_T(s, z)dz, \quad R_T(s, z) = (W_T(s) - zI)^{-1}
    \]
Contour integrals for bounds

- Projector can be given by contour integral:
  \[ P_T(s) = \frac{1}{2\pi i} \oint_{\Gamma_T(s)} R_T(s, z)dz \]
  \[ R_T(s, z) = (W_T(s) - zI)^{-1} \]

- Bound on difference of projectors needs contour threaded between eigenvalues for multiple steps.
  \[ \|DP_T(s)\| \leq \frac{2c_1(s)}{T\Delta_1(s)} \]
Theorem for $p = 1.5$

- For $A$ positive-definite & Hermitian, bound all terms from theorem as follows:
  \[
  \frac{\hat{c}_1(0)}{T \bar{\Lambda}(0)^2} = \frac{4\sqrt{\kappa}}{T} + O\left(\frac{\kappa}{T^2}\right) \\
  \frac{\hat{c}_1(1)}{T \bar{\Lambda}(1)^2} = \frac{4\kappa}{T} + O\left(\frac{\kappa}{T^2}\right) \\
  \frac{\hat{c}_2(1)}{T \bar{\Lambda}(1)^2} = \frac{4}{T} + O\left(\frac{1}{T^2}\right)
  \]

- Overall inequality becomes
  \[
  \|U_T(s) - U_T^A(s)\| \leq 5632 \frac{\kappa}{T} + O\left(\frac{\sqrt{\kappa}}{T}\right)
  \]

- For general $A$, gap has factor of $1/\sqrt{2}$ – replace $\bar{\Lambda}$ with $\bar{\Lambda}/\sqrt{2}$.

- Updated inequality is
  \[
  \|U_T(s) - U_T^A(s)\| \leq 15307 \frac{\kappa}{T} + O\left(\frac{\sqrt{\kappa}}{T}\right)
  \]
The discrete adiabatic theorem is for target eigenspace being in one region.

For quantum walk zero eigenvalue of Hamiltonian maps to $\pm 1$ for walk operator; eigenstates ($|k\rangle$ is eigenstate on system)

$$|\pm\rangle = \frac{1}{\sqrt{2}} (|0\rangle |k\rangle \pm i|0\perp\rangle)$$

$$\pi - \arcsin \frac{h}{\lambda} \quad \arcsin \frac{h}{\lambda}$$
Multiple eigenvalues problem

- The discrete adiabatic theorem is for target eigenspace being in one region.
- For quantum walk zero eigenvalue of Hamiltonian maps to $\pm 1$ for walk operator; eigenstates ($|k\rangle$ is eigenstate on system)

$$|\pm\rangle = \frac{1}{\sqrt{2}} (|0\rangle|k\rangle \pm i|0k^\perp\rangle)$$

$$|0\rangle|k\rangle = \frac{1}{\sqrt{2}} (|+\rangle + |−\rangle)$$

- We show phase between $|+\rangle$ and $|−\rangle$ is preserved.
- We show orthogonal degenerate eigenstates do not cross.

$$\begin{pmatrix} A(f)^{-1}b \\ 0 \end{pmatrix} \nu_s \begin{pmatrix} 0 \\ b \end{pmatrix}$$
Numerical testing for constant factor

- Theorems for QLSP use many simplifications with loose inequalities.
- Numerically test with the complete discrete adiabatic theorem.
- \( \| \Omega(1) - I \| \approx 1/2 \), and \( \kappa/T = 40/50000 = 1/1250 \), so \( \| \Omega(1) - I \| \geq 625\kappa/T \).
- About 9 times better than theorem.
Filtering solution

- To obtain solution with precision $\epsilon$ need to filter the eigenstate.
- Lin & Tong use quantum signal processing – drawback is computing rotation angles.
- We instead apply linear combination of unitaries $\sum_j w_j W_j^T$.
- LCU maps superposition of eigenstates as
  \[ \sum_k \psi_k |k\rangle \mapsto \sum_k \tilde{w}(\phi_k) \psi_k |k\rangle, \]
  \[ \tilde{w}(\phi) = \frac{1}{\sum_j w_j} \sum_j w_j e^{ij\phi} \]
- Error goes like $\tilde{w}(\phi)$ for incorrect phases – minimised for Dolph-Chebyshev window.
LCU with two qubits

- Normally LCU would require a large number of qubits for control.
- A single ancilla qubit is given as major advantage of quantum signal processing.
- We can do LCU using just two ancilla qubits.
- First consider control state in unary:
• Normally LCU would require a large number of qubits for control.
• A single ancilla qubit is given as major advantage of quantum signal processing.
• We can do LCU using just two ancilla qubits.
• Switch sequence of controlled operations for inverse preparation:
LCU with two qubits

- Normally LCU would require a large number of qubits for control.
- A single ancilla qubit is given as major advantage of quantum signal processing.
- We can do LCU using just two ancilla qubits.
- Shift order of operations so we use two qubits at a time:
We can use these results for solving the QLSP with complexity $O(\kappa \log(1/\epsilon))$ in terms of calls to a block encoding to $A$, as well as preparation of $|b\rangle$.

The procedure is:

1. Construct a quantum walk step from the block encoding of $A$.
2. Use the discrete adiabatic theorem to solve for $|A^{-1}b\rangle$ with fixed precision $1/2$. Complexity is $O(\kappa)$.  
3. Filter solution using LCU; complexity is $\kappa \log(2/\epsilon)$ steps of walk.
Kothari and Harrow have proven lower bound for sparse matrices (unpublished)

\[ \Omega(\sqrt{d}\kappa \log(1/\epsilon)) \]

Our complexity for general block-encoded matrices is

\[ \Omega(\kappa \log(1/\epsilon)) \]

For sparse matrices, this would need to be multiplied by complexity of block-encoding sparse matrices, typically \( d \).

Low has demonstrated \( \sqrt{d} \) scaling for Hamiltonian simulation up to logarithmic factors (STOC 2019).

Similar approach here would make scaling no longer optimal in \( \kappa \) and \( \epsilon \).
Conclusions

- Fastest possible quantum algorithm for the quantum linear systems problem in terms of $\kappa$ and $\epsilon$ (we don’t consider sparsity $d$).
- Lower bound of Kothari and Harrow is $\Omega(\sqrt{d}\kappa \log(1/\epsilon))$.
- Our discrete adiabatic algorithm with qubitisation gives speedup for all adiabatic algorithms too.
- Method for LCU with two ancilla qubits can be used generally.

Future work:
- There could be scope for improving constant factors.
- How to achieve optimal scaling in $d$ at the same time is open problem.