HHL Quantum Linear Systems Algorithm

Trace HHL's coherent reciprocal weighting and its output limitations.

Undergraduate | 22 min lesson | Dr. Charith Jayasekara

By the end, you'll be able to

- State conditioning, access and output assumptions.
- Calculate a diagonal-system solution state and success probability.
- Distinguish solution amplitudes from measured probabilities.

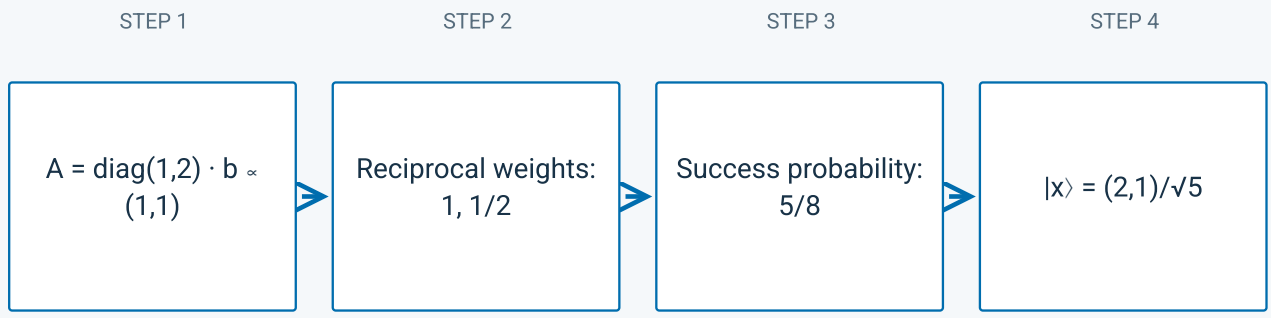
Solve for a state, not a printed list

HHL is a quantum algorithm for a linear-system task under specific access and conditioning assumptions. Given a matrix A and a vector b, the classical equation is A x equals b. The quantum procedure aims to prepare a normalised state proportional to the solution, enabling selected measurements about that solution. It does not automatically return every entry of x as classical numbers.

Prerequisites: eigenvector-ready linear algebra, phase kickback and phase estimation. The original paper explicitly focuses on extracting a property associated with the solution rather than printing the full vector. Harrow, Hassidim and Lloyd.

WORKED DIAGRAM · A26

Diagonal HHL example



The solution-state Z probabilities are 4/5 and 1/5. They are squared normalised magnitudes, not the classical solution entries.

The input and output setup

The basic explanation assumes A is Hermitian, invertible and scaled so its eigenvalues have bounded magnitude away from zero. Efficient access to A and efficient preparation of the state for b must be supplied. The ratio of largest to

smallest singular value is the condition number, kappa. Small eigenvalues make inversion sensitive and generally increase cost.

The output is approximately the solution state, together with a success mechanism and a way to estimate requested observables. Recovering the entire N-entry solution requires additional information extraction and can remove an apparent advantage based on logarithmic dependence on N. A fair classical comparison should solve the same output task under comparable data-access assumptions.

Worked example: a diagonal two-by-two system

Take the unscaled matrix with diagonal entries one and two, and b proportional to the column one, one. The classical solution is proportional to one, one half. Its normalised quantum state is

$$|x\rangle = \frac{2|0\rangle + |1\rangle}{\sqrt{5}}.$$

Its Z probabilities are four fifths and one fifth. These probabilities are not the solution entries themselves. They are their squared normalised magnitudes. Signs and phases would require other measurement settings, and the overall norm of the classical solution is not contained in a normalised state alone.

For the algorithm, scale A by two so the eigenvalues are one half and one. Choose the phase-estimation unitary $U = \exp(i\pi A_{\text{scaled}})$, giving phase fractions one quarter and one half respectively. Two counting qubits resolve this toy example exactly. Using $\exp(2\pi i A_{\text{scaled}})$ without another scaling would instead wrap eigenvalue one to phase zero and make reciprocal decoding ambiguous. Prepare plus for b. Phase estimation records the relevant eigenvalue coherently beside each eigenvector. A controlled ancilla rotation places an amplitude proportional to the reciprocal eigenvalue in the success branch. Choose the reciprocal scale constant one half, giving success amplitudes one and one half for the two eigenvalues.

Undo phase estimation to remove the eigenvalue register. Conditional on the success ancilla, the system amplitudes are proportional to one and one half, giving the state above. The success probability is the average of one squared and one half squared, or five eighths. Omitting that probability would hide part of the procedure's cost.

The scaled eigenvalues and rotation scale give the complete success probability.

$$A_{\text{scaled}} = \operatorname{diag}(1/2, 1), \quad C = 1/2, \quad |b\rangle = |+\rangle, \qquad P_{\text{success}} = \frac{1}{2}\left(1^2 + (1/2)^2\right) = \frac{5}{8}.$$

The success amplitudes are one and one half on the two eigenvectors, respectively. Normalising their resulting column gives the existing solution $(2|0\rangle + |1\rangle)/\sqrt{5}$.

Why the inverse works

Expand b in A's eigenbasis. The linear-system solution divides each eigenvector coefficient by its eigenvalue. Phase estimation makes an approximation to that eigenvalue available as a coherent control label. The rotation then changes the relative coefficient by the reciprocal factor. Uncomputation removes the label so the eigenvector components can form the intended solution state coherently.

Measuring the eigenvalue register before uncomputation would select an eigencomponent rather than implement the desired coherent inverse on their superposition. Likewise, merely learning a few eigenvalues classically is not the same algorithm. The crucial transformation reweights components without reading which component is present, while its ancilla success probability respects normalisation.

In the eigenbasis, inversion weights each component by the reciprocal eigenvalue; an ancilla supplies the bounded success amplitude.

$$|b\rangle = \sum_j \beta_j |u_j\rangle, \qquad |\widetilde{x}\rangle = \sum_j \frac{C\beta_j}{\lambda_j}|u_j\rangle, \qquad P_{\text{success}} = \|\widetilde{x}\|^2, \qquad |x\rangle = \frac{|\widetilde{x}\rangle}{\|\widetilde{x}\|}.$$

λ_j are nonzero eigenvalues of the scaled matrix and C is chosen so $|C/\lambda_j| \leq 1$. The tilde marks an unnormalised success branch. Phase estimation must be uncomputed before conditioning leaves the desired system state.

Condition number and precision

If the smallest relevant eigenvalue is tiny, its reciprocal is large. The rotation must use a bounded common scale, reducing success amplitude on other components. Phase estimation must also resolve eigenvalues accurately enough that reciprocal errors remain controlled. This is why condition number and accuracy appear in complexity statements rather than only the matrix dimension.

An ill-conditioned linear system is sensitive classically too. A small perturbation to b or A can cause a large change in x. Quantum inversion does not remove that mathematical sensitivity. Preconditioning can help, but constructing and applying a suitable preconditioner must be included in the access model and resource estimate.

A complete implementation plan

Verify A's properties and scaling. Specify sparse-oracle or block-encoding access and the b preparation. Choose eigenvalue-resolution and simulation tolerances. Perform coherent phase estimation, controlled reciprocal rotation and inverse phase estimation. Use postselection or a justified amplification strategy. Estimate the required observable on independently prepared successful outputs, including statistical uncertainty.

For the diagonal example, a full quantum implementation is unnecessary for solving the problem; direct arithmetic is cheaper. Its purpose is to make every amplitude and success probability checkable. An educational demonstration that hard-codes known eigenvalues should be labelled as a mechanism example rather than a scalable linear-system solver.

The normalised output does not retain every classical quantity

Choose the unnormalised right-hand side b equal to the column one, one for the original diagonal matrix. The classical solution is one, one half, whose Euclidean norm is square root of five divided by two. The quantum solution state divides by that norm and becomes the column two, one over square root of five.

Multiplying the original right-hand side by seven leaves the normalised quantum state unchanged, while multiplying the classical solution and its norm by seven. Thus no measurement on the normalised state alone can recover that missing overall scale without additional input information or estimation. This is a concrete output limitation, not merely a technical caveat.

When an application needs a particular observable of the normalised solution, that may be exactly the desired output. When it needs physical amplitudes with an absolute scale or every vector entry, include the extra estimation work. Match the algorithm's output to the application before discussing a speedup.

Complexity without the headline shortcut

The original HHL analysis provides favourable dimension dependence under sparsity, efficient preparation, bounded conditioning and suitable output requirements. Modern linear-system methods improve several parameter dependences using other techniques. Do not quote one universal "exponential speedup" without specifying the access model, kappa, precision and measurement task. Modern matrix-transformation framework.

Exercises and worked answers

Exercise 1. For the worked solution state, what is the Z expectation?

> Check the worked answer
>
> It is four fifths minus one fifth, equal to three fifths. This is a small observable that can be estimated without printing the full classical solution vector.

Exercise 2. If the input is entirely the eigenvector with scaled eigenvalue one, what is the success probability with reciprocal scale one half?

> Check the worked answer
>
> It is one quarter. The success amplitude is one half, even though the conditional output direction is unchanged. Input dependence of success is part of the nonunitary inverse implementation and must be counted.

Check your understanding

PAUSE & THINK

What are the Z probabilities of $(2|0\rangle+|1\rangle)/\sqrt{5}$?

- [ ] 2/3 and 1/3
- [ ] 4/5 and 1/5
- [ ] 1/2 and 1/2

Answer: 4/5 and 1/5. Square the normalised amplitudes: 4/5 and 1/5. The classical vector's overall norm is not supplied by this state alone.

Sources

- Harrow, Hassidim and Lloyd
- Modern matrix-transformation framework