Quantum annealing for polynomial systems of equations

Chia Cheng Chang (張家丞),
Arjun Gambhir,
Travis S. Humble,
and Shigetoshi Sota

1 Interdisciplinary Theoretical and Mathematical Sciences Program (iTHEMS), RIKEN, 2-1 Hirosawa, Wako, Saitama 351-0198, Japan
2 Department of Physics, University of California, Berkeley, California 94720, USA
3 Nuclear Science Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA
4 Physics Division, Lawrence Livermore National Laboratory, Livermore, CA 94550, USA
5 Quantum Computing Institute, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831, USA
6 Computational Materials Science Research Team, RIKEN Advanced Institute for Computational Science (AICS), Kobe, Hyogo 650-0047, Japan

Numerous fields require numerically solving a system of linear equations. For equations stemming from large, sparse matrices, this is classically done with iterative methods and judicious preconditioning. Convergence of such algorithms can be highly variable and depends in part, on the condition number of the matrix. With the arrival of quantum computing in the Noisy Intermediate-Scale Quantum era, we present a quantum annealing algorithm which directly solves systems of linear equations with favorable scaling with respect to condition number. The algorithm can also be extended to solve systems of polynomial equations. We discuss the theoretical framework for this method and perform experiments of the algorithm with a quantum annealer. Scaling with respect to matrix rank, condition number, and search precision is studied. Finally, we define an iterative annealing process and demonstrate its efficacy in solving a linear system to a tolerance of 10^{-8}.

INTRODUCTION.

Many problems in science, engineering, and mathematics can be reduced to solving systems of equations with notable examples in modeling and simulation of physical systems and the verification and validation of engineering designs. Conventional methods for solving linear systems range from exact methods, such as matrix diagonalization, to iterative methods, such as fixed-point solvers. The advent of quantum computing has opened up the possibility of new methods for solving these challenging problems. For example, a quantum algorithm for solving systems of linear equations was established for gate-based quantum computers [1] and demonstrated with small-scale problem instances [2]. Additionally, an algorithm for solving linear systems within the adiabatic quantum computing model [3] was experimentally demonstrated [4], followed by a more recent proposal [5].

In this Letter, we present an approach for solving a more general system of polynomial equations using quantum annealing. Akin to adiabatic quantum optimization [6], quantum annealing prepares a quantum statistical distribution that approximates a sought-after solution by applying a slowly changing, time-dependent Hamiltonian [7], where measurements drawn from the distribution represent candidate solutions. Unlike adiabatic quantum computing, quantum annealing permits non-adiabatic dynamics at non-zero temperature, making this approach easier to realize experimentally but also more challenging to distinguish quantum mechanically [7–14]. While examples of non-trivial advantages have been observed for fixed-size problem instances [15–19], more general statements about computational complexity remain unresolved [20].

Based on the principles of quantum annealing, we demonstrate a novel probabilistic solver for ill-conditioned systems of linear equations. We use the discretized Dirac equation \( D\phi = \chi \) from lattice quantum chromodynamics (QCD) as a motivating example, in which the Dirac operator \( D \) encodes the fundamental interactions between quarks and gluons that lead to formation of hadrons. The solution to the discretized equation is currently the only approach to evaluating QCD non-perturbatively, and yields the quark propagator \( \phi \) for a given source \( \chi \). However, well-known numerical challenges slow convergence with conventional solvers [21, 22]. We use quantum annealing to solve this system and characterize the performance from experimental demonstrations with a commercial quantum annealer. Our implementation recovers the correct solution with high precision and accuracy for problem sizes that fit within the available hardware.

Polynomial Systems of Equations.

We consider the system of \( N \) polynomial equations

\[
P^{(1)}_{ij} x_j + P^{(2)}_{ijk} x_j x_k + \cdots = P^{(0)}_i
\]

where \( i \in \{1, \ldots, N\} \), and \( P^{(n)} \) is a rank \( n + 1 \) tensor of known real-valued coefficients for the polynomial of order \( n \), and the real-valued vector \( x \) denotes the solution. Truncating to first order recovers a linear system of equations, i.e., \( P^{(1)}_{ij} x_j = P^{(0)}_i \).

Existing approaches for solving Eq. (1) include direct diagonalization using Gauss-Jordan elimination or iterative methods such as conjugate-gradient. In practice, direct diagonalization is limited in computational efficiency, as those methods scale sharply with the size of the matrix. By contrast, iterative methods may have greater computational efficiency but the performance and stability are often sensitive to the input matrix. For example, the Dirac operator

\[
\chi \partial_\mu A^\mu - \frac{i}{2} \gamma_5 \gamma^\mu (\partial_\mu A^\mu) + \mu \phi = 0
\]

is the fundamental interaction between quarks and gluons in QCD, where \( A^\mu \) is the gluon field strength and \( \phi \) is the quark density. While this equation is unsolvable analytically, it is a common starting point for many lattice simulations.

\[
\chi \partial_\mu A^\mu - \frac{i}{2} \gamma_5 \gamma^\mu (\partial_\mu A^\mu) + \mu \phi = 0
\]
\[ D \] becomes more singular as the quark mass is lowered to its physical value. The condition number of the resulting \( D \) matrix slows the convergence of the linear solver and causes low quark mass calculations to be prohibitively expensive.

Preconditioning improves convergence by transforming the input as \( M^{-1}Ax = M^{-1}b \), where the preconditioner \( M \) must be inexpensive to invert and \( M^{-1} \) should be “close” to \( A^{-1} \), so that \( M^{-1}A \) resembles a matrix close to unity. Identifying an effective preconditioner plays an important role in numerical convergence of iterative methods [23–27]. For lattice QCD applications, the low-lying spectrum of the Dirac operator slows iterative convergence and preconditioning has been used to project out these low-lying modes. Acquiring the low-lying eigenpairs or singular triplets of \( D \) is in general computationally expensive because it requires the use of additional iterative methods that also suffer from critical slowing down and poor scaling with the physical volume. Solutions to address this issue include EigCG [21, 22], inexact deflation [23], and adaptive multigrid [24–27].

RESULTS

Quantum Annealing for Polynomial Solvers

Quantum annealing offers an alternative approach to solving ill-conditioned systems of equations. By recasting Eq. (1) as the scalar minimization problem

\[
\text{arg min}_x \left[ \frac{1}{2} P_{11}^{(1)} x_1 x_1 + \frac{1}{3} P_{12}^{(2)} x_1 x_2 x_2 + \ldots + P_{NN}^{(N)} x_N x_N \right],
\]

we reduce finding the solution to minimizing an \( x \)-variable objective function. We establish the connection to quantum annealing by first approximating the objective function using an \( R \)-bit representation for each real-valued variable \( x_j \) as

\[
\hat{x}_j = a_j \sum_{r=0}^{R-1} 2^r \psi_{r,j} + b_j.
\]

where \( \psi_{r,j} \in \{0, 1\} \) and \( a_j \) and \( b_j \) control the range and numerical shift, respectively, such that \( \hat{x}_j \in [b_j + 2^{R-1}a_j] \) in increments of \( a_j \). Direct substitution yields the multi-body objective function

\[
H_{\text{Ising}}(\psi) = h_{i} \psi_{i} + J_{ij} \psi_{i} \psi_{j} + K_{ijk} \psi_{i} \psi_{j} \psi_{k} + \ldots
\]

where its ground state solves a system of polynomial equations. For current commercial quantum annealers, ancillary binary variables are required to decompose multi-linear terms into two-body interactions. This quadratization process introduces more variables into the problem with additional penalty terms [28–30].

Quantum Annealing for Linear Solvers

Restricting the transformation of Eq. (3) to a system of linear equations simplifies Eq. (4) by involving only terms up to two-body interactions. As a result the explicit binary quadratic optimization problem (QUBO) is \( H^{\text{QUBO}}(\psi) = \psi_i Q_{ij} \psi_j \) with

\[
Q = \frac{1}{2} \begin{pmatrix}
 a_{11} P_{11}^{(1)} & \ldots & a_{1N} P_{1N}^{(1)} \\
 \vdots & \ddots & \vdots \\
 a_{N1} P_{N1}^{(1)} & \ldots & a_{NN} P_{NN}^{(1)}
\end{pmatrix} \otimes \begin{pmatrix}
 2^{0} & 2^{0} & \ldots & 2^{0} & 0^{R-1} \\
 0^{R-1} & 2^{0} & \ldots & 2^{0} & 0^{R-1} \\
 \vdots & \vdots & \ddots & \vdots & \vdots \\
 0^{R-1} & 0^{R-1} & \ldots & 2^{0} & 0^{R-1}
\end{pmatrix} - \begin{pmatrix}
 a_{1} S_{1} \\
 \vdots \\
 a_{N} S_{N}
\end{pmatrix} \otimes \begin{pmatrix}
 2^{0} \\
 0^{R-1}
\end{pmatrix}
\]

where \( S_n = P_{nn}^{(n)} - b_n \sum_i P_{ni}^{(n)} \). Note that the matrix \( Q \) is real-valued and symmetric. In addition, constant terms that arise from the substitution of Eq. (3) are omitted, as these simplifications shift the energy spectrum by a constant but leave the minimizing solution unchanged.

Linear Least Squares

Least squares minimization is a standard approach for regression analysis of an overdetermined system [31]. In general, the unknown parameters of interest, dependent variables and independent variables can be continuous and correlated. Generalized least squares offers a strategy for defining the best-fit such that the parameters in the regression function best describe the given data.

Given a set of \( N \) identical and independent observations of the

\[
\text{independent } \{x_i : i \in \{1, \ldots, X\}\} \\
\text{dependent } \{y_{i,g} : i \in \{1, \ldots, X\}, g \in \{1, \ldots, N\}\}
\]

variable, the mean and covariance of \( y_i \) follows

\[
\langle y_i \rangle = \frac{1}{N} \sum_{g=1}^{N} y_{i,g} \\
S_{ij} = \langle (y_i - \langle y_i \rangle)(y_j - \langle y_j \rangle) \rangle
\]

where the angle brackets denote the expectation value over \( N \) observations. A fitting function \( F(x, p) \) may be defined with respect to the set of \( P \) unknown parameters \( p = \{p_n : n \in \{1, \ldots, P\}\} \), and a corresponding objective function for generalized least squares may be defined as

\[
\chi^2 = [F(x, p) - \langle y \rangle]_i S_{ij}^{-1} [F(x, p) - \langle y \rangle]_j
\]

where the optimal value for the set \( p \) is determined by minimizing the value of \( \chi^2 \).

Restriction to linear least squares demands that the fitting function is linear in the unknown parameters, and therefore
may be written in the form
\[ F(x_i, p) = \sum_{n=1}^{P} p_n f_n(x_i) \] (7)
where \( f_n(x_i) \) can be any function. The solution for linear regression is obtained by expanding Eq. (6) with Eq. (7) and yields
\[ p_n f_n(x_i) - y_i | S_{ij}^{-1} p_m f_m(x_j) - y_j \] (8)
The extrema of the objective function can be determined by taking the derivative of Eq. (8) with respect to \( p_n \) yielding a matrix equation of the form \( A^{(1)} p_j = A^{(0)} \) analogous to Eq. (1) where
\[ P^{(1)} = \begin{pmatrix} f_0(x_i)S_{ij}^{-1} f_0(x_j) & \ldots & f_0(x_i)S_{ij}^{-1} f_P(x_j) \\ \vdots & \ddots & \vdots \\ f_P(x_i)S_{ij}^{-1} f_0(x_j) & \ldots & f_P(x_i)S_{ij}^{-1} f_P(x_j) \end{pmatrix} \]
\[ P^{(0)} = \begin{pmatrix} f_0(x_i)S_{ij}^{-1} y_j \\ \vdots \\ f_P(x_i)S_{ij}^{-1} y_j \end{pmatrix} \] (9)
The solution to least squares minimization can then be mapped to a QUBO problem following Eq. (5), and amenable to methods of quantum annealing.

**DISCUSSION**

**Linear Least Squares Example**

As an example, consider the following artificially generated data
\[ E[D(x)] = 8 + 4x + 7x^2 \] (10)
\[ \text{Var}[D(x)] = E[D(x)]/10 \] (11)
\[ \text{Corr}[D(x_i), D(x_j)] = 0.9|x_i - x_j| \] (12)
where \( x \in \mathbb{Z} : x \in [0, 49] \). The Toeplitz correlation matrix [32] is chosen to simulate a correlated time-series dataset, where the correlations decay exponentially as a function of \( x \). Following the notation in Eq. (7), we assume a linear fit
\[ F(x, A) = A_0 + A_1 x + A_2 x^2, \] (13)
and we estimate the parameters \( A_n \) given the data \( D(x) \). Using Eq. (3), we express each parameter \( A_i \) as a 4-bit unsigned integer
\[ A_i = \psi_{1,i} + 2\psi_{2,i} + 4\psi_{3,i} + 8\psi_{4,i} \] (14)
and construct the problem Hamiltonian following Eqs. (5) and (9). The required 12 logical spins (3 parameters \( \times \) 4-bit representation) support a total of 4096 possible solutions. Explicit evaluation finds the true ground-state to have energy \( E_0 = -1.418 \) and eigenstate
\[ \Psi_0 = (0 0 0 1 0 0 1 0 1 1 1 0) \] (15)
which corresponds to the parameter values
\[ A_0 = 8 \leftrightarrow (0 0 0 1) \] (16)
\[ A_1 = 4 \leftrightarrow (0 0 1 0) \] (17)
\[ A_2 = 7 \leftrightarrow (1 1 1 0) \] (18)

These correct coefficients for the generating function in Eq. (10) verify the design of the algorithm.

We next test the algorithm by solving the objective function using quantum annealing. The target Hamiltonian of Eq. (10) is mapped into the D-Wave 2000Q, as discussed above. Results for 100,000 independent evaluations are acquired using an annealing schedule with \( T = 200 \) \( \mu s \). The histogram (blue) and cumulative distribution (red) of solutions are presented in Fig. 1, with 0.5% of the results reproducing the ground state given by Eq. (15). In addition, the lowest 0.8% of the true eigenvalue spectrum is obtained by 10% of these solutions with the overall results biased towards the lower-lying energy eigenspectrum.

**Conditioned Systems of Linear Equations**

In this section we show results and scaling of a classical method and the quantum annealer. One of the criteria for categorizing the “difficulty” of a linear system is condition number. The condition number of a matrix is defined as the ratio of maximum and minimum singular values.
\[ \kappa(A) = \frac{\sigma_{\text{max}}(A)}{\sigma_{\text{min}}(A)} \] (19)

In the case of symmetric matrices, this is equivalent to the ratio of largest and smallest eigenvalues. All of our numerical experiments in the subsequent sub-sections are performed with symmetric matrices.

We vary our test matrices in two ways: 1) vary the rank of the matrix while holding the condition number fixed, 2) the dimension is held constant with varying condition number. A common metric to assess how accurately the linear system has been solved is the \( L_2 \)-norm of the relative residual.
\[ ||r_{\text{relative}}|| = \frac{||Ax_{\text{approx}} - b||}{||b||} \] (20)

For conjugate gradient, a tolerance for Eq. (20) is utilized as a terminating criterion and the number of iterations when this point is reached is recorded. For quantum annealing the role of the relative residual is more subtle. The annealer is run many times and the lowest energy eigenpair is returned. The eigenvector from this set is substituted for \( x_{\text{approx}} \), allowing a relative residual to be defined for the total anneal.

**Classical Solutions** – For the examples with a classical linear solver, conjugate gradient is used on a matrix of rank 12, with varying condition number. Although conjugate gradient is not the optimal choice for classically solving such systems, the scaling comparison in condition number with the quantum algorithm is informative. Fig. 2 shows slightly worse than square root scaling of conjugate gradient with condition number.
The matrices from this result are constructed by creating a random unitary matrix of rank 12, denoted as $U$. A linear spacing of real eigenvalues is chosen with the maximum equaling 5 and the minimum selected precisely to give the condition number that is desired. The matrices are trivially formed as $A = UΛU^†$. A common right-hand side is taken for all $A$: a vector of length 12 with linearly spaced decimals between 1 and -1.

**Quantum annealing** – In the following section, we demonstrate the scaling of the annealing algorithm under varying problem size, condition number, and precision of the search space. We conclude by applying the algorithm iteratively on a fixed problem and study the convergence of the relative residual.

**Problem size** – We observe that as the problem size increases linearly, the relative residual remains relatively constant. This is consistent with the notion that at fixed condition number, the range of numerical values in the result remain unchanged, and the algorithm provides constant precision under such conditions. However, for problem sizes beyond 16 linear equations, hardware limitations devoid us of the possibility of sampling the correct ground state, as demonstrated by the divergence of the annealed result versus the theoretical solution. In the on-set of this limitation, the residual grows more rapidly as expected.

In Fig. 3a, we study the scaling behavior for a fixed condition number of 1.1 using a 2-bit representation for each of $n$ parameters. Due to prior knowledge of the conjugate-gradient solution, the search space for all $n$ parameters are fixed for the set of problems, and are known to encompass the minimum and maximum results of the solution vector $x$. Additionally, knowledge of the result allows us to identify the ground-state QUBO solution by minimizing the difference between the conjugate-gradient and QUBO results (the forward error), and studies the theoretical scaling of the algorithm absent of current hardware limitations. Due to the small condition number of this study, minimizing the forward error is equivalent to minimizing the backwards error.

With increased problem size, we observe that the percentage of annealed solutions which return the ground state decreases exponentially. This indicates the solution for a dense matrix may require exponentially more evaluations to obtain for current quantum annealers. The observed scaling is consistent with the assumption that the energy gap exponentially vanishes with increasing size for a dense Hamiltonian. In particular, beyond $n = 16$, only one out of 100,000 evaluations yield the resulting annealed solution, demonstrating that the real ground-state is well beyond the reach of the available statistics.

**Condition number** – Fig. 3b demonstrates the scaling of the algorithm with respect to changing condition number. The problem size is fixed to a system of 12 linear equations, and a 2-bit precision search is employed. The condition number affects the solution vector $x$, and therefore for this study we restrict the search range to span exactly the minimum and maximum values of $x$. The chosen search range keeps the resulting relative residual approximately constant under varying condition number. For linear systems with larger condition numbers, minimizing the forward error is no longer a reliable estimate of the residual of the backwards error, and is therefore dropped from this study.

With increasing condition number, we observe that the percentage of solutions that converge to the lowest-lying state is a relatively constant value as demonstrated by a less than one order-of-magnitude change between the different examples. This behavior is in stark contrast with the scaling observed in Fig. 3a, and suggests that with increasing condition number, the ground state is exponentially easier to identify. This is in amazing contrast to the classical result from Fig. 2, in which convergence to the solution decreases as condition number is raised.

**Precision of search** – Fig. 3c explores the behavior of the algorithm as the search precision is increased for a system of four linear equations and a condition number of 1.1. We observe that the relative residual exponentially decreases, as expected due to sampling an exponential number of solutions. In contrast however, increasing the search precision linearly increases the size of the target Hamiltonian, and results in requiring exponentially more evaluations from the annealer in order to resolve the ground state. Similarly to Fig. 3a, we observe that the forward error for problem sizes beyond 5-bits of search precision starts to deviate from the backward error, an indication that the limits of hardware control have been reached.

**Iterative approach** – Finally, we explore the possibility of iteratively applying the algorithm in order to decrease the relative residual of the final solution. We demonstrate this technique on a system of 4 linear equations, with a condition number of 1.1, and perform a 4-bit search for each unknown variable. For this study, we initiate the search space ranging from -1 to 1. At the end of each iteration, we narrow the optimization to two neighboring values of the result allowed by the search space. Fig. 4 shows how the search space is refined with each iteration of the algorithm and converges to the conjugate gradient solution. Fig. 3d shows that the relative residual exponentially decreases with the application of each iteration, while the number of anneals required to sample the ground state stays relatively constant. The solution from quantum annealing at the final (ninth) iteration agrees with conjugate-gradient at single precision accuracy.

**METHODS**

**Conjugate Gradient** The method of conjugate gradient [33] is an iterative algorithm that builds an optimal polynomial for the solution within an order-$q$ Krylov subspace.

$$K_q = \text{span}\{b, Ab, A^2b, A^3b, \ldots A^{q-1}b\}$$  (21)

The first basis vector in this procedure $p_0$ is chosen to be the negative gradient of $\nabla x^†Ax - x^†b$, where $\dagger$ refers to the Hermitian conjugation. All other basis vectors $p_i \neq 0$ will be conjugate to $p_0$, thereby satisfying $p_i^†Ap_0 = 0$, thus the name Conjugate Gradient. This may be used to solve positive-definite symmetric systems or to solve the normal
The Hamiltonian is initialized as the previously well-characterized ground state of the target Hamiltonian, for example, the linear interpolation over \( s = t/T \) as
\[
H(s) = (s - 1)H^{\text{init}} + sH^{\text{problem}},
\]
where the QUBO function is expressed by the target Hamiltonian \( H^{\text{problem}} \), while the quantum state of system is initialized as the previously well-characterized ground state of the Hamiltonian \( H^{\text{init}} \). Under idealized conditions, pure-state evolution from time \( s = 0 \) to \( s = 1 \) transforms the initial quantum state to the sought-after ground state of the problem Hamiltonian. The minimum time \( T \) required to evolve the time-dependent Hamiltonian for values of \( s \in [0, 1] \) which yields a vector \( \Psi \) that is \( \epsilon \) close to the ground state of \( H(1) \) is given by the adiabatic condition [36]
\[
T \geq O \left( \frac{\|H^{\text{init}} - H^{\text{problem}}\|^2}{\epsilon \min_{s \in [0,1]} \Delta(H(s))^3} \right),
\]
where the numerator is given by the Frobenius norm, and \( \Delta(H(s)) \) is the energy gap between the instantaneous ground-state and first excited-state eigenvalues. As a result, the time complexity of adiabatic quantum computing depends on the scaling of the eigenvalue spectrum of \( H(s) \). However, the time-dependent spectrum for a generic \( H(s) \) is \textit{a priori} unknown and experimental implementations must use a value for \( T \) that meets desired tolerances.

Because of the absence of detailed Hamiltonian information, quantum annealing is assumed to operate under a relaxation of the condition for adiabatic time-evolution. This relaxation permits coupling of the lowest-lying energy level with higher-lying states, which compromises the promise of terminating in the ground state. In addition, quantum annealing operates at non-zero temperature and may be influenced by external perturbations such as fluctuating control fields and poorly-characterized Hamiltonian interactions. As a result, the ground state is typically prepared with non-unit probability by quantum annealing, and the process must be repeated multiple times so that a distribution of solutions representing possible eigenvalues and eigenstates is generated. Notably, QA often resolves the low-lying spectrum of \( H^{\text{problem}} \), and for a generic problem, the interesting statistic is to measure the percentage of measurements which resolve the ground-state, and serves as the QA analogue of Eq. (23).

We use quantum annealing to solve systems of equations by mapping the corresponding QUBO to an Ising model Hamiltonian. Due to hardware topology, we restrict our applications to only linear systems. We validate the algorithm using the D-Wave 2000Q commercial quantum annealer. This hardware is based on cryogenically cooled superconducting electronic elements that implement a programmable Ising model. Each quantum register element expresses a single Ising spin variable, but the D-Wave 2000Q supports only a limited connectivity between these elements. The underlying connectivity pattern is represented by the Chimera graph shown in Fig. 5 [37]. In particular, the \( i \)-th spin variable may be assigned a bias \( h_i = Q_{ii} \) and can be coupled to a unique set of six neighboring registers through the coupling \( J_{ij} = Q_{ij} \). A densely connected Hamiltonian can be embedded into the sparsely connected Chimera topology by using secondary constraints to build chains of strongly correlated elements in which \( J_{ij}^{\text{constraint}} \gg J_{ij}^{\text{Ising}} \). This coupling constraint favors chains of spin elements which behave as a single spin variable [38].

Previous studies have identified the optimal mappings for a fully-connected graph into the Chimera structure [39, 40]. For the D-Wave 2000Q quantum annealer, approximately 64 logical spin variables may be represented within the 2048 physical spin elements. Our subsequent examples use the dwave-sapi2 Python library [41], which is a software tool kit that facilitates cloud access to the 2000Q and supports a heuristic embedding method for the available hardware.

**DATA AVAILABILITY**

This is a theoretical paper and there is no experimental data available beyond the example data described in the paper.

**ACKNOWLEDGMENTS**

We thank Hidetoshi Nishimori (Tokyo Tech), Tetsuo Hatsuda (RIKEN), and Chih-Chieh Chen (ITRI) for discussions. This work was performed under the auspices of the U.S. Department of Energy by LLNL under Contract No. DE-AC52-07NA27344 (AG). Access to the D-Wave 2000Q computing system was provided by Oak Ridge National Laboratory. TSH acknowledges support from the Department of Energy, Office of Science, Early Career Research Project and the ORNL Directed Research and Development funds. This manuscript has been authored by UT-Battelle, LLC, under Contract No. DE-AC05-00OR22725 with the U.S. Department of Energy. The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this manuscript, or allow others to do so, for the United States Government purposes. The Department of Energy will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan.

**AUTHOR CONTRIBUTIONS**

Initial idea was proposed by C.C.C. Design of test cases was done by A.G. and C.C.C. Conjugate-gradient solves was done by A.G.. Calculation done on the DWave 2000Q
was performed by C.C.C. and T.S.H.. Cross checks with density matrix renormalization group was performed by S.S. and C.C.C.. All authors contributed to writing and editing of the final manuscript.

ADDITIONAL INFORMATION

Competing Interests: The authors declare no competing interests.

[1] A. W. Harrow, A. Hassidim, and S. Lloyd, Physical Review Letters 103, 150502 (2009), arXiv:0811.3171 [quant-ph].
[2] Y. Zheng, C. Song, M.-C. Chen, B. Xia, W. Liu, Q. Guo, L. Zhang, D. Xu, H. Deng, K. Huang, et al., Physical review letters 118, 210504 (2017).
[3] Y. Subasi, R. D. Somma, and D. Orsucci, arXiv preprint arXiv:1805.10549 (2018).
[4] J. Wen, X. Kong, S. Wei, T. Xin, B. Wang, K. Li, Y. Zhu, and G. Long, arXiv preprint arXiv:1806.03295 (2018).
[5] A. Borle and S. J. Lomonaco, ArXiv e-prints (2018), arXiv:1809.07649 [quant-ph].
[6] E. Farhi, J. Goldstone, S. Gutmann, and M. Sipser, eprint arxiv:quant-ph/0001106 (2000), quant-ph/0001106.
[7] T. Kadowaki and H. Nishimori, Phys. Rev. E 58, 5355 (1998), cond-mat/9804280.
[8] E. Farhi, J. Goldstone, S. Gutmann, J. Lapan, A. Lundgren, and D. Preda, Science 292, 472 (2001), quant-ph/0104129.
[9] N. Dickson, M. Johnson, M. Amin, R. Harris, F. Altomare, A. Berkley, P. Bunyk, J. Cai, E. Chapple, P. Chavez, F. Ciocata, T. Cirip, P. Debuen, M. Drew-Brook, C. Enderud, S. Gildert, F. Hamze, J. Hilton, E. Hoskinson, K. Karimi, E. Ladizinsky, N. Ladizinsky, T. Lanting, T. Mahon, R. Neufeld, T. Oh, I. Perminov, C. Petroff, A. Przybysz, C. Rich, P. Spear, A. Tcaciuc, M. Thom, E. Tolkacheva, S. Uchaikin, J. Wang, A. Wilson, Z. Merati, and G. Rose, Nature Communications 4 (2013), 10.1038/ncomms2920.
[10] M. H. Amin, D. V. Averin, and J. A. Nesteroff, Phys. Rev. A 79, 022107 (2009), arXiv:0708.0384 [cond-mat.mes-hall].
[11] T. Albash and D. A. Lidar, Phys. Rev. A 91, 062320 (2015), arXiv:1503.08767 [quant-ph].
[12] W. Wang, J. Machta, and H. G. Katzgraber, Phys. Rev. B 93, 224414 (2016), arXiv:1603.00543 [cond-mat.dis-nn].
[13] K. Nishimori, H. Nishimori, A. J. Ochoa, and H. G. Katzgraber, Phys. Rev. E 94, 032105 (2016), arXiv:1605.03303 [cond-mat.dis-nn].
[14] J. Marshall, E. G. Rieffel, and I. Hen, Physical Review Applied 8, 064025 (2017), arXiv:1703.03902 [quant-ph].
[15] T. F. Rønnow, Z. Wang, J. Job, S. Boixo, S. V. Isakov, D. Wecker, J. M. Martinis, D. A. Lidar, and M. Troyer, Science 345, 420 (2014).
[16] H. G. Katzgraber, F. Hamze, and R. S. Andrist, Physical Review X 4, 021008 (2014).
[17] V. S. Denchev, S. Boixo, S. V. Isakov, N. Ding, R. Babbush, V. Smelyanskiy, J. Martinis, and H. Neven, Physical Review X 6, 031015 (2016).
[18] S. Mandrà, Z. Zhu, W. Wang, A. Perdomo-Ortiz, and H. G. Katzgraber, Phys. Rev. A 94, 022337 (2016).
[19] T. Albash and D. A. Lidar, Phys. Rev. X 8, 031016 (2018).
[20] S. Mandrà and H. G. Katzgraber, Quantum Science and Technology 3, 04LT01 (2018).
[21] A. Stathopoulos and K. Orginos, SIAM J. Sci. Comput. 32, 439 (2010), arXiv:0707.0131 [hep-lat].
[22] A. Abdel-Rehim, K. Orginos, and A. Stathopoulos, Proceedings, 27th International Symposium on Lattice field theory (Lattice 2009): Beijing, P.R. China, July 26-31, 2009, PoS LAT2009, 036 (2009), arXiv:0911.2285 [hep-lat].
[23] M. Lüscher, Journal of High Energy Physics 2007, 081 (2007).
[24] A. Brandt (2008).
[25] J. Brannick, R. C. Brower, M. A. Clark, J. C. Osborn, and C. Rebbi, Phys. Rev. Lett. 100, 041601 (2008), arXiv:0707.4018 [hep-lat].
[26] A. Frommer, K. Kahl, S. Krieg, B. Leder, and M. Rottmann, SIAM Journal on Scientific Computing, 36 (2013).
[27] M. A. Clark, B. Joó, A. Strelchenko, M. Cheng, A. Gambhir, and R. Brower, (2016), arXiv:1612.07873 [hep-lat].
[28] E. Boros and P. L. Hammer, Dis. App. Math. 123, 155 (2002).
[29] E. Boros and A. Gruber, in ISAIM (2012).
[30] T. S. Humble, A. J. McCaskey, R. S. Bennink, J. J. Billings, E. D’Azevedo, B. D. Sullivan, C. F. Klymk, and H. Seddiqui, Computational Science and Discovery 7, 015006 (2014).
[31] A. C. Aitken, Proceedings of the Royal Society of Edinburgh 55, 42–48 (1936).
[32] J. Hardin, S. R. Garcia, and D. Golan, ArXiv e-prints (2011), arXiv:1106.5834 [math.ST].
[33] M. R. Hestenes and E. Stiefel, Journal of Research of the National Bureau of Standards 49, 409 (1952).
[34] C. Chang et al., Nature 558, 91 (2018), arXiv:1805.12130 [hep-lat].
[35] D. B. Kaplan, Phys. Lett. B288, 342 (1992), arXiv:hep-lat/9206013 [hep-lat].
[36] D. Aharonov, W. van Dam, J. Kempe, Z. Landau, S. Lloyd, and O. Regev, eprint arxiv:quant-ph/0405098 (2004), quant-ph/0405098.
[37] P. I. Bunyk, E. M. Hoskinson, M. W. Johnson, E. Tolkacheva, F. Altomare, A. J. Berkley, R. Harris, J. P. Hilton, T. Lanting, A. J. Przybysz, and J. Whittaker, IEEE Transactions on Applied Superconductivity 24, 2318294 (2014), arXiv:1401.5504 [quant-ph].
[38] V. Choi, ArXiv e-prints (2008), arXiv:0804.4884 [quant-ph].
[39] V. Choi, Quantum Inf. Process. 7, 193 (2008).
[40] C. F. Klymk, B. D. Sullivan, and T. S. Humble, Quantum Inf. Process. 13, 709 (2014).
[41] D.-W. S. Inc., “Developer guide for python,” (2018).
[42] D.-W. S. Inc., “Qubist,” (2018).
Figure 1 | Histogram and cumulative distributions for solutions to the linear least squares example problem. The solutions are discrete by construction and no binning is performed to create the histogram. A total of 100,000 measurements are performed on the quantum annealer.

Figure 2 | The number of conjugate gradient iterations grows slightly worse than $\sqrt{\kappa(A)}$. The stopping criterion is a tolerance of $10^{-6}$ for the norm of the relative residual. All matrices are rank 12, with smaller eigenvalues as $\kappa(A)$ increases, but identical eigenvectors. The same right-hand side is solved for all cases.
Figure 3 | a, (Top) The black dashed line shows the theoretical minimum relative residual of the algorithm as predicted by minimizing the forward error, given the search precision and search range used for the test. The red crosses are results of the lowest energy state from 100,000 quantum annealing measurements. Physical measurements deviate from the theoretical minimum as problem size grows. (Bottom) The corresponding percentage of measurements in the minimum energy state are shown in blue. The vertical axis is shown in a logarithmic scale.

b, Analogous to Fig. 3a but for varying condition number. The forward error prediction (dashed black line) is omitted and not a reliable measure of the relative residual for larger condition numbers. Note that the vertical axis of the bottom plot showing the percentage of measurements observed in the lowest-lying state is on a linear scale.

c, This plot is analogous to Fig. 3a but for varying search precision. Note that both vertical axes are on a logarithmic scale.

d, (Top) The relative residual exponentially decreases with each iteration of the algorithm. By the ninth iteration the result reaches single precision. (Bottom) The percentage of quantum annealing solutions in the lowest-lying state. The algorithm successfully resolves the solution at single precision for this example without issue.
**Figure 4** | Search region over successive iterations of the algorithm. The shaded regions indicate the search region for each component of the solution vector $x$. The gray horizontal line is the result from a conjugate-gradient solver. a, Search region for the first five iterations for all parameters. b, Search region zoomed in on a single parameter reaching single precision accuracy.

**Figure 5** | Example Chimera graph for a D-Wave 2000Q Quantum Computer solving a system of 12 linear equations. The lattice sites are qubits, the values on the sites set the biases $h_i$, and the connections set the couplers $J_{ij}$. The figure is generated with the online solver visualizer provided by D-Wave Qubist [42].