Practical Quantum Algorithms for Cryptanalysis

Ben Priestley

January 19, 2024

1 Quantum-accelerated Sieving for Integer Factorisation

1.1 Claus Peter Schnorr (2021)'s classical factoring algorithm

Previously, Claus-Peter Schnorr had proposed a sieving-based classical factorisation algorithm that reduces the problem of finding useful smooth-relation pairs (sr-pairs) to a closest vector problem (CVP) on the prime lattice (C. P. Schnorr 1991; Claus Peter Schnorr 2013), and the approach has since been extensively explored by the scientific community (e.g. see Vera 2010). The dimension of the lattice was shown to scale polynomially in the bit-length n of the integer to be factored, which serves as a substantial bottleneck for the method to have any serious applications in breaking modern-day factoring-based cryptosystems.

In follow-up work, Claus Peter Schnorr (2021) claimed to have reduced the required dimension of the lattice to $\mathcal{O}(n/\log n)$. The results are highly debated and Ducas (2021) performed an experimental analysis of the algorithm, indicating that the proposed claims do not hold. We now outline the proposed algorithm as detailed in Grebnev et al. (2023):

1. Construct a factor base $P_n := \{p_i\}_{i=0,\dots,n}$ of the first n primes for some fixed n (and with $p_0 := -1$ to enable negative representation)[1].
2. Generate a set of $n+1$ useful relations $\{(u_j, v_j)\}_{j=1,\dots,n+1}$ with respect to P_n. A pair of integers (u, v) are called a *useful relation* if u and $u - vN$ are p_n-smooth, and an integer u is called p_n-smooth if all of its prime factors belong to P_n; that is, $u = \prod_{i=0}^{n} p_i^{e_i}$ for some non-negative integers e_i.
3. Construct matrices $E = (e_{i,j})$ and $E' = (e'_{i,j})$ with elements taken from the factorsiations of $u_j = \prod_{i=0}^{n} p_i^{e_{i,j}}$ and $u_j - Nv_j = \prod_{i=0}^{n} p_i^{e'_{i,j}}$.
4. Find solutions $\tau = (\tau_1, \dots, \tau_{n+1})^T \neq 0$ of the system of linear equations $(E' - E)\tau \equiv 0 \mod 2$, and for each linear independent τ, construct $X_\tau := \prod_{i=0}^{n} p^{\frac{1}{2}\sum_{j=1}^{m}(e'_{i,j} - e_{i,j})\tau_j} \mod N$ which definitely satisfies $X_\tau^2 \equiv 1 \mod N$.
5. According to the classic Fermat's idea, if it appears that $X_\tau \not\equiv \pm 1 \mod N$, then non-trivial factors of N can be calculated efficiently as $\gcd(X_\tau \pm 1, N)$, where $\gcd(\cdot, \cdot)$ denotes the greatest common divisor.

Claus Peter Schnorr (2021)'s work proposes a lattice-based approach to the collection of useful relations (step 2 above) by considering the solution $\vec{e} = \sum_{i=1}^{n+1} e_i \vec{b}_i$ to the shortest vector problem (SVP) of a lattice $\Lambda \subset \mathbb{Z}^{n+1}$ given by the basis matrix

$$\begin{pmatrix} \vec{b}_1 & \dots & \vec{b}_{n+1} \end{pmatrix} := \begin{pmatrix} f(1) & 0 & \dots & 0 & 0 \\ 0 & f(2) & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & f(n) & 0 \\ N' \ln p_1 & N' \ln p_2 & \dots & N' \ln p_n & N' \ln N \end{pmatrix},$$

where $N' := N^{1/(n+1)}$, p_i is the i-th prime, and $f : [1, \dots, n] \to [1, \dots, n]$ is a random permutation. It is claimed that (u, v) defined as $u := \prod_{e_i > 0} p_i^{e_i}$ and $v := \prod_{e_i > 0} p_i^{-e_i}$ form an sr-pair with high probability, and hence we may solve the factorisation problem by solving a number of these

[1] It is assumed that $p_n \ll \sqrt{N}$, and thus N is not itself p_n-smooth (otherwise N could be easily factored in polynomial time)

SVP problems, which themselves are randomised via the permutation f on the diagonal. With a sufficiently large n, sufficiently many SVPs may be generated and solved, and so sufficiently many useful relations can be obtained to form a system of equations.

The construction of the SVP is not restricted to the above; we may scale the rightmost row in any number of ways to modify the lattice's construction (Ducas 2021; Grebnev et al. 2023). To align with Yan et al. (2022), we instead consider the closest vector problem (CVP) formulation, as also proposed more generally in Claus Peter Schnorr (2021), given by the lattice and target vector

$$B_{n,c} := \begin{pmatrix} f(1) & 0 & \dots & 0 \\ 0 & f(2) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & f(n) \\ \lfloor 10^c \ln p_1 \rceil & \lfloor 10^c \ln p_2 \rceil & \dots & \lfloor 10^c \ln p_n \rceil \end{pmatrix} \quad \text{and} \quad t := \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 10^c \ln N \end{pmatrix},$$

where $\lfloor \cdot \rceil$ denotes the rounding operation, $f(i)$ for $i = 1, \dots, n$ is now a random permutation of $(\lceil 1/2 \rceil, \lceil 2/2 \rceil, \dots, \lceil n/2 \rceil)$, and c is a "precision parameter" that is conjectured to be positively correlated with the probability of finding vectors close to t (Khattar and Yosri 2023).

1.2 Yan et al. (2022)'s quantum-accelerated subroutine

Yan et al. (2022) acknowledge the time-complexity bottleneck of Claus Peter Schnorr (2021) – namely, that finding sr-pairs by solving a CVP is a slow and arduious task – and offer a quantum-accelerated subroutine to solve the CVP more rapidly. Unfortunately, they reuse the weak claim of Claus Peter Schnorr (2021) that a sublinear lattice scheme would be sufficient, and so propose a "sublinear-resource quantum integer factorisation (SQIF)" algorithm that is the subject of intense debate (Aaronson 2023; Khattar and Yosri 2023; Atallah et al. 2023; Grebnev et al. 2023)

The SQIF algorithm accelerates Claus Peter Schnorr (2021)'s algorithm by utilising a quantum approximate optimisation algorithm (QAOA) subroutine to quickly solve the CVP and thus speed up the slowest part of the process. Their acceleration is implemented as follows:

1. Apply Babai's algorithm (Babai 1986) to find a lattice vector $b_{op} \in \Lambda(B_{n,c})$ which is *approximately* closest to the target vector t. The details of this are yet another layer deep of unnecessary, so we'll summarise the relevant parts:
   (a) Perform lattice reduction (LLL-reduction due to A. K. Lenstra, H. W. Lenstra, and Lovász 1982, to be specific), resulting in a reduced basis $D = [d_1 \cdots d_n]$ and corresponding Gram-Schmidt orthogonal basis $\tilde{D} = [\tilde{d}_1 \cdots \tilde{d}_n]$.
   (b) Perform a "size-reduction" of the target vector t using the reduced basis. We will then have an approximately closest vector $b_{op} = \sum_{i=1}^{n} c_i d_i$ where the coefficient $c_i = \lceil \mu_i \rfloor = \lceil \langle d_i, \tilde{d}_i \rangle / \langle \tilde{d}_i, \tilde{d}_i \rangle \rfloor$ is obtained by rounding to the nearest integer to the Gram-Schmidt coefficient μ_i (since our lattice must be an integer combination).

2. Define the all-to-all connected Ising spin-glass Hamiltonian H_P capturing the quality of the neighbouring solutions v_{new} in the unit hypercube centered on b_{op} (in the reduced basis D). The quality of each $v_{new} = \sum_{i=1}^{n} (c_i + \text{sign}(\mu_i - c_i) z_i) d_i = \sum_{i=1}^{n} (c_i + \lceil \mu_i - c_i \rceil z_i) d_i$ is given by its Euclidean distance to t, and hence we have a QUBO problem over the steps[2] z_i;

$$F(z_1, \dots, z_n) = \|t - v_{new}\|^2 = \left\| t - b_{op} + \sum_{i=1}^{n} \lceil \mu_i - c_i \rceil z_i d_i \right\|^2. \tag{1}$$

   This can then be mapped to the Hamiltonian using $z_i \mapsto \hat{z}_i = (I - \sigma_z^i)/2$, which gives us

$$H_P = \left\| t - b_{op} + \sum_{i=1}^{n} \lceil \mu_i - c_i \rceil \hat{z}_i d_i \right\|^2 = \sum_{j=1}^{n+1} \left| t_j - b_{op}^j + \sum_{i=1}^{n} \lceil \mu_i - c_i \rceil \hat{z}_i d_{i,j} \right|^2, \tag{2}$$

   where σ_z^i denotes the Pauli-Z operator $|0\rangle\langle 0| - |1\rangle\langle 1|$ acting on the i-th qubit.

3. Sample low-energy eigenstates by QAOA (Farhi, Goldstone, and Gutmann 2014) to yield an improved solution to the CVP.

[2] We have written an equivalent Hamiltonian to Yan et al. (2022), but we include the direction of the step directly in the definition of v_{new}. The original formulation uses $\|t - v_{new}\|^2 = \|t - b_{op} - \sum_{i=1}^{n} z_i d_i\|^2$ with z_i being conditionally in $\{0, +1\}$ or $\{0, -1\}$ on the basis of its rounding. This is convoluted, so we explicitly write the condition into equation (2).

2 The Insufficient Scalability of the Quantum Subroutine

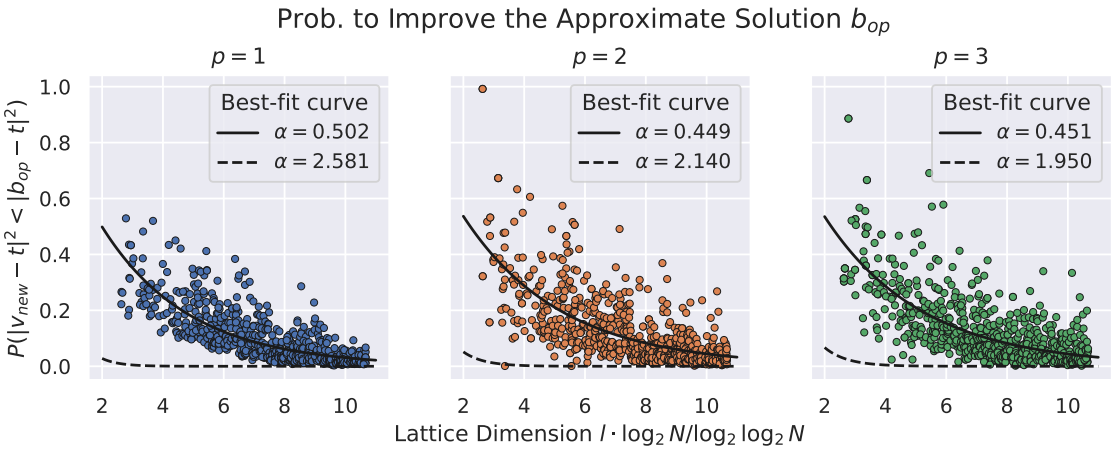

Figure 1: Caption

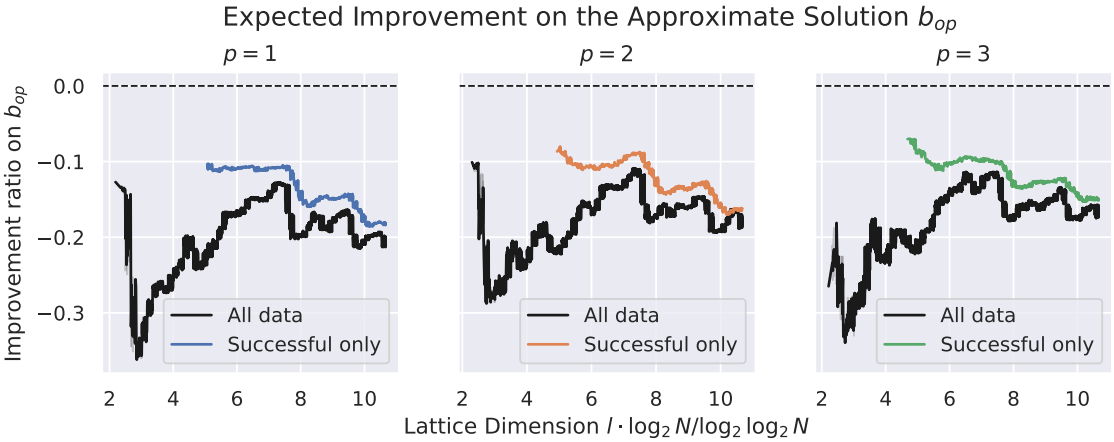

Figure 2: Caption

3 Theoretical Arguments Against a Sublinear Lattice

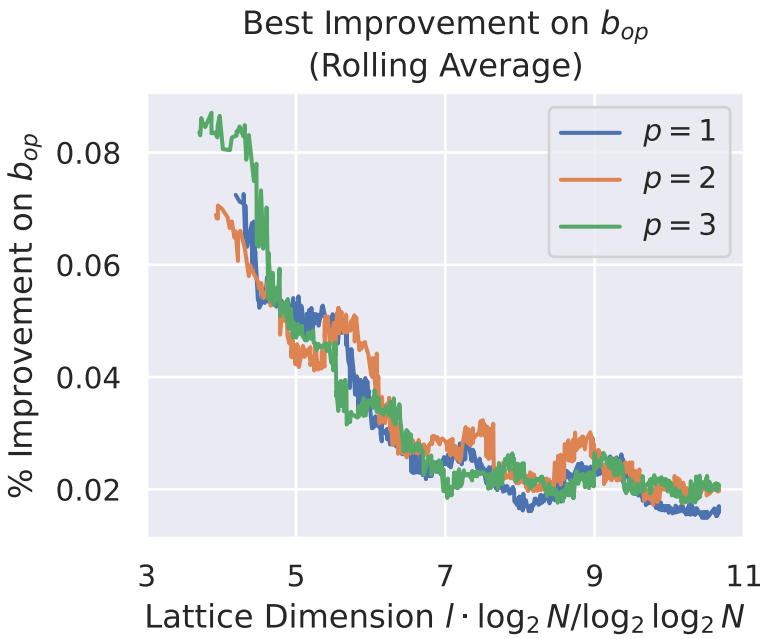

Figure 3: Caption

References

Lenstra, Arjen K, Hendrik Willem Lenstra, and László Lovász (1982). "Factoring polynomials with rational coefficients". In: *Mathematische annalen* 261.ARTICLE, pp. 515–534.

Babai, László (1986). "On Lovász'lattice reduction and the nearest lattice point problem". In: *Combinatorica* 6, pp. 1–13.

Schnorr, C. P. (1991). "Factoring Integers and Computing Discrete Logarithms via Diophantine Approximation". In: *Advances in Cryptology — EUROCRYPT '91*. Ed. by Donald W. Davies. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 281–293. ISBN: 978-3-540-46416-7.

Vera, Antonio Ignacio (2010). "A Note on Integer Factorization Using Lattices". In: *arXiv preprint arXiv:1003.5461*.

Schnorr, Claus Peter (2013). "Factoring integers by CVP algorithms". In: *Number Theory and Cryptography: Papers in Honor of Johannes Buchmann on the Occasion of His 60th Birthday*, pp. 73–93.

Farhi, Edward, Jeffrey Goldstone, and Sam Gutmann (2014). "A quantum approximate optimization algorithm". In: *arXiv preprint arXiv:1411.4028*.

Peruzzo, Alberto et al. (2014). "A variational eigenvalue solver on a photonic quantum processor". In: *Nature communications* 5.1, p. 4213.

McClean, Jarrod R, Jonathan Romero, et al. (2016). "The theory of variational hybrid quantum-classical algorithms". In: *New Journal of Physics* 18.2, p. 023023.

McClean, Jarrod R, Sergio Boixo, et al. (2018). "Barren plateaus in quantum neural network training landscapes". In: *Nature communications* 9.1, p. 4812.

Variational Principle (Aug. 2020). [Online; accessed 2023-10-18].

Cerezo, Marco et al. (2021). "Variational quantum algorithms". In: *Nature Reviews Physics* 3.9, pp. 625–644.

Ducas, Léo (2021). *Lducas/schnorrgate: Testing Schnorr's factorization claim in sage*. URL: https://github.com/lducas/SchnorrGate.

Schnorr, Claus Peter (2021). *Fast Factoring Integers by SVP Algorithms, corrected*. Cryptology ePrint Archive, Paper 2021/933. https://eprint.iacr.org/2021/933. URL: https://eprint.iacr.org/2021/933.

Fontana, Enrico et al. (Sept. 2022). "Non-trivial symmetries in quantum landscapes and their resilience to quantum noise". In: *Quantum* 6, p. 804. DOI: 10.22331/q-2022-09-15-804. URL: https://doi.org/10.22331%2Fq-2022-09-15-804.

Yan, Bao et al. (2022). *Factoring integers with sublinear resources on a superconducting quantum processor*. arXiv: 2212.12372 [quant-ph].

Aaronson, Scott (Jan. 2023). *Cargo Cult Quantum Factoring*. URL: https://scottaaronson.blog/?p=6957.
Atallah, Mostafa et al. (2023). *Integer Factorization through Func-QAOA*. arXiv: 2309.15162 [quant-ph].
Grebnev, Sergey V. et al. (2023). "Pitfalls of the Sublinear QAOA-Based Factorization Algorithm". In: *IEEE Access* 11, pp. 134760–134768. ISSN: 2169-3536. DOI: 10.1109/access.2023.3336989. URL: http://dx.doi.org/10.1109/ACCESS.2023.3336989.
Khattar, Tanuj and Nour Yosri (2023). "A comment on" Factoring integers with sublinear resources on a superconducting quantum processor"". In: *arXiv preprint arXiv:2307.09651*.
IBM (n.d.). *Variational algorithm design*. URL: https://learning.quantum-computing.ibm.com/course/variational-algorithm-design.

A More Details on VQA Design

We'll now give some more detail about the components of a VQA, how they interact, and the theory that brings things together to give the main idea behind variational algorithms. We keep this discussion as general as possible.

Cost functions

First, let us discuss the cost function. For any set of parameters $\vec{\theta}$, the cost function should evaluate how 'good' that set is with regard to the encoded problem, hopefully with $C(\vec{\theta}_a) < C(\vec{\theta}_b)$ when $\vec{\theta}_a$ is a set of parameters that is 'closer'[3] to the best solution of the problem (which we may denote by $\vec{\theta}^*$) than $\vec{\theta}_b$, and $\min C(\vec{\theta}^*)$ should correspond to an optimal solution[4]. Implicit is the desire that the cost only be efficient to compute on a quantum device.

It is beneficial to take a brief interlude into the variational theorem before continuing, as it very much speaks to our goal of minimising the cost function.

The variational theorem

Definition A.1 (The Variational Theorem (or Principle))**.** *The ground-state energy, denoted E_0, is no greater than the expectation value of the time-independent Hamiltonian $\hat{\mathcal{H}}$, calculated with the parameterised wave function $|\psi(\vec{\theta})\rangle$ (Variational Principle 2020); that is*

$$\langle\hat{\mathcal{H}}\rangle(\vec{\theta}) := \langle\psi(\vec{\theta})|\hat{\mathcal{H}}|\psi(\vec{\theta})\rangle \geq E_0. \tag{3}$$

In the context of a variational algorithm, we can then optimise the set $\vec{\theta}$ to vary $|\psi(\vec{\theta})\rangle$ until the expectation value $\langle\hat{\mathcal{H}}\rangle(\vec{\theta})$ is minimised to obtain approximations to the wavefunction and the energy of the ground-state. Let us now develop the necessary mathematical rigor[5] to understand that this principle is true.

It is common to state the theorem in terms of energy minima as, for physical reasons, it assumes a finite lower bound must exist (even as the dimensionality tends to infinity) but that an upper bound does not generally exist. The Hamiltonian $\hat{\mathcal{H}}$ is but one observable following these constraints; we could generalise the theorem to any quantum observable provided it is also constrained in this way. We will abstract slightly to refer to to eigenvalues and eigenstates (rather than energies and energy-states) to keep the idea more general. Although, we will continue to use the Hamiltonian as it is is relevant moving forward, and so we lean on the fact that the energies of $\hat{\mathcal{H}}$ corresponding to the eigenvalues are increasing; $\lambda_0 = E_0 \leq \lambda_1 = E_1 \leq \cdots$.

Suppose, as is perhaps the best-known application of VQA (Cerezo et al. 2021; McClean, Romero, et al. 2016; Peruzzo et al. 2014), that we would like to estimate the low-lying eigenstates and eigenvalues of a given Hamiltonian $\hat{\mathcal{H}}$[6]. Then consider the spectral decomposition

$$\hat{\mathcal{H}} = \sum_{k=0}^{N-1} \lambda_k |\phi_k\rangle\langle\phi_k|, \tag{4}$$

[3] This is not closeness in the geometric sense, but in the 'quality' sense. It is a better solution because its quality is closer to that of the best solution.

[4] This corresponds to the 'faithful' and 'operationally meaningful' desiderata for cost functions outlined by Cerezo et al. (2021).

[5] See IBM (n.d.) and *Variational Principle* (2020) for more, so long as you are comfortable to return to a discussion of physics.

[6] This is a very genuine thing to want to know, supposing $\hat{\mathcal{H}}$ encodes our optimisation problem, and thus we have that $C(\vec{\theta}) = \langle\psi(\vec{\theta})|\hat{\mathcal{H}}|\psi(\vec{\theta})\rangle$ (by the variational theorem that we are yet to derive).

where N is the dimensionality of the space of states, $|\phi_k\rangle$ is the k-th eigenstate, and λ_k is the corresponding eigenvalue[7].

Using this decomposition as a substitution, we can then write the expectation value of the Hamiltonian, calculated with some trail state $|\psi\rangle$, as

$$\begin{aligned}\langle\psi|\hat{\mathcal{H}}|\psi\rangle &= \langle\psi|\left(\sum_{k=0}^{N-1}\lambda_k|\phi_k\rangle\langle\phi|\right)|\psi\rangle \\ &= \sum_{k=0}^{N-1}\lambda_k\langle\psi|\phi_k\rangle\langle\phi_k|\psi\rangle \\ &= \sum_{k=0}^{N-1}\lambda_k|\langle\psi|\phi_k\rangle|^2.\end{aligned} \tag{5}$$

Then, noticing that $\lambda_0 \leq \lambda_k$ for any k, we yield the inequality

$$\langle\psi|\hat{\mathcal{H}}|\psi\rangle = \sum_{k=0}^{N1}\lambda_k|\langle\psi|\phi_k\rangle|^2 \geq \sum_{k=0}^{N-1}\lambda_0|\langle\psi|\phi_k\rangle|^2 = \lambda_0\sum_{k=0}^{N-1}|\langle\psi|\phi_k\rangle|^2. \tag{6}$$

The set of eigenstates $\{|\phi_k\rangle\}_{k=0}^{N-1}$ form an orthonormal basis, thus the sum probability for measuring each eigenstate $\sum_{k=0}^{N-1}|\langle\psi|\phi_k\rangle|^2$ is unit (an axiom of probability). Therefore, equation (6) reduces effortlessly to

$$\langle\psi|\hat{\mathcal{H}}|\psi\rangle \geq \lambda_0\sum_{k=0}^{N-1}|\langle\psi|\phi_k\rangle|^2 = \lambda_0. \tag{7}$$

We have been purposefully vague about the trial state $|\psi\rangle$. In fact, the idea that has culminated in equation (7) holds for any normalised state, so it follows that we may consider parameterised states $|\psi(\vec{\theta})\rangle$ and the inequality will still hold, so long as $\langle\psi(\vec{\theta})|\psi(\vec{\theta})\rangle = 1$ for any choice of $\vec{\theta}$.

To connect this with the structure of a VQA, we define the cost by $C(\vec{\theta}) := \langle\psi(\vec{\theta})|\hat{\mathcal{H}}|\psi(\vec{\theta})\rangle$ and know that the minimum will always satisfy

$$\min_{\vec{\theta}} C(\vec{\theta}) = \min_{\vec{\theta}}\langle\psi(\vec{\theta})|\hat{\mathcal{H}}|\psi(\vec{\theta})\rangle \geq \lambda_0. \tag{8}$$

Recalling that our goal is to find $|\phi_0\rangle$. Well, the closest one can get to the corresponding eigenvalue λ_0 using the parameterised states $|\psi(\vec{\theta})\rangle$ is the minimum value of $C(\vec{\theta})$. We may only recover the eigenstate exactly if there exists a perfect set $\vec{\theta}^*$ such that $\hat{\mathcal{H}}|\psi(\vec{\theta}^*)\rangle = \lambda_0|\phi_0\rangle$. Otherwise, we recover the optimal approximation.

This is the variational theorem that gives rise to the family of variational algorithms. Let us now conclude this interlude and return to the components of a VQA, continuing with the ansatz.

Ansatze

The word *ansatz* is German, roughly meaning "an assumption made to help solve a problem". Mathematically, it is the combination of a reference state $|\rho\rangle = U_R|0\rangle$ and a variational form $U_V(\vec{\theta})$, defined as $U_A(\vec{\theta}) := U_V(\vec{\theta})U_R$, where we have presupposed that $|0\rangle$ is the initialised default state (IBM n.d.). Implementationally, it is a parametric quantum circuit transforming the default state (here, $|0\rangle$) to a target trial state $|\psi(\vec{\theta})\rangle$; that is, $|\psi(\vec{\theta})\rangle = U_A(\vec{\theta})|0\rangle$. In this way, it explicitly defines the search space by determining the nature of the parameters $\vec{\theta}$ and how they are trained (Cerezo et al. 2021), and thus determines the collection of possible states.

It is common practice to impose constraints on the variational form (and hence the ansatz) to reduce the number of parameters required when the dimensionality is high; the dimensionality D of the space governing the distinct quantum states that an n-qubit system may take increases exponentially ($D = 2^{2n}$). Specifically for NISQ devices, circuit depth is also a cause for concern, given the greater accumulation of noise in deeper circuits.

We will not spend any time discussing the common ansatze in the literature (see Cerezo et al. 2021), but we should distinguish between 'problem-inspired' and 'problem-agnostic' ansatze. A problem-agnoistic (or 'heuristic') ansatz is employed when we are without any specific information

[7] Physically, the k-th energy level of the system described by $\hat{\mathcal{H}}$.

for our problem. Restricting the dimensionality of the search space is more difficult without information, so a heuristic approach is the best bet. It is left to the discretion of the designer to balance quality and efficiency, with a larger parameter space taking longer to search through but more likely to contain a better solution. On the other hand, when we have some knowledge of the problem, we can be more deliberate with the reduction of dimensionality to improve efficiency without loss of accuracy by using a problem-inspired ansatz.

(Classical) optimisers

VQA is known to present its own challenges in addition to those one would expect to face anyway in an optimisation problem, such as a stochastic environment due to measurement budgets, noise, or barren plateaus (Cerezo et al. 2021; McClean, Boixo, et al. 2018). These quantum-specific challenges have inspired the development of 'quantum-aware' optimisers (Fontana et al. 2022) – classical optimisers specifically tailored to the kinds of peculiar landscapes encountered in quantum settings.

Of course, any optimiser may be used, but the success of the VQA depends very much on the quality of optimiser. At this stage, this work is not concerning itself with the optimiser, and so we spend no further time discussing it.

Putting it all together

Here is a summary to wrap things up. We refer to the specific objective of finding the minimum eigenstate (i.e. the ground state) of a system.

Suppose we have some problem that we can elegantly map to the Hamiltonian $\hat{\mathcal{H}}$ such that the cost function $C(\vec{\theta}) := \langle\psi(\vec{\theta})|\hat{\mathcal{H}}|\psi(\vec{\theta})\rangle$ describes the expectation value of $\hat{\mathcal{H}}$ with respect to the parameterised trial state $|\psi(\vec{\theta})\rangle$. In a variational algorithm, we look to minimise $C(\vec{\theta})$, evaluated here by computing $\langle\hat{\mathcal{H}}\rangle(\vec{\theta})$, which involves the use of the ansatz to yield the trail state by $|\psi(\vec{\theta})\rangle = U_A(\vec{\theta})|0\rangle$. A classical optimiser is used to search the space of solutions, but we are concerned only with the quantum aspects; it suffices to let this classical optimiser be a black-box procedure that updates the set of parameters $\vec{\theta}$ in each iteration.