

Continuous-time quantum walks for MAX-CUT are hot

Robert J. Banks¹, Ehsan Haque², Farah Nazer², Fatima Fethallah², Fatima Ruqaya², Hamza Ahsan², Het Vora², Hibah Tahir², Ibrahim Ahmed², Isaac Hewins², Ishaq Shah², Krish Baranwal², Mannan Arora², Mateen Asad², Mubasshirah Khan², Nabian Hasan², Nuh Azad², Salgai Fedaiie², Shakeel Majeed², Shayam Bhuyan², Tasfia Tarannum², Yahya Ali², Dan E. Browne³, and P. A. Warburton^{1,4}

¹London Centre for Nanotechnology, UCL, London WC1H 0AH, UK

²Newham Collegiate Sixth Form Centre, 326 Barking Rd, London, E6 2BB, UK

³Department of Physics and Astronomy, UCL, London WC1E 6BT, UK

⁴Department of Electronic & Electrical Engineering, UCL, London WC1E 7JE, UK

By exploiting the link between time-independent Hamiltonians and thermalisation, heuristic predictions on the performance of continuous-time quantum walks for MAX-CUT are made. The resulting predictions depend on the number of triangles in the underlying MAX-CUT graph. We extend these results to the time-dependent setting with multi-stage quantum walks and Floquet systems. The approach followed here provides a novel way of understanding the role of unitary dynamics in tackling combinatorial optimisation problems with continuous-time quantum algorithms.

1 Introduction

Continuous-time quantum walks (CTQWs) were first proposed for quantum computation by Farhi et al. [1]. Since then CTQWs have been applied to numerous other problems [2–8], including unstructured search, and found to recover the same scaling as Grover’s algorithm [9].

More recently CTQWs have been applied to a wider range of combinatorial optimisation problems, such as Sherrington-Kirkpatrick spin glasses and the random energy model [10]. For both of these problems CTQWs were found to have a better scaling than Grover’s algorithm. CTQWs have also been applied to MAX-2-SAT [11, 12]. The approach is relatively simple: a time-independent Hamiltonian is applied and the resulting distribution is sampled to try and find

the solution or approximate solutions to the optimisation problem. Due to the exponentially growing Hilbert-space, previous work on combinatorial optimisation problems with CTQW has largely been limited to tens of qubits [10–12], with the exception of unstructured search. However, it has been shown CTQWs can always do better than random guessing [13].

In this paper we explore the performance of CTQWs on MAX-CUT, though we expect many of the arguments to hold for a wide range of combinatorial optimisation problems. In Sec. 3 we introduce CTQWs for MAX-CUT in more depth. By exploiting the effective low dimensionality of the evolution for short run-times we explore the performance of CTQWs in this limit, independent of problem size in Sec. 4.

Given the intractability of the Schrödinger equation for large problem sizes and long times, in Sec 5 we turn to the explanatory power of the Eigenstate Thermalisation Hypothesis (ETH) to provide some insight. First proposed by Deutsch [14] and Srednicki [15], the ETH conjectures that an isolated system will reach thermal equilibrium. See [16] for a review. The ETH is a hypothesis not expected to hold for every system. Indeed it is known not to hold for integrable systems [14, 17–20] nor systems that exhibit many-body localisation [21–23]. The ETH has been observed both numerically [24–34] and experimentally [35–39] for a variety of systems. In Sec. 6 we turn the tools developed for a single-stage CTQW to multi-stage quantum walks (MSQWs) [13]. We also briefly discuss some implications of the ETH for gate-based implementation of CTQWs in Sec. 7.

Robert J. Banks: robert.banks.20@ucl.ac.uk

The key take-home of this paper is that by assuming thermalisation we can analytically reason about the behaviour of CTQWs, using a phenomenologically motivated model discussed in Sec. 5.2. CTQWs for optimisation being a relatively new and promising method for tackling combinatorial optimisation problems [10–13, 9], this approach provides physical insight where exact diagonalisation is prohibitively difficult.

We emphasise all the discussion in this paper relates to CTQWs in the closed-system setting. Throughout this paper we adopt the convention $\hbar = 1$ and $k_B = 1$. The Pauli matrices are denoted by X , Y and Z . We also make use of the python packages QuTip [40, 41] and NetworkX [42].

2 Introducing the Eigenstate Thermalisation Hypothesis

The aim of this section is to provide a very brief introduction to the ETH, focusing on the results pertinent to this paper. This section draws heavily from [16]. For a more comprehensive review the reader is referred to the original work. This section is split into two parts: the first will discuss what is meant by a steady state, while the second part will introduce the consequences of the ETH.

2.1 The steady-state

Typically, classical random walks tend to steady states [43]. Due to the oscillatory behaviour of the Schrödinger equation we generally do not expect this to be true for the state-vector of a CTQW. Instead of looking at the state-vector we might look at the expectation value of an observable \mathcal{O} , with initial state $|\psi_i\rangle$ and Hamiltonian H :

$$\begin{aligned} \langle \mathcal{O}(t) \rangle &= \langle \psi_i | e^{iHt} \mathcal{O} e^{-iHt} | \psi_i \rangle \\ &= \sum_{m,n} e^{i(E_m - E_n)t} \langle \psi_i | E_m \rangle \langle E_m | \mathcal{O} | E_n \rangle \langle E_n | \psi_i \rangle \\ &= \sum_{E_m = E_n} \langle \psi_i | E_m \rangle \langle E_m | \mathcal{O} | E_n \rangle \langle E_n | \psi_i \rangle + \\ &\quad \sum_{E_m \neq E_n} e^{i(E_m - E_n)t} \langle \psi_i | E_m \rangle \langle E_m | \mathcal{O} | E_n \rangle \langle E_n | \psi_i \rangle, \end{aligned} \quad (1)$$

where $|E_k\rangle$ denotes an eigenstate of H . It follows that if the expectation of \mathcal{O} is going to tend

to a certain value, it will tend to its long time-averaged value:

$$\begin{aligned} \langle \bar{\mathcal{O}} \rangle &:= \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \langle \psi_i | e^{iHt} \mathcal{O} e^{-iHt} | \psi_i \rangle \\ &= \sum_{E_m = E_n} \langle \psi_i | E_m \rangle \langle E_m | \mathcal{O} | E_n \rangle \langle E_n | \psi_i \rangle. \end{aligned} \quad (2)$$

From Eq. 1 it is clear that for the system to reach Eq. 2 there must be sufficient dephasing on sufficiently short time scales, such that the second sum in Eq. 1 becomes negligible compared to the first sum. In practice for any finite system we expect there to be fluctuations around the steady state value of \mathcal{O} (i.e., around $\langle \bar{\mathcal{O}} \rangle$). In short, a system is in its steady state if $\langle \mathcal{O}(t) \rangle = \langle \bar{\mathcal{O}} \rangle$ for all t after a certain value, up to negligible fluctuations in $\langle \mathcal{O}(t) \rangle$. For discussion about the time taken to reach equilibrium, please see [44].

Though Eq. 2 is useful, it requires complete knowledge of the spectrum of H , which quickly becomes intractable for large problem sizes. It is also perhaps not clear from this expression which observables correspond to steady states. The next section begins to address some of these questions.

2.2 The Eigenstate Thermalisation Hypothesis

The ETH is an attempt to explain how thermalisation occurs in a closed-system. The details can be found in [16]. A consequence is that for a large class of observables the steady state value is equal to averaging over the microcanonical ensemble in the thermodynamic limit. The fluctuations from this value are exponentially suppressed with the degrees of freedom in the system [16]. In certain cases a single eigenstate is sufficient to carry out the microcanonical averaging [15].

As mentioned prior, not all Hamiltonians exhibit ETH and not all observables thermalise [45]. Typically, the Hamiltonian is assumed to be highly non-degenerate [15, 14, 24] and to provide a locally conserved quantity [18]. The observables that exhibit thermalisation are typically considered to be local or few-body operators [15, 24]. For example, consider the highly non-local operator for a non-trivial system [18]:

$$\mathcal{O}_{jk} = -i (|E_j\rangle \langle E_k| - |E_k\rangle \langle E_j|), \quad (3)$$

where $E_j \neq E_k$. This will clearly not approach a steady state.

In summary, the ETH tells us that for large systems the steady state, exhibited by local observables, can be well approximated by the microcanonical ensemble. Though there are known exceptions, this behaviour is thought to be very common [22, 46]. Note that it is possible for a system to approach a steady state that is not a thermal state [18].

For ease of computation, throughout this paper we make use of the canonical ensemble, which is equivalent to the microcanonical ensemble in the thermodynamic limit [24, 28]. That is to say, throughout this paper we assume the pure quantum state is locally well approximated by a thermal Gibbs state:

$$\rho_\beta(H) = \frac{e^{-\beta H}}{\text{Tr } e^{-\beta H}}, \quad (4)$$

with temperature $1/\beta$.

A CTQW is a time-independent Hamiltonian and in most settings unlikely to be integrable. We might therefore expect it to exhibit thermalisation for local observables. We provide numerical evidence for this in Sec. 5.1. By exploiting the observation that the system is well approximated by a thermal state we make predictions on the performance of CTQWs.

3 Quantum walks for MAX-CUT

3.1 Set-up of the CTQW

In this section we introduce MAX-CUT for CTQWs. Given a graph $G = (V, E)$, a cut separates the nodes into two disjoint sets. The weight of the cut is given by the number of edges which connect the two sets. The aim of MAX-CUT is to find the maximum cut.

This can be encoded as finding the ground state of an Ising Hamiltonian,

$$H_p = \sum_{(i,j) \in E} Z_i Z_j. \quad (5)$$

In this paper we focus on graphs with fixed degrees and binomial graphs. To generate the binomial graphs each edge is selected with a certain probability, throughout this paper this probability is taken to be $2/3$. Binomial graphs are also known as Erdős Rényi random graphs.

As the corresponding driver Hamiltonian we take the transverse-field Hamiltonian:

$$H_d = - \sum_{i=1}^n X_i, \quad (6)$$

where $n = |V|$. This corresponds to a quantum walk on the Boolean hypercube.

The CTQW Hamiltonian is given by:

$$H_{QW} = H_d + \gamma H_p, \quad (7)$$

where $\gamma > 0$ is a parameter that needs to be set¹. The system is prepared in the ground-state of H_d , here denoted by $|+\rangle$, and evolved under H_{QW} . The resulting state, after a time t , is

$$|\psi(t)\rangle = e^{-iH_{QW}t} |+\rangle. \quad (8)$$

To assess the performance of Eq. 8 we look at $\langle H_p \rangle$. If $\langle H_p \rangle = E_0^{(p)}$, where $E_0^{(p)}$ is the ground state energy of H_p , then the CTQW has found the ground-state. Random guessing would correspond to $\langle H_p \rangle = 0$. The value of $\langle H_p \rangle$ can be related to the length of cut by:

$$\langle H_p \rangle = -2 \times (\text{Average cut value from the CTQW} - \text{Average cut value from random guessing}). \quad (9)$$

In short, the more negative $\langle H_p \rangle$ is, the better the performance of the CTQW.

Energy is conserved during the evolution. Since the evolution starts in the ground-state of $\langle H_d \rangle$, $\langle H_d \rangle$ cannot decrease initially. Assuming non-trivial dynamics, this means that $\langle H_p \rangle$ must decrease to conserve energy. Hence, why Eq. 8 will do better than random guessing [13].

A second consequence of the conservation of energy is $\langle H_{QW} \rangle = -n = E_0^{(d)}$, where $E_0^{(d)}$ is the ground-state energy of H_d . This means that the evolution cannot completely occupy the ground-state of the CTQW Hamiltonian. This is in stark contrast to a number of other popular approaches to tackling optimisation problems with NISQ algorithms [47–49].

Finally, we note the Hamiltonian in Eq. 8 commutes with the spin-flip operator:

$$G = \prod_{k=1}^n X_k. \quad (10)$$

¹Normally γ is appended to H_d not H_p . Here we take the opposite approach so changing γ does not change $\langle H \rangle$.

Since, $|+\rangle$ is an eigenstate of G with eigenvalue plus one, the evolution is restricted to the plus one eigenspace of G . This can be utilised in simulation to reduce the size of the Hilbert space by half. It also means that the CTQWs cannot reach a separable computational basis state, resulting in the ideal final state having non-zero entanglement. We also touch upon this in Appendix B in the context of the ETH. In the next section we discuss how to optimise γ to give the best possible performance.

3.2 Optimising the free parameter

The CTQW contains a free-parameter γ . Heuristically this controls the amount of dynamics present in the evolution. If γ is too large, we have approximate evolution under H_p , leaving the measurement statistics unchanged. Correspondingly $\langle H_p \rangle \approx 0$. If γ is too small we have approximate evolution under H_d , resulting in trivial dynamics and $\langle H_p \rangle \approx 0$. Clearly for a CTQW to be successful γ must be chosen somewhere between these two limits.

In this section we attempt to optimise the free parameter γ to give the best time-averaged value of $\langle H_p \rangle$. More explicitly the time-averaged value is given by:

$$\langle \bar{H}_p \rangle = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \langle H_p(t) \rangle dt. \quad (11)$$

Numerically, to evaluate $\langle \bar{H}_p \rangle$ we calculate:

$$\langle \bar{H}_p \rangle = \sum_{E_m=E_n} \langle + | E_n \rangle \langle E_m | + \rangle \langle E_n | H_p | E_m \rangle, \quad (12)$$

where $|E_k\rangle$ is an eigenstate of H_{QW} with corresponding eigenvalue E_k . The intuition behind these equations can be found in Sec. 2.

In the rest of this section we find by brute force optimisation the optimal choice of γ for different graph choices. We also make comparisons to a reasonable heuristic choice of γ . To begin with we focus on a specific example of a 12-qubit binomial graph shown in Fig. 1, before looking at statistics gathered from multiple graph instances.

Starting with a specific choice of γ for the graph instance shown in Fig. 1, Fig. 2 shows a CTQW with $\gamma = 1$. The blue line in Fig. 2 shows the average value of H_p from the Schrödinger equation, or more explicitly:

$$\langle H_p(t) \rangle = \langle + | e^{iH_{QW}t} H_p e^{-iH_{QW}t} | + \rangle. \quad (13)$$

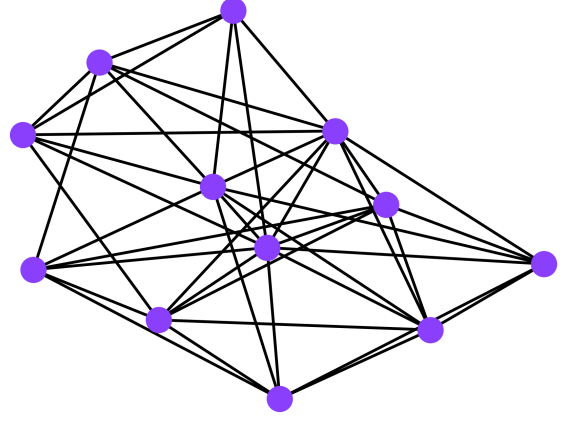


Figure 1: The single 12-qubit binomial graph instance discussed in Sec. 3.2.

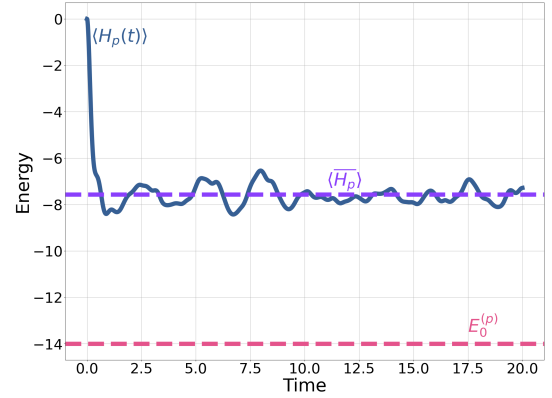


Figure 2: A time domain plot of $\langle H_p \rangle$ (the blue line) for a randomly generated 12-qubit instance shown in Fig. 1. The dashed pink line shows the minimum energy of H_p . The dashed purple line is the time-averaged value, according to Eq. 12.

The dashed pink line shows the ground state energy of H_p (i.e. $E_0^{(p)}$). For the majority of the evolution shown $\langle H_p(t) \rangle$ is fluctuating around the steady-state value, $\langle \bar{H}_p \rangle$ (the dashed purple line). We can compare this to the ground state probability, shown in Fig. 3. The ground state probability shows significant oscillations, illustrating why we avoid optimising over this quantity.

Fig. 4 shows how $\langle \bar{H}_p \rangle$ varies with γ for the same problem instance as Fig. 2. The optimal γ (i.e. γ_{opt}) occurs at the minimum value of $\langle \bar{H}_p \rangle$ (i.e., $\langle \bar{H}_p \rangle_{min}$). For the problem instance considered in Fig. 4, $\gamma_{opt} \approx 0.90$ and $\langle \bar{H}_p \rangle_{min} \approx -7.65$.

In previous works, the free parameter in a CTQW has been heuristically chosen in an attempt to maximise dynamics. One approach to

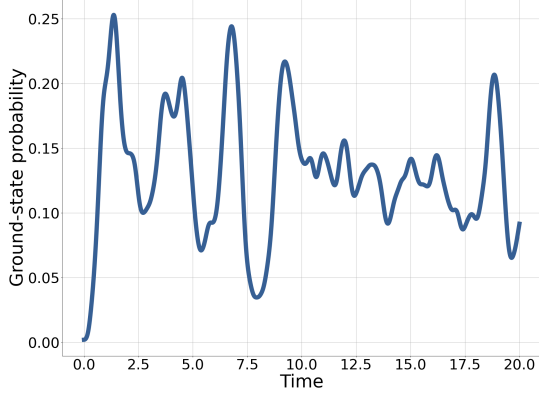


Figure 3: A time domain plot of the ground-state probability for the randomly generated 12-qubit instance shown in Fig. 1.

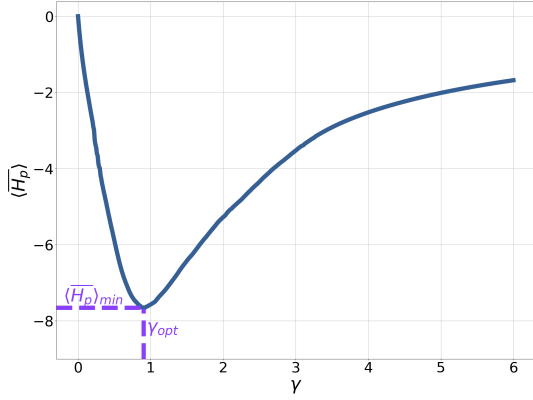


Figure 4: The time-averaged value $\langle \bar{H}_p \rangle$ as γ is changed for the 12 qubit graph shown in Fig. 1. The dashed purple line shows the location of γ_{opt} and $\langle \bar{H}_p \rangle_{min}$.

do this has been to match the energy scales of the driver and problem Hamiltonian [10, 11]. For a MAX-CUT problem, this could be interpreted as $\text{Tr} \gamma^2 H_p^2 = \text{Tr} H_d^2$ such that:

$$\gamma_{heur}^2 \kappa_2 = n, \quad (14)$$

where $\kappa_2 = |E|$ is the number of edges in the graph.

Focusing first on the performance of the binomial graph shown in Fig. 4, evaluating Eq. 14 gives $\gamma_{heur} = 0.5$, corresponding to $\langle \bar{H}_p \rangle_{min} \approx -5.88$. This is a significant reduction in the performance of the CTQW.

Fig. 5 shows the optimal γ for one hundred randomly generated problem instances for binomial graphs with problem sizes ranging between ten

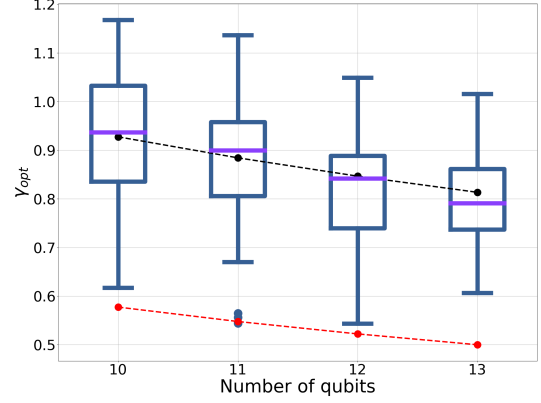


Figure 5: The optimal value of γ for one hundred problem instances for each problem size for binomial graphs. The red dashed line reflects the heuristic from Eq. 14, explicitly $\gamma = \sqrt{3/(n-1)}$. The dashed black line corresponds to fitting the curve $\gamma_{opt} = an^{-1/2}$ to the medians of the data, yielding $a \approx 2.93$.

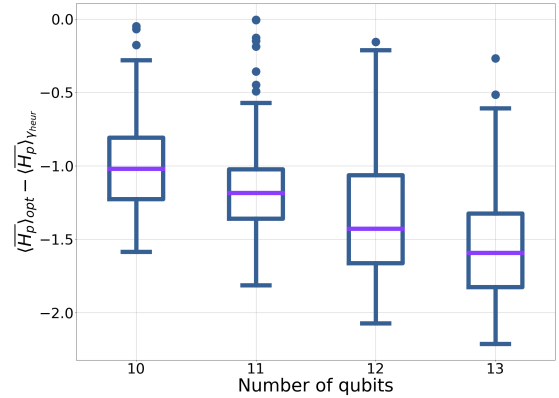


Figure 6: The difference in $\langle \bar{H}_p \rangle$ for γ_{opt} and γ_{heur} for the same binomial graphs as Fig. 5.

and thirteen qubits. The optimal γ is seen to decrease with the problem size. The range of optimal values of γ is on the order of 0.1. The dashed red line represents the heuristic in Eq. 14 assuming the number of edges is $n(n-1)/3$. Clearly Eq. 14 is significantly underestimating the optimal γ .

As mentioned in the introduction of this section, it is expected that the optimal γ will balance H_p and H_d , hence why it is reasonable to assume $\gamma \propto n^{-1/2}$. This is the same functional dependence on n as Eq. 14. The dashed black line in Fig. 5 shows $\gamma \propto n^{-1/2}$ fitted to the available data. The curve is not inconsistent with the data.

Fig. 6 shows how $\langle \bar{H}_p \rangle$ changes between γ_{opt}

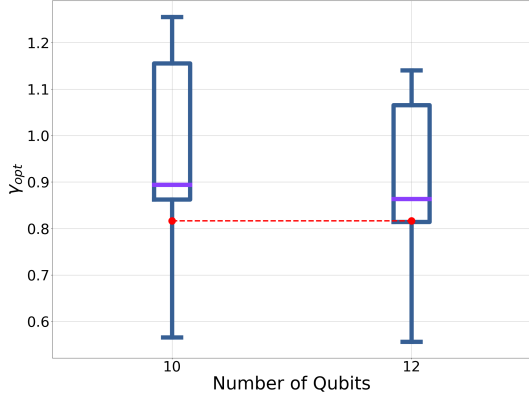


Figure 7: The optimal value of γ for three-regular graphs. In ascending problem size, there are: 16, 45 problem instances. The red dashed line is the heuristic from Eq. 14, explicitly $\gamma = \sqrt{2/3}$.

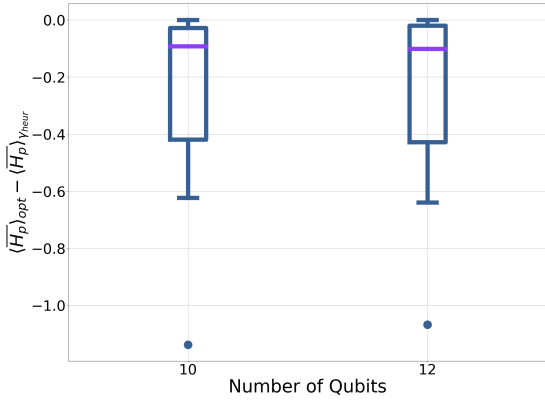


Figure 8: The difference in $\langle \bar{H}_p \rangle$ for γ_{opt} and γ_{heur} for the same three-regular graphs as Fig. 7.

and γ_{heur} for the same problem instances as Fig. 5. The difference in performance is on the order of one and appears to increase with the problem size. There is clearly scope for improvement on the heuristic choice of γ .

We now turn to three-regular graphs, Fig. 7 shows γ_{opt} for different problem sizes. For regular graphs the number of edges for a given problem size is fixed and scales with n . Hence, γ_{heur} evaluates to be the same for all instances of a d -regular graph. This is the red dashed line in the figure. For three-regular graphs γ_{heur} also appears to be generally underestimating γ_{opt} but less so than in the binomial case. This is reflected in the change in performance, as shown in Fig. 8.

In this section we have focused on finding the optimal γ . The closer γ is to the optimal value,

the better the approximation ratio is. We have also explored the use of a heuristic based around the mantra of maximising dynamics. The heuristic choice of γ has elucidated how the optimal choice of γ might scale with n . This is particularly pertinent to the discussion in Sec. 5.2. In Sec. 5.4 we show how the heuristic choice of γ (i.e. Eq. 14) can be recovered assuming thermalisation and how it might be improved upon.

An alternative, more involved, method for finding heuristic values of γ also based on the notion of maximising dynamics can be found in [13].

For the majority of this paper we are interested in the steady-state behaviour of the CTQW. However, there may be some advantage to running the CTQW on a very short time scale, with the aim of having a short time to solution. In the next section, we analytically investigate the very short time behaviour of CTQWs for MAX-CUT.

4 The very-short-time limit

A CTQW consists of a relatively simple, time-independent Hamiltonian. Following [50, 51] in this section we evaluate the torsion of the quantum evolution. The torsion is a local measure of how much the evolution deviates from a two-dimensional space. By working in this sub-space we can make short-time predictions for $\langle H_p \rangle$.

The torsion depends only on expectations of powers of H_{QW} , hence are constants over the duration of the evolution for a time-independent Hamiltonian. For the details of the calculation see Appendix A. The torsion of the evolution is given by:

$$\mathcal{T} = 2\gamma^4 \left[\kappa_2 (\kappa_2 - 1) - \frac{18\kappa_3^2}{\kappa_2} + 12\kappa_4 \right], \quad (15)$$

where κ_2 is the number of edges in the graph, κ_3 the numbers of triangles, and κ_4 the number of squares.

We can see that the torsion is reduced by the number of triangles in the graph. This suggests that the CTQW explores the solution space slower for graphs with a large number of triangles. The error from leaving a two dimensional space in a time Δt can be approximated by [50]:

$$\varepsilon_{2D} = \mathcal{T} \Delta t^4. \quad (16)$$

So for times much less than $\mathcal{T}^{-\frac{1}{4}}$ we can approximate the evolution of CTQWs by a two dimensional system. The details can be found in Appendix A. Under these assumptions, with $\gamma > 0$, the value of $\langle H_p \rangle$ can be calculated as:

$$\langle H_p \rangle = -\frac{4\kappa_2\gamma}{\omega^2} \sin^2 \omega t, \quad (17)$$

where

$$\omega^2 = \gamma^2\kappa_2 + \frac{(2\kappa_2 + 3\gamma\kappa_3)^2}{\kappa_2^2}. \quad (18)$$

As we can see, for very short times, $\langle H_p \rangle$ depends very little on the properties of the underlying MAX-CUT graph, depending primarily on the number of edges in the graph. The frequency term ω does depend on the number of triangles but for large problem sizes it is reasonable to expect $\omega^2 \approx \gamma^2\kappa_2$. Physically, it is reasonable, that at short times the CTQW only sees triangles and edges. Longer evolution would result in the approach seeing larger loops in the graph.

From a computational point of view this suggests that any useful short-time sampling must not be on a timescale much smaller than $\mathcal{T}^{-\frac{1}{4}}$ since you would like the CTQW to see the whole graph. A rough estimate for a CTQW to see the whole structure of the graph might be $l\mathcal{T}^{-\frac{1}{4}}$, where l is the length of the largest loop in the problem graph. This is bounded from above by $n\mathcal{T}^{-\frac{1}{4}}$.

In Fig. 9 we compare Eq. 17 to numerical simulation with a 12-qubit binomial graph with $\gamma = 0.05$. The dashed purple line shows the location of $\mathcal{T}^{-\frac{1}{4}}$. For short times there is good quantitative agreement between the numerical simulation (the blue line) and the two-level prediction (the pink line). At longer times there is reasonable qualitative agreement, with the two-dimensional approximation capturing the oscillatory nature.

Examining a different 12-qubit binomial graph with $\gamma = 1$ gives Fig. 10. Again, we can see good agreement between the numerical simulation (the blue line) and the two-level prediction (the pink line) for short times. The dashed purple line corresponds to $\mathcal{T}^{-\frac{1}{4}}$ with the dashed green line corresponding to $n\mathcal{T}^{-\frac{1}{4}}$. Here the two-level approximation provides poor qualitative insight into the CTQW outside of the short time limit, with the CTQW approaching an approximate steady

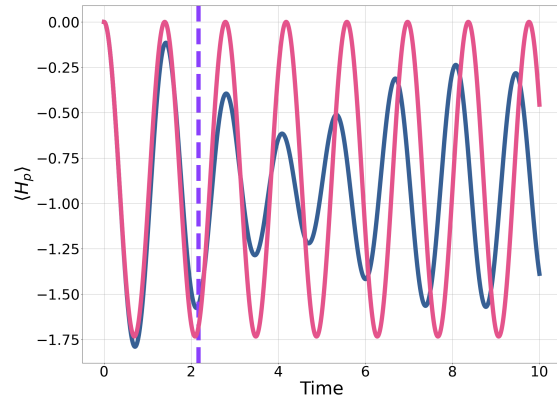


Figure 9: The performance of a CTQW on a 12 qubit binomial graph with $\gamma = 0.05$. The blue line shows the result of direct integration of the Schrödinger equation. The pink line shows the result of the two-dimensional approximation (i.e, Eq. 17) The dashed purple line shows the location of $\mathcal{T}^{-\frac{1}{4}}$.

state. Noticeably, by $n\mathcal{T}^{-\frac{1}{4}}$, the system has settled into the steady-state.

In this section we have explored the short-time limit of CTQW for MAX-CUT. The argument has relied on a geometric property of the quantum evolution (i.e., the torsion). By integrating the Schrödinger equation in the reduced two-dimensional sub-space we have made a prediction on the performance of CTQWs at short times, without limiting the analysis to small problem sizes. For longer times, Eq. 17 breaks down. In general we are unable to integrate the Schrödinger equation for arbitrarily large problem sizes. For long times we turn to the Eigenstate Thermalisation Hypothesis (ETH).

5 Continuous-time Quantum walks in the Thermodynamic limit

As described in Sec. 2 a closed quantum system under evolution by a constant Hamiltonian can result in a state that is locally indistinguishable from a thermal state. In terms of CTQWs we are primarily interested in $\langle \bar{H}_p \rangle$, a sum over local observables. Thus if a CTQW close to the optimal γ exhibits thermalisation, then $\langle \bar{H}_p \rangle$ should correspond to that of a thermal state. In Sec. 5.1 we provide numerical evidence that CTQWs do exhibit thermalisation even for small problem sizes. The aim of the rest of the section is to use this

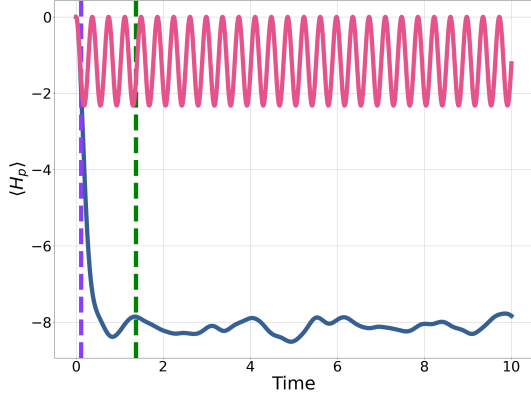


Figure 10: The performance of a CTQW on a 12 qubit binomial graph with $\gamma = 1$. The blue line shows the result of direct integration of the Schrödinger equation. The pink line shows the result of the two-dimensional approximation (i.e, Eq. 17) The dashed purple line shows the location of $\mathcal{T}^{-\frac{1}{4}}$. The dashed green line shows the location of $n\mathcal{T}^{-\frac{1}{4}}$, with $n = 12$.

insight to estimate the optimal choice of γ , and ideally the associated value of $\langle \bar{H}_p \rangle$. To do this, in Sec. 5.2 a model for the density of states (DOS) is provided. The DOS is then used to predict the temperatures and entropy of entanglement associated with CTQWs in Sec. 5.3. In Sec. 5.4 we use the model to estimate $\langle \bar{H}_p \rangle$ and to provide estimates for the optimal choice of γ .

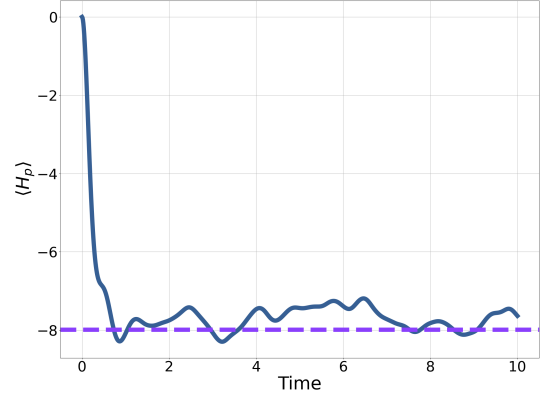
5.1 Continuous-time quantum walks are well modelled by thermal states

Here we demonstrate, for a few examples, that CTQWs are well modelled by thermal states, taking $\langle \bar{H}_p \rangle$ as a measure of success of the CTQW. If we know the temperature associated with the CTQW, this provides us with a second route to find the same value of $\langle \bar{H}_p \rangle$ by preparing a thermal state, thus side-stepping the need for complete unitary dynamics.

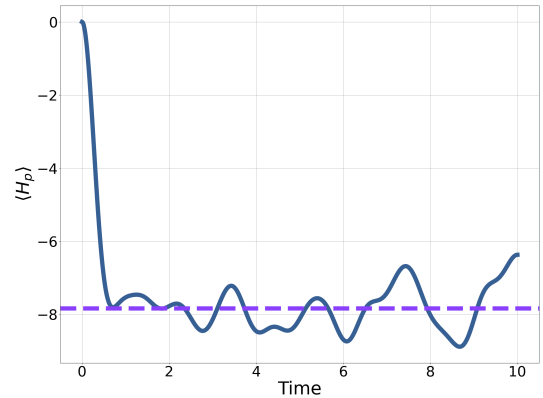
Since the Hamiltonian is constant during the evolution under a CTQW, energy is conserved in a closed-system. Hence the following must hold true for the thermal state [18]:

$$\text{Tr} [H_{QW} \rho_\beta (H_{QW})] = -n, \quad (19)$$

where $\rho_\beta(H_{QW})$ denotes a thermal state with Hamiltonian H_{QW} and inverse temperature β . This equation fixes β , therefore there are no free parameters to fit.



(a) A 12 qubit binomial graph.

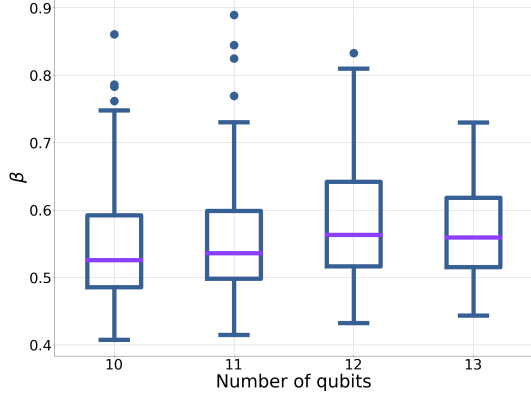


(b) A 12 qubit three-regular graph.

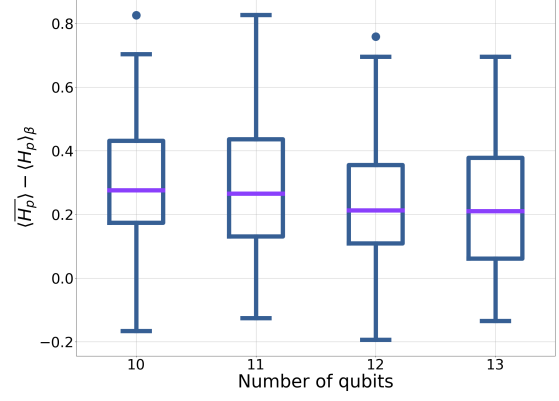
Figure 11: The performance of a CTQW with γ optimised to give the best value of $\langle \bar{H}_p \rangle$. The dashed purple line shows the thermal state prediction. The temperature is fixed using Eq. 19.

Fig. 11 shows $\langle H_p \rangle$ the Schrödinger evolution for a binomial graph and a three-regular graph and the corresponding prediction for $\langle H_p \rangle$ from the thermal state. For both problem instances it appears that $\langle H_p \rangle$ is fluctuating around a steady state value. Importantly, despite being far from the thermodynamic limit at only 12-qubits, the thermal state prediction is capturing the steady state behaviour well.

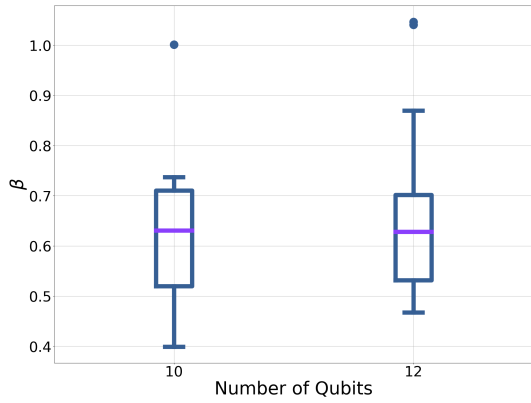
Fig. 12 shows the solution to Eq. 19 for the optimal choice of γ for multiple instances of binomial and three-regular graphs. For both cases the inverse temperatures are quite small, with $\beta < 1$ for almost all instances. This means CTQWs correspond to Gibbs states with high temperatures.



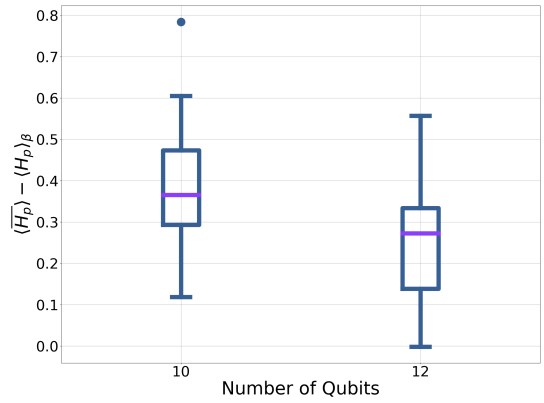
(a) Binomial graphs.



(a) Binomial graphs.



(b) Three-regular graphs.



(b) Three-regular graphs.

Figure 12: The inverse temperature associated with CTQWs (i.e. the solution to Eq. 19). For each problem instance γ is equal to γ_{opt} , shown in Fig. 5 and Fig. 7.

Figure 13: The corresponding thermal prediction for the values of β shown in Fig. 12

As mentioned in Sec. 2, we expect the ETH to hold true in the thermodynamic limit. Here we are working with of order ten qubits, hence we expect there to be finite-size effects meaning there will be some error between the Schrödinger equation and the thermal prediction. This is captured in Fig. 13. Generally, the thermal prediction, $\langle H_p \rangle_\beta$ is overestimating the performance of the CTQW. As n is increased we would expect this error to decrease. However, even for these very small systems the error is relatively small, with no significant outliers in this data set.

5.2 Modelling the density of states

In the previous section we numerically demonstrated that even for relatively small systems, $\langle \bar{H}_p \rangle$ is well predicted by a thermal state. Numerically solving Eq. 19 is however difficult, requiring finding the exponential of a matrix that increases exponentially with the problem size. In Sec. 5.3 we will attempt to estimate the temperature associated with a CTQW. Before this is possible we need to model the density of states (DOS) for H_{QW} .

To make progress we assume that the DOS is well modelled by a continuous distribution. We justify this assumption by arguing that the eigenstates of H_{QW} are likely to become exponentially close as the system size grows. This follows from the largest eigenvalue of H_{QW} being polynomial in n , and there being exponentially many states. Assuming no (or little) degeneracy, the difference between energy levels must shrink exponentially with the problem size. We also assume that the DOS is well modelled by a uni-modal distribution. This assumption will break down if γ is too large or too small. However, as seen in Sec. 3.2 we expect useful values of γ to correspond to somewhere between these two limits.

Intuitively, for large systems, we would expect the energy of a randomly chosen eigenstate to be close to the average energy. The average energy, μ , of H_{QW} is given by:

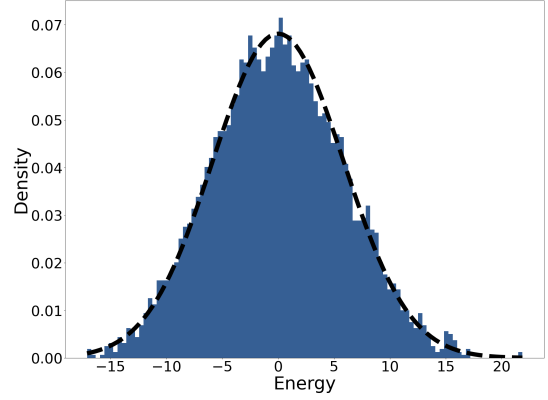
$$\begin{aligned} \mu &= \frac{1}{2^n} \sum_k E_k = \frac{1}{2^n} \text{Tr } H_{QW} \\ &= 0. \end{aligned}$$

Now consider the variance, σ^2 :

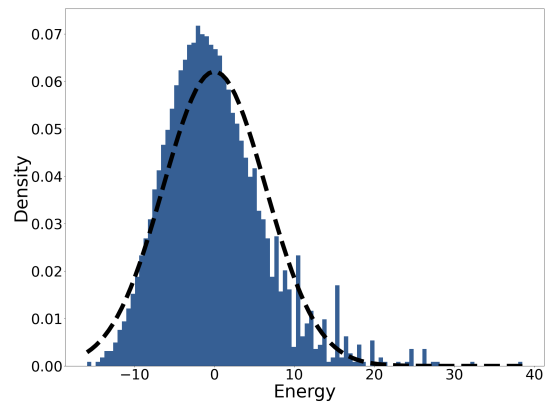
$$\begin{aligned} \sigma^2 &= \frac{1}{2^n} \sum_k E_k^2 = \frac{1}{2^n} \text{Tr } H_{QW}^2 \\ &= n + \gamma^2 \kappa_2. \end{aligned}$$

Therefore, as n goes to infinity, $\sigma/2^n$ tends to zero. So we expect a tightly peaked distribution for large n . The details for calculating moments of the DOS can be found in Appendix C.

The DOS for thermalising systems has previously been modelled with a Gaussian distribution [52]. Since the CTQW exhibits thermalisation we adopt this approach. Fig. 14 shows a Gaussian fit to the DOS for a 12-qubit three-regular graph (Fig. 14a) and a binomial graph (Fig. 14b) with γ optimised for each problem. Visually this appears



(a) A 12-qubit three-regular graph with $\gamma = 1.11$.



(b) A 12-qubit binomial graph with $\gamma = 0.78$.

Figure 14: Histogram for the DOS for two graphs with optimised γ . The energies have been binned into 100 bins. The DOS has been normalised such that the total density is equal to one. The dashed black line shows the fitted Gaussian distribution.

to be an acceptable approximation for the regular graph, less so for the binomial graph. Given the simplicity of this approximation we utilise it throughout the paper to make simple analytic predictions.

To verify the Gaussianity of the DOS, we look at two typical tests of normality: the skewness, s , of the distribution and the excess kurtosis k . The skewness is the ratio of the third moment of the distribution to the cube of the standard deviation. The kurtosis is the ratio of the fourth moment of the distribution to the variance squared. The excess kurtosis is the kurtosis minus the expected kurtosis of a Gaussian distribution, which is equal to 3.

Starting with the skewness:

$$s = \frac{6\gamma^3\kappa_3}{(n + \gamma^2\kappa_2)^{3/2}}. \quad (20)$$

Further details can be found in Appendix C. Note that triangle-free graphs have no skewness and that for positive γ , the skewness is always non-negative.

For a d -regular graph, $\kappa_2 = dn/2$ and κ_3 is bounded by $\mathcal{O}(d^2n)$. So the skewness, following the heuristic in Eq. 14 (provided γ is held constant) scales approximately as $n^{-1/2}$. Therefore it will tend to zero as the problem size is increased.

Conversely, for binomial graphs, provided γ scales proportional to $1/\sqrt{n}$, then the skewness will not scale with n . Hence, the Gaussian approximation will not hold as well for a binomial graph.

Examining now the excess kurtosis gives:

$$k = -2 \frac{n + 4\gamma^2\kappa_2 + \gamma^4(\kappa_2 - 12\kappa_4)}{(n + \gamma^2\kappa_2)^2} \quad (21)$$

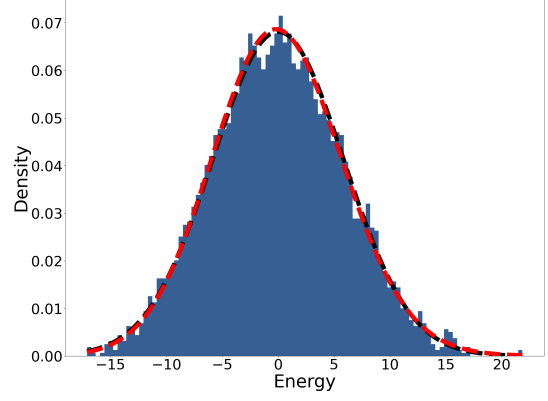
Under the same assumptions above for the scaling of the skewness, the excess kurtosis will tend to zero as n tends to infinity for regular graphs. For binomial graphs the excess kurtosis is unlikely to vanish.

For large regular graphs the Gaussian approximation looks to hold well. The same cannot be said to be true for binomial graphs. This is a consequence of regular graphs looking locally tree-like in the infinite sized limit [53]. Hence it may be necessary to consider other models for the DOS for binomial graphs and for regular graphs away from the infinite-size limit.

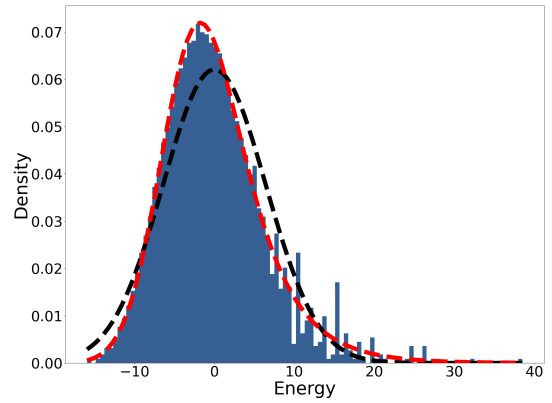
Here we propose using an exponentially modified Gaussian (EMG) distribution [54, 55] to model the DOS for a CTQW. The EMG is a skewed unimodal distribution. The skew can only be positive, as is the case with the DOS for a CTQW. In the correct limit it can recover a Gaussian. It has the following probability density function:

$$p(x; m, \nu, \lambda) dx = \frac{\lambda}{2} e^{\frac{\lambda}{2}(2m + \lambda\nu^2 - 2x)} \times \operatorname{erfc} \left(\frac{m + \lambda\nu^2 - x}{\sqrt{2\nu^2}} \right) dx, \quad (22)$$

where erfc is the complementary error function. The fitting parameters m , ν and λ can be ascer-



(a) A 12 qubit three-regular graph with $\gamma = 1.11$.

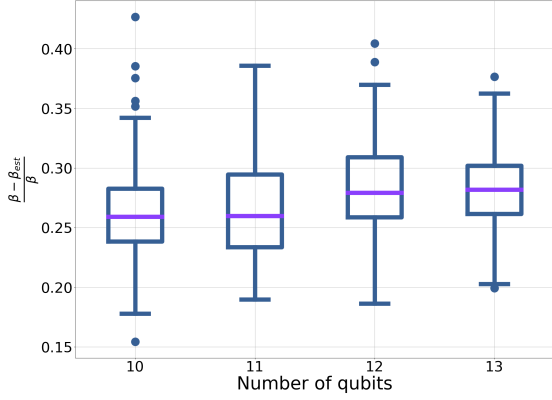


(b) A 12 qubit binomial graph with $\gamma = 0.78$.

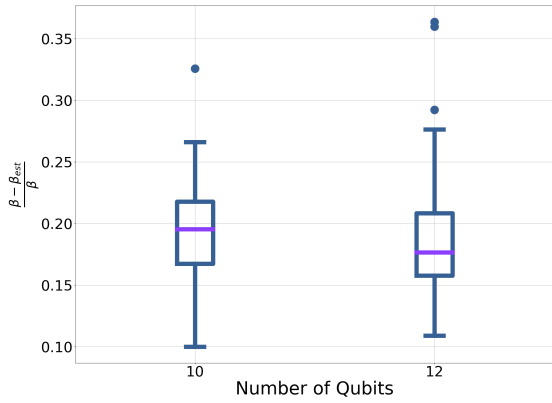
Figure 15: Histogram for the DOS for two graphs with γ optimised. The energies have been binned into 100 bins. The DOS has been normalised such that the total density is equal to one. The dashed black line shows the fitted Gaussian distribution. The dashed red line shows the EMG distribution.

tained from n , σ and s . For the details the reader is referred to Appendix D.2.

The Gaussian approximation only includes information about the problem-size and therefore will make the same prediction for a large number of graphs. For instance, all regular graphs with the same degree have the same DOS under the Gaussian approximation. By incorporating the skewness into the model we are incorporating more information about the graph structure, namely the number of triangles in the problem. Visually, as shown in Fig. 15, we can see that the EMG distribution (the dashed red line) provides a better fit than the Gaussian distribution for the binomial graph. The EMG distribution also still models the DOS for the regular graph well.



(a) Binomial graphs



(b) Three-regular graphs

Figure 16: The error in the predicted inverse temperature assuming a Gaussian DOS (i.e. Eq. 24). For each problem instance γ is equal to γ_{opt} , shown in Fig. 5 and Fig. 7.

During the rest of this paper we use the DOS to make analytic predictions about the behaviour of a CTQW in a closed-system setting.

5.3 Predictions from the density of states

Having provided a model for the DOS for H_{QW} , in this section we make two predictions in relation to CTQWs, namely the temperature and the entropy of entanglement.

5.3.1 Estimating the temperature

Finding the temperature of the CTQW requires finding the solution to Eq. 19. To find approximate solutions we use the DOS models to evaluate the partition function, \mathcal{Z} . Eq. 19 then be-

comes [56]:

$$-n = -\frac{\partial \ln \mathcal{Z}}{\partial \beta}. \quad (23)$$

Assuming a Gaussian DOS then the estimated inverse temperature is given by:

$$\beta_{est} = \frac{n}{n + \gamma^2 \kappa_2}. \quad (24)$$

This is perhaps what one would estimate as the temperature on dimensional grounds alone. The full details can be found in Appendix D.1.

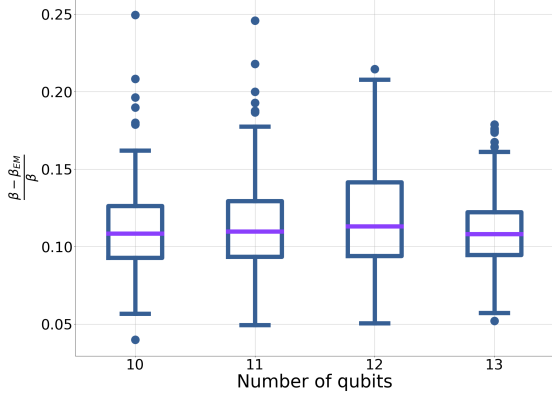
Taking the EMG to model the density of states gives:

$$\beta_{EM} = \frac{n\Delta - \sigma^2 + \sqrt{(n\Delta + \sigma^2)^2 - 4n\Delta^3}}{2\Delta(\sigma^2 - \Delta^2)}, \quad (25)$$

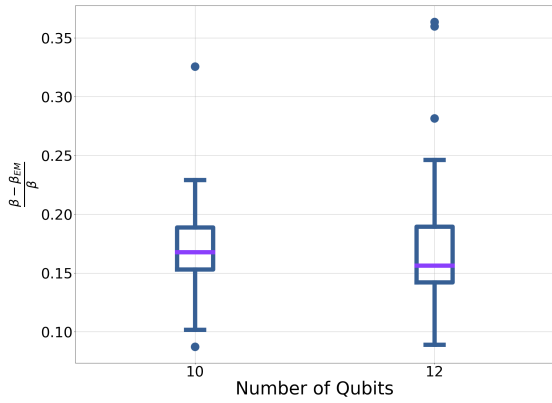
where $\sigma^2 = n + \gamma^2 \kappa_2$ and $\Delta = \gamma(3\kappa_3)^{1/3}$. The details can be found in Appendix D.2.

Fig. 16 shows the error in temperature between β_{est} and β (the numerical solution to Eq. 19). The approximate solutions are consistently underestimating the inverse temperature. The error for binomial graphs is substantial, with typical errors being between 25% and 30% (Fig. 16a). For three-regular graphs, this reduces to somewhere between 15% and 25% (Fig. 16b). Using the EMG distribution (i.e. Eq. 25) provides little improvement on the estimate for the three-regular graphs (Fig. 17b). This is unsurprising given how close to Gaussian the DOS is for these problems. For binomial graphs the improvement is more substantial with typical errors being between 10% and 15%. As n is increased, we expect that the DOS will be better modelled by a continuous DOS, hence we would expect the error in β to be improved.

By assigning the quantum evolution a temperature, we have mapped the challenge of understanding a dynamical problem to a static problem. So far in this section we have shown how to reasonably estimate the associated temperature. If the temperature is too high, then the associated thermal state can be efficiently classically approximated. Results by Crosson et al. [57] suggest that for values of $\beta \leq 0.1$ a classical computer could simulate the associated thermal state efficiently for a three-regular graph, suggesting that for a CTQW to provide an advantage, it must operate outside this regime.



(a) Binomial graphs



(b) Regular graphs

Figure 17: The error in the predicted inverse temperature assuming an EMG DOS (i.e. Eq. 25). For each problem instance γ is equal to γ_{opt} , shown in Fig. 5 and Fig. 7.

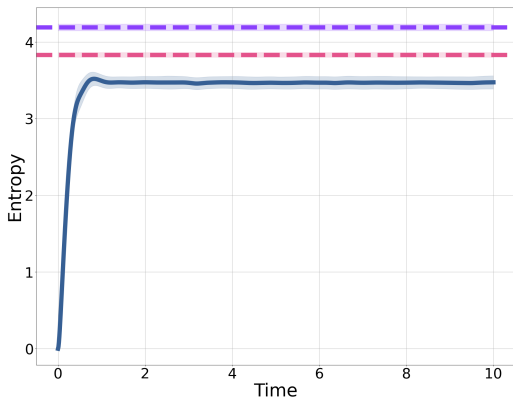


Figure 18: The entanglement entropy averaged over a hundred 14-qubit binomial instances (solid blue line). The dashed purple line shows S^{norm} and the dashed pink line S^{EM} . The shaded regions show a single standard deviation.

5.3.2 Estimating the entropy

Often considered of high interest for a quantum algorithm is the entanglement in the system [58]. This can be quantified by the entanglement entropy. Here we calculate the entanglement entropy by tracing out half of the qubits, which are randomly selected, and calculating the von Neumann entropy of the reduced density operator.

We compare the entanglement entropy to the thermodynamic entropy derived from the modelled DOS, given by:

$$S = \beta \langle H_{QW} \rangle + \ln \mathcal{Z}. \quad (26)$$

Since entropy is an extensive quantity we take half of the above and compare it to the entanglement entropy. Although the ETH cannot be straightforwardly applied here as the entropy is not a local observable, the two have been shown to be linked in previous works [59–61].

For the Gaussian distribution Eq. 26 reduces to:

$$S^{norm} = n \ln 2 - \beta_{est} n + \frac{\sigma^2 \beta_{est}^2}{2}, \quad (27)$$

and for the EMG:

$$S^{EM} = n \ln 2 + \beta_{EM} (\Delta - n) + \frac{\beta_{EM}^2}{2} (\sigma^2 - \Delta^2) - \ln(1 + \beta_{EM} \Delta) \quad (28)$$

Fig. 18 shows the entanglement entropy averaged over a hundred 14-qubit binomial graphs as a function of time. As is clear from the figure the entanglement entropy is approximately constant. The dashed purple line shows S^{norm} and the dashed pink line S^{EM} . Both are overestimating the entanglement entropy. However, they appear to be reasonable estimates with S^{EM} providing the better estimate. As mentioned in Sec. 3, the ideal final state will be a low-entanglement state. CTQWs lack any mechanism to dissipate entanglement, as shown by Fig. 18

5.4 Estimating $\langle \bar{H}_p \rangle$

In Sec. 5.3.1 we saw how the temperatures associated with a CTQW could be found. By assuming the system is well modelled by a Gibbs distribution, it follows that

$$\langle \bar{H}_p \rangle = -\frac{1}{\beta} \frac{\partial \ln \mathcal{Z}}{\partial \gamma}, \quad (29)$$

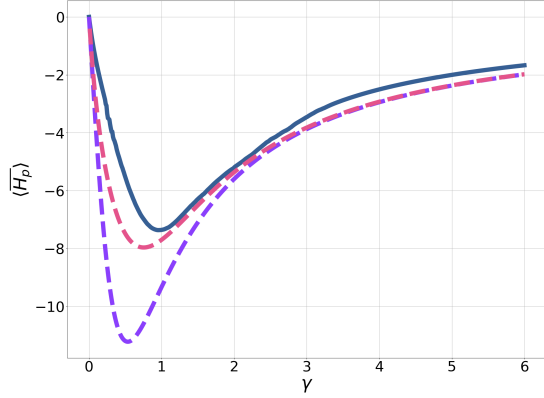


Figure 19: The solid blue line shows $\langle \bar{H}_p \rangle$ for a 12-qubit binomial graph. The dashed purple (pink) line shows the prediction from Eq. 29 assuming a Gaussian (EMG) DOS.

holds approximately true. The derivation behind this equation can be found in Appendix E. For the Gaussian approximation this gives:

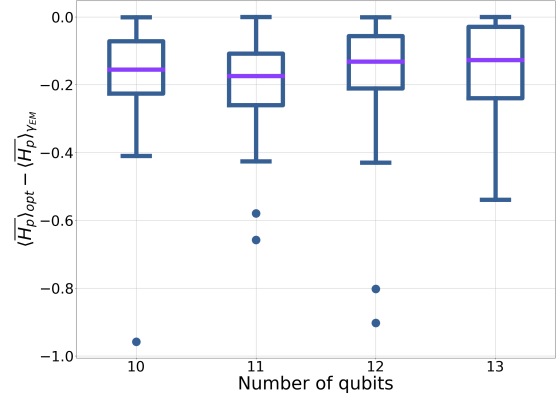
$$\langle \bar{H}_p \rangle_{est} = -\frac{\gamma n \kappa_2}{n + \gamma^2 \kappa_2}, \quad (30)$$

which is optimised by

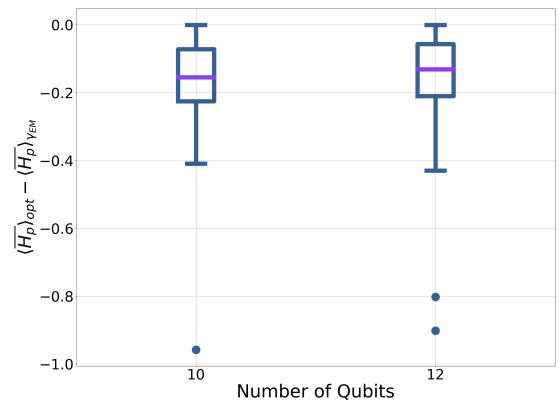
$$\gamma_{est}^{opt} = \sqrt{\frac{n}{\kappa_2}}. \quad (31)$$

The details can be found in Appendix D.1. Note that this is the same result as the heuristic of maximising the dynamics (i.e. Eq. 14). The same approach can be taken for the EMG model. The corresponding result is too cumbersome to be written out in full.

Fig. 19 compares $\langle \bar{H}_p \rangle$ to the predictions from the Gaussian and EMG predictions for a single 12-qubit binomial graph. Though both overestimate the performance of the CTQW, both provide reasonable estimates on the optimal γ . The EMG clearly provides a better model for $\langle \bar{H}_p \rangle$ than the Gaussian model. Indeed it is remarkably close to the direct numerical calculation of $\langle \bar{H}_p \rangle$. Note that although the thermal model appears to provide good estimates for $\langle \bar{H}_p \rangle$ for large and small γ , in this regime $\langle H_p(t) \rangle$ will not necessarily display steady-state behaviour. For example in Fig. 9, γ is very small and there are significant oscillations present. However, this regime of γ , as discussed in Sec. 3, is unlikely to be of practical interest to CTQWs.



(a) Binomial graphs



(b) Three-regular graphs

Figure 20: The difference in $\langle \bar{H}_p \rangle$ between the optimal choice of γ and the prediction from the EMG DOS (γ_{EM}).

For the performance of γ_{est}^{opt} the reader is referred to Sec. 3. For the EMG distribution there are two important quantities to examine:

- How much does the performance of $\langle \bar{H}_p \rangle$ change between the optimal γ and the γ predicted by the EMG DOS?
- What is the difference between the prediction of $\langle \bar{H}_p \rangle$ and the true value?

The first of these is addressed for small problem sizes in Fig. 20. The performance is improved over γ_{heur} from simply balancing the drive and problem Hamiltonians, especially for binomial graphs. Hence, we believe we have found a good heuristic method for finding γ for CTQWs applied to MAX-CUT.

Fig. 21 shows the error in the performance between the true value of $\langle \bar{H}_p \rangle$ and the predicted

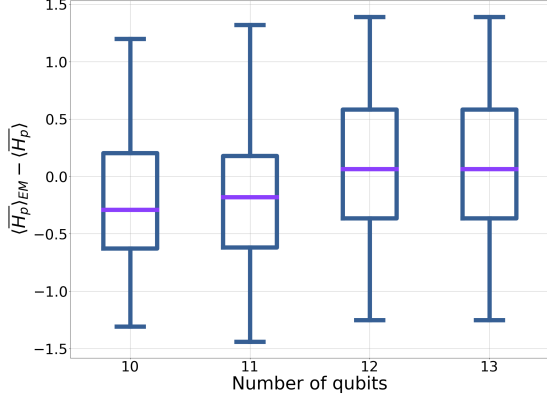


Figure 21: The difference between the prediction of $\langle H_p \rangle$ with an EMG DOS using Eq. 29 ($\langle H_p \rangle_{EM}$) and the true value. For each instance the value of γ has been chosen such that $\langle H_p \rangle_{EM}$ is minimised. Each bar shows one hundred binomial graph instances.

value from the EMG approximation, for the γ predicted to give the best possible performance from the EMG DOS. The error is relatively small for these small problem sizes. Given the tractability of Eq. 29, and evidence of small errors, it is possible to estimate the performance for larger problem sizes. This is shown in Fig. 22.

By treating closed-system CTQWs as thermalising systems we have shown they thermalise. From here we have been able to provide useful heuristic choices for optimising γ and to make practical predictions for $\langle \bar{H}_p \rangle$. For further discussion of the performance of CTQWs see Appendix F. We have also provided an alternative approach to understanding CTQW that might be applicable to other combinatorial optimisation problems.

6 Multi-stage quantum walks

6.1 The set-up

So far we have focused on time-independent Hamiltonians. In this section we introduce time-dependence into the CTQW, by implementing multi-stage quantum walks (MSQWs). In this case γ is increased monotonically in a step-wise fashion. With each step of γ the system is allowed to reach equilibrium. For instance an l -stage

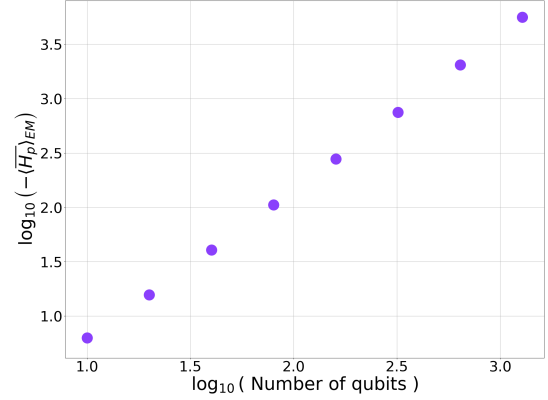


Figure 22: Estimated scaling of the CTQW for binomial graphs assuming thermalisation and an EMG DOS. Each data point shows the median performance of one hundred binomial graph instances.

quantum walk would have the following schedule:

$$\gamma(t) = \begin{cases} \gamma_1 & \text{for } 0 \leq t < t_1 \\ \gamma_2 & \text{for } t_1 \leq t < t_2 \\ \dots & \\ \gamma_k & \text{for } t_{k-1} \leq t < t_k \\ \dots & \\ \gamma_l & \text{for } t_{l-1} \leq t \end{cases} \quad (32)$$

with $\gamma_l > \dots > \gamma_k > \dots > \gamma_2 > \gamma_1$. This schedule can be seen as being motivated by approaches like Quantum Annealing [47]. For each step, while γ is held constant energy is conserved. Therefore we can modify Eq. 19 to be:

$$\text{Tr} \left[H_{QW}^{(k)} \rho_\beta \left(H_{QW}^{(k)} \right) \right] = \langle \psi(t_{k-1}) | H_{QW}^{(k)} | \psi(t_{k-1}) \rangle, \quad (33)$$

where $H_{QW}^{(k)}$ is the Hamiltonian during the k^{th} stage of the quantum walk (i.e. with γ_k) and $|\psi(t)\rangle$ is the state-vector associated with the MSQW. Fig. 23 shows a five-stage CTQW for a 12-qubit binomial graph. Each stage increases γ by 0.5, starting with 0.5. Each stage has a duration of twenty units of time. With each stage $\langle H_p \rangle$ (the solid blue line) improves and quickly tends to an approximate steady-state. Solving Eq. 33 to find the temperature and associated performance gives the purple line in the figure. For this instance, we have good qualitative and reasonable quantitative agreement between the thermal prediction and the Schrödinger equation. During the first stage of the MSQW, the thermal prediction is quite far from the prediction of the

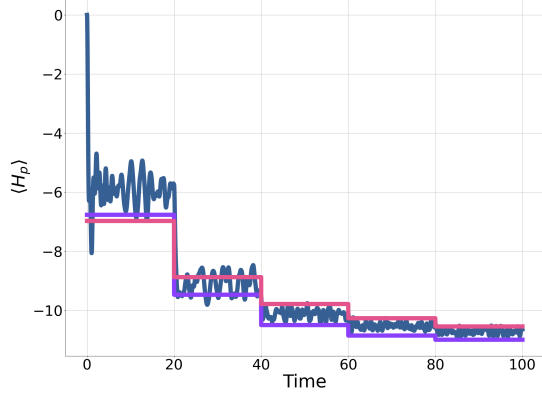


Figure 23: A five-stage CTQW on a 12-qubit binomial graph. The purple line is the numerical prediction assuming that the MSQW is well modelled by a CTQW. The pink line shows the prediction from Eq. 29 and modelling the DOS as an EMG distribution.

Schrödinger equation, suggesting the state is far from thermal. This is to be expected as γ is small and the driver Hamiltonian dominates, breaking the assumption that the DOS is continuous. Despite this, as γ is increased the system thermalises. Perhaps most interestingly, as γ is further increased the state remains thermal despite the problem Hamiltonian becoming more dominant. The inverse temperatures associated with this MSQW are $\beta = 1.21, 0.72, 0.53, 0.41, 0.32$. This corresponds to heating the system, suggesting that despite the final stage providing the best performance, it might be the distribution easiest to simulate classically [57]. Given the thermal behaviour, it might be reasonable to assume that the techniques developed earlier in this paper for CTQW can be applied to MSQW.

6.2 Making predictions

In Sec. 5 the tools for making predictions about CTQWs were developed. Here we apply these tools to the MSQW. Here we only use the EMG distribution to model the DOS, since this also captures the Gaussian model.

To make predictions Eq. 33 needs to be approximated to find β . Since finding $|\psi(t)\rangle$ is likely to be numerically intractable for large systems we

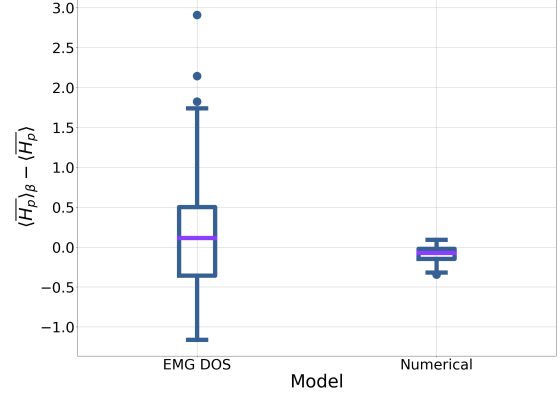


Figure 24: The difference between the stationary state calculated from the Schrödinger equation and a thermal state for five-stage CTQW on 100 12-qubit binomial graphs. “EMG DOS” refers to the prediction with the energy fixed by Eq. 34. The numerical approach gives the prediction from using the numerically determined energy, numerically fixing the temperature and calculating $\langle H_p \rangle$.

approximate Eq. 33 by:

$$\begin{aligned}
 & \langle \psi(t_{k-1}) | H_{QW}^{(k)} | \psi(t_{k-1}) \rangle \\
 &= \langle \psi(t_{k-1}) | H_d + \gamma_k H_p | \psi(t_{k-1}) \rangle \\
 &= \langle \psi(t_{k-1}) | H_d + \gamma_{k-1} H_p | \psi(t_{k-1}) \rangle \\
 & \quad + (\gamma_k - \gamma_{k-1}) \langle \psi(t_{k-1}) | H_p | \psi(t_{k-1}) \rangle \\
 & \approx \langle H_{QW}^{(k-1)} \rangle_{(k-1)} + (\gamma_k - \gamma_{k-1}) \langle \bar{H}_p \rangle_{(k-1)}, \quad (34)
 \end{aligned}$$

where $\langle \cdot \rangle_{k-1}$ denotes the expectation during the $(k-1)^{\text{th}}$ stage. Therefore, we can estimate Eq. 33 by recursively evaluating Eq. 34 using Eq. 29 to calculate $\langle \bar{H}_p \rangle_{(k-1)}$. In Fig. 23 the solid pink line shows the prediction, using Eq. 34, to good agreement.

In Fig. 24 we consider the final value of $\langle \bar{H}_p \rangle$ for a hundred five-stage CTQW on 12-qubit binomial graphs compared to the analytical prediction (i.e. Eq. 34) and the numerical prediction assuming thermalisation. The schedule is the same as the schedule used in Fig. 23. Consider first the numerical prediction. Assuming thermalisation, the correct final steady state for all the problem instances is predicted within an error of 0.5. In contrast the EMG DOS model only achieves this error bound for approximately 50% of the instances. This approach suffers from cumulative errors in $\langle H_p \rangle$ at each stage of the MSQW (e.g. Eq. 34), which can result in large errors in the prediction.

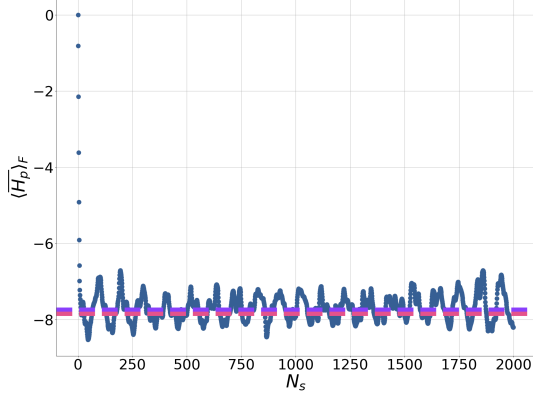


Figure 25: A Floquet system simulating a CTQW for a 12-qubit binomial graph, with $\gamma = 1$ and $\tau = 0.1$. The dashed purple line shows the time averaged value (i.e. Eq. 36). The dashed pink line shows the prediction from Eq. 29, assuming an EMG DOS.

Here we have numeric evidence for MSQW exhibiting thermalisation. The EMG DOS model struggles to capture the performance for all instances, as well as the time-independent case. This is due to the difficulty in determining the energy of the system. To improve this approach a better model for the DOS or prediction of $\langle H_p \rangle$ needs to be developed, especially for the small γ case.

7 Implications for gate-based implementation

One approach for realising a CTQW might be to simulate it on a gate-based quantum computer. The simplest approach to achieve this is to Trotterise the evolution [62], such that the Hamiltonian alternates between H_d and H_p . The resulting state-vector is prepared by the quantum computer:

$$|\psi_F(t)\rangle = \left(e^{-i\tau H_d/2} e^{-i\gamma\tau H_p/2} \right)^{N_s} |+\rangle, \quad (35)$$

where $t = N_s\tau$ and N_s is an integer. Here N_s acts as a proxy for circuit depth. Setting τ presents a trade-off between minimising gate-depth with a large τ , and accuracy of the simulation that requires τ to be small. Alternatively, Eq. 35 could be viewed as a QAOA [63] schedule with each step having the same set of angles. In the previous sections the Hamiltonians have been held constant for extended intervals in time. This has provided

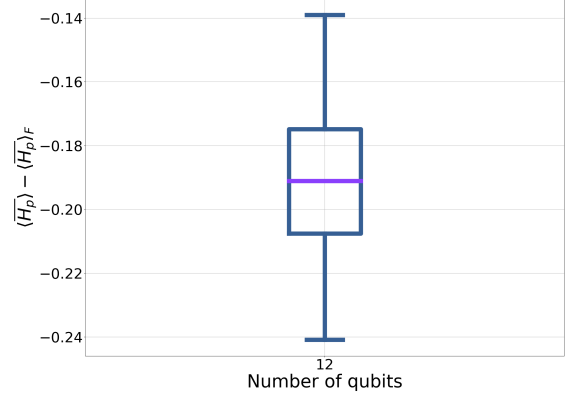


Figure 26: The reduced performance resulting from implementing the CTQW as a Floquet system for one-hundred 12-qubit binomial graphs. For the Floquet systems $\tau = 0.2$. For all the problem instances $\gamma = 1$.

a local conserved quantity, resulting in a finite-temperature thermal state. The Trotterised system is not time-independent and energy is not conserved. In this section we investigate Eq. 35. Since the Hamiltonian is periodic in time, it is an example of a Floquet system (a very brief introduction can be found in Appendix G).

To analyse the long time behaviour of the Floquet system we evaluate:

$$\langle \bar{H}_p \rangle_F = \sum_{\epsilon_\alpha = \epsilon_\beta} c_\alpha c_\beta^* \langle \phi(\tau)_\beta | H_p | \phi(\tau)_\alpha \rangle, \quad (36)$$

where $|\phi(t)_\alpha\rangle$ are the Floquet modes, ϵ_α the corresponding quasi-energies, and $c_\alpha = \langle \phi(0)_\alpha | + \rangle$ [64]. This is illustrated in Fig. 25 for a 12 qubit binomial graph with $\tau = 0.1$ and $\gamma = 1$. The solid blue line shows the Schrödinger equation. Within a few intervals of τ the system has already thermalised to $\langle \bar{H}_p \rangle_F = -7.74$. This is shown by the dashed purple line (i.e. the time averaged value from Eq. 36). The performance of the Floquet system is marginally worse than the value of $\langle \bar{H}_p \rangle = -7.79$ for the corresponding CTQW. Indeed Fig. 26 shows the difference between $\langle \bar{H}_p \rangle_F$ and the corresponding value for a CTQW for one-hundred binomial qubit instances. For the Floquet instances $\tau = 0.2$. For each problem instance the CTQW provided the better performance.

Depending on the size of τ , we might expect two things to occur for $\langle \bar{H}_p \rangle_F$ [65]. For small τ the system does not have time to thermalise, so $|\psi_F(t)\rangle$ mimics the behaviour of a CTQW.

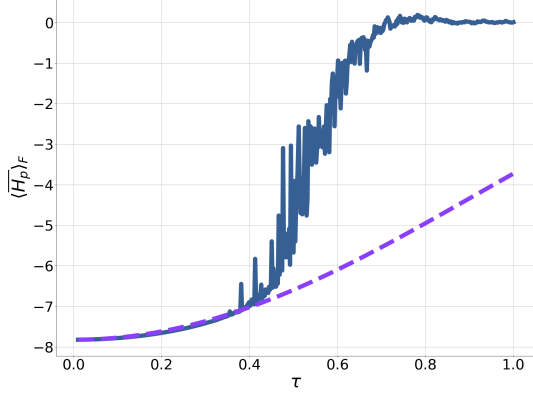


Figure 27: The steady state value for a single 12-qubit binomial graph with $\gamma = 1$ as τ is increased. The blue line shows the prediction from Eq. 36. The dashed purple line shows the prediction using the time-independent Hamiltonian shown in Eq. 40.

For large τ , the system thermalises at each step. Since energy is not conserved, the system tends to the infinite temperature Gibbs state, resulting in $\langle \tilde{H}_p \rangle_F = 0$. This is shown by the blue line in Fig. 27 for a single 12-qubit binomial graph as τ is increased.

7.1 The thermal model

In this section we provide an approximate time-independent model to capture the stroboscopic behaviour of the Floquet system assuming the Trotter-step is small. Previous works have focused on using a Magnus expansion [66–68], for ease of computation here we take an alternative approach.

The stroboscopic behaviour is captured by the single period unitary:

$$U_F(\tau) = e^{-i\tau H_d/2} e^{-i\gamma\tau H_p/2}, \quad (37)$$

from here it is a trivial exercise to recover the associated Hamiltonian:

$$H_F(t) = iU_F^\dagger(t) \frac{d}{dt} U_F(t) \quad (38)$$

$$= \frac{1}{2} \left[U_p^\dagger(t) H_d U_p(t) + \gamma H_p \right] \quad (39)$$

where $U_p(t) = e^{-iH_p\gamma t/2}$. Assuming τ is small, making the approximation $H_F(t) \approx H_F(\tau/2)$ gives a time-independent Hamiltonian. We can also neglect the constant scaling of the Hamiltonian that is not going to change the long time

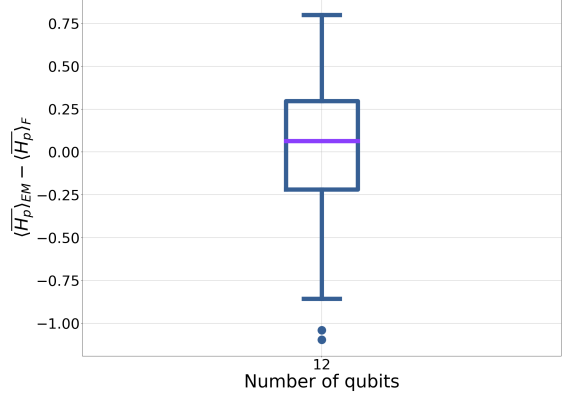


Figure 28: The difference between the true steady state and the steady state predicted assuming the system is well modelled by a time-independent Hamiltonian with a EMG distribution. The figure shows the results for one hundred 12-qubit binomial graph instances with $\gamma = 1$ and $\tau = 0.2$.

behaviour. Hence we are left with the following time-independent Hamiltonian:

$$\tilde{H}_F = U_p^\dagger(\tau/2) H_d U_p(\tau/2) + \gamma H_p \quad (40)$$

From here we can follow the steps outlined in Sec. 5 for the CTQW. Note that

$$\begin{aligned} \text{Tr } \tilde{H}_F^k(t) &= \text{Tr} \left[U_p^\dagger(\tau/2) H_d U_p(\tau/2) + \gamma H_p \right]^k \\ &= \text{Tr} [H_d + \gamma H_p]^k \\ &= \text{Tr } H_{QW}^k, \end{aligned} \quad (41)$$

so we can use the moments for the DOS derived for the CTQW. Therefore any change in performance must come from the energy of the initial state:

$$\langle + | \tilde{H}_F | + \rangle = - \sum_{k=1}^n \cos^{\text{deg}(k)}(\gamma\tau/4), \quad (42)$$

where $\text{deg}(k)$ is the degree of the k^{th} node. Notably, for binomial graphs (or any graph where $\text{deg}(k)$ increases with n) as n goes to infinity, $\langle + | \tilde{H}_F | + \rangle$ goes to zero. This corresponds to an infinite temperature thermal state. Consequently, in the thermodynamic limit, under the above assumptions, any attempt to approximate a CTQW with a Floquet system will fail. In contrast, regular graphs will have more stability. The numerical prediction for thermalising under \tilde{H}_F can be seen in Fig. 27 (dashed purple line) for

a single 12-qubit binomial graph. The numerical prediction agrees well with the true numerical value of $\langle \bar{H}_p \rangle_F$ (solid blue line) up until the crossover to an infinite temperature Gibbs state.

As established in the earlier section on CTQWs, they correspond to Gibbs states. Therefore, we use the same model as for the CTQW with a modified energy constraint (i.e. Eq. 42). This is shown by the dashed pink line in Fig. 25, for a single 12-qubit binomial graph to good agreement. The results for one hundred 12-qubit binomial graphs with $\tau = 0.2$ and $\gamma = 1$ can be seen in Fig. 28. As shown in the figure the median error between the numerically calculated value of $\langle \bar{H}_p \rangle_F$ and the analytic value, assuming an EMG DOS and thermalisation, is close to zero.

Perhaps most usefully, this thermal model provides some intuition on how mitigate Trotter errors for this system. Essentially, we wish to start in the effective ground-state of the Floquet system. Eq. 40, suggests this might be achieved by changing the initial state from $|+\rangle$ to $U_p^\dagger(\tau/2)|+\rangle$. This is illustrated for a single 12-qubit binomial graph in Fig. 29 with $\gamma = 1$. The dashed purple line shows the steady-state value for a CTQW. The solid blue shows the steady state of the Floquet system with the initial state being the $|+\rangle$ state. The pink line shows the same Floquet system with the modified initial state $U_p^\dagger(\tau/2)|+\rangle$. Indeed, for the short time steps, the modified initial state better matches the CTQW. Indeed it even provides a slight advantage for a small interval in τ .

8 Conclusion

Throughout this paper we have attempted to understand the performance of CTQWs outside of what is classically simulable for non-integrable models. For the short-time limit, we were able to exploit the effective low dimensionality to demonstrate that the performance can be characterised by underlying graph properties, such as the number of edges in the graph.

For the steady-state we conjectured that the system thermalises. This means despite the state-vector consisting of 2^n complex amplitudes, for a given γ , the value of $\langle H_p \rangle$ will depend on one real number, the energy. This provides some insight into the computational mechanisms involved in CTQWs.

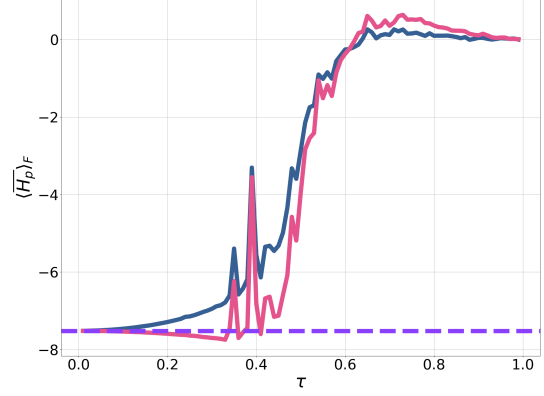


Figure 29: Improved performance from changing the initial state. The dashed purple line shows the steady-state value for a CTQW. The solid blue line shows the steady state of the Floquet system with initial state $|+\rangle$. The solid pink line shows the steady state of the Floquet system with initial state $U_p^\dagger(\tau/2)|+\rangle$.

Assuming thermalisation, classical statistical physics provides an alternative route to understanding CTQWs for MAX-CUT on a broad range of graphs within a closed-system setting. By associating the unitary dynamics with a temperature, it provides an alternative route to achieve the same value of $\langle \bar{H}_p \rangle$, either through dissipative dynamics or classical simulation. Here we have introduced an EMG distribution to account for some of the frustration in the model to make analytic predictions. By exploiting this model we were able to find reasonable estimates for the optimal choice of γ , that utilised properties of the underlying graph. Importantly, we were able to make predictions far away from what it is easy to directly and completely simulate. We have also demonstrated some preliminary results for MSQWs.

Further to this we were able to extend these results to Floquet systems as an approach to implementing CTQWs. For all the problems considered the Floquet system performed worse than the corresponding CTQW with the same γ . Our results predict that in the thermodynamic limit, MAX-CUT on binomial graphs will thermalise to an infinite temperature Gibbs state, independent of the step size. This is consistent with known work on thermalisation in Floquet systems [65]. The Floquet systems considered in this paper can be thought of as a QAOA [63] ansatz with restricted angle choices. For QAOA we wish to

avoid the system thermalising to an infinite temperature Gibbs state, hence it would be prudent to avoid a Floquet ansatz. Further to this it is reasonable to assume large step sizes might be detrimental for high-depth QAOA circuits. It has been observed that a QAOA circuit with randomly chosen step sizes also thermalises to an infinite temperature state [58], suggesting that thermalisation is not solely the consequence of the discrete time symmetry of the Floquet system. How far the results of this paper might extend to other NISQ algorithms such as QAOA and Quantum Annealing is an ongoing area of research.

Acknowledgements

We gratefully acknowledge Itay Hen, Michael Kolodrubetz, Matthew Cooper, Henry Chew, Natasha Feinstein, Leon Guerrero and Pias Tubrinuan for inspiring discussion and helpful comments. Robert Banks would like to acknowledge Orbyts initiative for providing the opportunity to collaborate with an outstanding team of students at Newham Collegiate Sixth Form Centre. Many thanks to Serena Maugeri, Abbie Bray, and William Dunn for making this possible and the Mathematical and Physical Sciences faculty at UCL through the “Take bold action for inclusion 2022-2023” grant for funding it. The authors acknowledge the use of the UCL Myriad High Performance Computing Facility (Myriad@UCL), and associated support services, in the completion of this work. This work was supported by the Engineering and Physical Sciences Research Council through the Centre for Doctoral Training in Delivering Quantum Technologies [grant number EP/S021582/1]. Dan Browne was supported by the Engineering and Physical Sciences Research Council [grant numbers EP/S005021/1 and EP/T001062/1]. Paul Warburton was supported by the Engineering and Physical Sciences Research Council [grant number EP/W00772X/2]. For the purpose of open access, the author has applied a Creative Commons Attribution (CC BY) licence to any Author Accepted Manuscript version arising.

References

- [1] Edward Farhi and Sam Gutmann. “Quantum computation and decision trees”. *Phys. Rev. A* **58**, 915–928 (1998).
- [2] Andrew M. Childs. “Universal computation by quantum walk”. *Phys. Rev. Lett.* **102**, 180501 (2009).
- [3] Kunkun Wang, Yuhao Shi, Lei Xiao, Jingbo Wang, Yogesh N. Joglekar, and Peng Xue. “Experimental realization of continuous-time quantum walks on directed graphs and their application in pagerank”. *Optica* **7**, 1524–1530 (2020).
- [4] Yunkai Wang, Shengjun Wu, and Wei Wang. “Controlled quantum search on structured databases”. *Phys. Rev. Res.* **1**, 033016 (2019).
- [5] Yang Wang, Shichuan Xue, Junjie Wu, and Ping Xu. “Continuous-time quantum walk based centrality testing on weighted graphs”. *Scientific Reports* **12**, 6001 (2022).
- [6] Andrew M. Childs, Richard Cleve, Enrico Deotto, Edward Farhi, Sam Gutmann, and Daniel A. Spielman. “Exponential algorithmic speedup by a quantum walk”. In *Proceedings of the thirty-fifth annual ACM symposium on Theory of computing*. ACM (2003).
- [7] Josh A. Izaac, Xiang Zhan, Zhihao Bian, Kunkun Wang, Jian Li, Jingbo B. Wang, and Peng Xue. “Centrality measure based on continuous-time quantum walks and experimental realization”. *Phys. Rev. A* **95**, 032318 (2017).
- [8] T. Loke, J. W. Tang, J. Rodriguez, M. Small, and J. B. Wang. “Comparing classical and quantum pageranks”. *Quantum Information Processing* **16**, 25 (2016).
- [9] Andrew M. Childs and Jeffrey Goldstone. “Spatial search by quantum walk”. *Phys. Rev. A* **70**, 022314 (2004).
- [10] Adam Callison, Nicholas Chancellor, Florian Mintert, and Viv Kendon. “Finding spin glass ground states using quantum walks”. *New Journal of Physics* **21**, 123022 (2019).
- [11] Puya Mirkarimi, Adam Callison, Lewis Light, Nicholas Chancellor, and Viv Kendon. “Comparing the hardness of max 2-sat problem instances for quantum and classical algorithms”. *Phys. Rev. Res.* **5**, 023151 (2023).

- [12] Adam Callison. “Continuous-time quantum computing”. *PhD thesis*. Imperial College London. (2021).
- [13] Adam Callison, Max Festenstein, Jie Chen, Laurentiu Nita, Viv Kendon, and Nicholas Chancellor. “Energetic perspective on rapid quenches in quantum annealing”. *PRX Quantum* **2**, 010338 (2021).
- [14] J. M. Deutsch. “Quantum statistical mechanics in a closed system”. *Phys. Rev. A* **43**, 2046–2049 (1991).
- [15] Mark Srednicki. “Chaos and quantum thermalization”. *Phys. Rev. E* **50**, 888–901 (1994).
- [16] Joshua M Deutsch. “Eigenstate thermalization hypothesis”. *Reports on Progress in Physics* **81**, 082001 (2018).
- [17] Marcos Rigol. “Breakdown of thermalization in finite one-dimensional systems”. *Phys. Rev. Lett.* **103**, 100403 (2009).
- [18] Fabian H L Essler and Maurizio Fagotti. “Quench dynamics and relaxation in isolated integrable quantum spin chains”. *Journal of Statistical Mechanics: Theory and Experiment* **2016**, 064002 (2016).
- [19] Marlon Brenes, Tyler LeBlond, John Goold, and Marcos Rigol. “Eigenstate thermalization in a locally perturbed integrable system”. *Phys. Rev. Lett.* **125**, 070605 (2020).
- [20] Jae Dong Noh. “Eigenstate thermalization hypothesis and eigenstate-to-eigenstate fluctuations”. *Phys. Rev. E* **103**, 012129 (2021).
- [21] David A. Huse, Rahul Nandkishore, Vadim Oganesyan, Arijeet Pal, and S. L. Sondhi. “Localization-protected quantum order”. *Phys. Rev. B* **88**, 014206 (2013).
- [22] Rahul Nandkishore and David A. Huse. “Many-body localization and thermalization in quantum statistical mechanics”. *Annual Review of Condensed Matter Physics* **6**, 15–38 (2015). [arXiv:https://doi.org/10.1146/annurev-conmatphys-031214-014726](https://doi.org/10.1146/annurev-conmatphys-031214-014726).
- [23] Ehud Altman. “Many-body localization and quantum thermalization”. *Nature Physics* **14**, 979–983 (2018).
- [24] Marcos Rigol, Vanja Dunjko, and Maxim Olshanii. “Thermalization and its mechanism for generic isolated quantum systems”. *Nature* **452**, 854–858 (2008).
- [25] Giulio Biroli, Corinna Kollath, and Andreas M. Läuchli. “Effect of rare fluctuations on the thermalization of isolated quantum systems”. *Phys. Rev. Lett.* **105**, 250401 (2010).
- [26] Lea F. Santos and Marcos Rigol. “Onset of quantum chaos in one-dimensional bosonic and fermionic systems and its relation to thermalization”. *Phys. Rev. E* **81**, 036206 (2010).
- [27] R. Steinigeweg, J. Herbrych, and P. Prelovšek. “Eigenstate thermalization within isolated spin-chain systems”. *Phys. Rev. E* **87**, 012118 (2013).
- [28] Hyungwon Kim, Tatsuhiko N. Ikeda, and David A. Huse. “Testing whether all eigenstates obey the eigenstate thermalization hypothesis”. *Phys. Rev. E* **90**, 052105 (2014).
- [29] R. Steinigeweg, A. Khodja, H. Niemeyer, C. Gogolin, and J. Gemmer. “Pushing the limits of the eigenstate thermalization hypothesis towards mesoscopic quantum systems”. *Phys. Rev. Lett.* **112**, 130403 (2014).
- [30] Keith R. Fratus and Mark Srednicki. “Eigenstate thermalization in systems with spontaneously broken symmetry”. *Phys. Rev. E* **92**, 040103 (2015).
- [31] Abdellah Khodja, Robin Steinigeweg, and Jochen Gemmer. “Relevance of the eigenstate thermalization hypothesis for thermal relaxation”. *Phys. Rev. E* **91**, 012120 (2015).
- [32] Rubem Mondaini and Marcos Rigol. “Eigenstate thermalization in the two-dimensional transverse field ising model. ii. off-diagonal matrix elements of observables”. *Phys. Rev. E* **96**, 012157 (2017).
- [33] Toru Yoshizawa, Eiki Iyoda, and Takahiro Sagawa. “Numerical large deviation analysis of the eigenstate thermalization hypothesis”. *Phys. Rev. Lett.* **120**, 200604 (2018).
- [34] David Jansen, Jan Stolpp, Lev Vidmar, and Fabian Heidrich-Meisner. “Eigenstate thermalization and quantum chaos in the holstein polaron model”. *Phys. Rev. B* **99**, 155130 (2019).
- [35] S. Trotzky, Y-A. Chen, A. Flesch, I. P. McCulloch, U. Schollwöck, J. Eisert, and I. Bloch. “Probing the relaxation towards equilibrium in an isolated strongly correlated one-dimensional bose gas”. *Nature Physics* **8**, 325–330 (2012).
- [36] Govinda Clos, Diego Porras, Ulrich War-

- ring, and Tobias Schaetz. “Time-resolved observation of thermalization in an isolated quantum system”. *Phys. Rev. Lett.* **117**, 170401 (2016).
- [37] Adam M. Kaufman, M. Eric Tai, Alexander Lukin, Matthew Rispoli, Robert Schittko, Philipp M. Preiss, and Markus Greiner. “Quantum thermalization through entanglement in an isolated many-body system”. *Science* **353**, 794–800 (2016).
- [38] G. Kucsko, S. Choi, J. Choi, P. C. Maurer, H. Zhou, R. Landig, H. Sumiya, S. Onoda, J. Isoya, F. Jelezko, E. Demler, N. Y. Yao, and M. D. Lukin. “Critical thermalization of a disordered dipolar spin system in diamond”. *Phys. Rev. Lett.* **121**, 023601 (2018).
- [39] Yijun Tang, Wil Kao, Kuan-Yu Li, Sangwon Seo, Krishnanand Mallayya, Marcos Rigol, Sarang Gopalakrishnan, and Benjamin L. Lev. “Thermalization near integrability in a dipolar quantum newton’s cradle”. *Phys. Rev. X* **8**, 021030 (2018).
- [40] J.R. Johansson, P.D. Nation, and Franco Nori. “Qutip: An open-source python framework for the dynamics of open quantum systems”. *Computer Physics Communications* **183**, 1760–1772 (2012).
- [41] J.R. Johansson, P.D. Nation, and Franco Nori. “Qutip 2: A python framework for the dynamics of open quantum systems”. *Computer Physics Communications* **184**, 1234–1240 (2013).
- [42] Aric A. Hagberg, Daniel A. Schult, and Pieter J. Swart. “Exploring network structure, dynamics, and function using networkx”. In Gaël Varoquaux, Travis Vaught, and Jarrod Millman, editors, Proceedings of the 7th Python in Science Conference. Pages 11 – 15. Pasadena, CA USA (2008). url: https://conference.scipy.org/proceedings/SciPy2008/paper_2/.
- [43] Feng Xia, Jiaying Liu, Hansong Nie, Yonghao Fu, Liangtian Wan, and Xiangjie Kong. “Random walks: A review of algorithms and applications”. *IEEE Transactions on Emerging Topics in Computational Intelligence* **4**, 95–107 (2020).
- [44] Henrik Wilming, Thiago R. de Oliveira, Anthony J. Short, and Jens Eisert. “Equilibration times in closed quantum many-body systems”. Page 435–455. Springer International Publishing. (2018).
- [45] James R. Garrison and Tarun Grover. “Does a single eigenstate encode the full hamiltonian?”. *Physical Review X* **8** (2018).
- [46] Peter Reimann. “Eigenstate thermalization: Deutsch’s approach and beyond”. *New Journal of Physics* **17**, 055025 (2015).
- [47] Tameem Albash and Daniel A. Lidar. “Adiabatic quantum computation”. *Reviews of Modern Physics* **90** (2018).
- [48] Philipp Hauke, Helmut G Katzgraber, Wolfgang Lechner, Hidetoshi Nishimori, and William D Oliver. “Perspectives of quantum annealing: methods and implementations”. *Reports on Progress in Physics* **83**, 054401 (2020).
- [49] Leo Zhou, Sheng-Tao Wang, Soonwon Choi, Hannes Pichler, and Mikhail D. Lukin. “Quantum approximate optimization algorithm: Performance, mechanism, and implementation on near-term devices”. *Phys. Rev. X* **10**, 021067 (2020).
- [50] Laba and Tkachuk. “Geometric characteristics of quantum evolution: curvature and torsion”. *Condensed Matter Physics* **20**, 13003 (2017).
- [51] Kh.P. Gnatenko, H.P. Laba, and V.M. Tkachuk. “Geometric properties of evolutionary graph states and their detection on a quantum computer”. *Physics Letters A* **452**, 128434 (2022).
- [52] Luca D’Alessio, Yariv Kafri, Anatoli Polkovnikov, and Marcos Rigol. “From quantum chaos and eigenstate thermalization to statistical mechanics and thermodynamics”. *Advances in Physics* **65**, 239–362 (2016).
- [53] Edward Farhi, David Gosset, Itay Hen, A. W. Sandvik, Peter Shor, A. P. Young, and Francesco Zamponi. “Performance of the quantum adiabatic algorithm on random instances of two optimization problems on regular hypergraphs”. *Physical Review A* **86** (2012).
- [54] Mark Jeanson and Joe Foley. “Review of the exponentially modified gaussian (emg) function since 1983”. *Journal of Chromatographic Science* **29**, 258–266 (1991).
- [55] Yuri Kalambet, Yuri Kozmin, Ksenia Mikhailova, Igor Nagaev, and Pavel

- Tikhonov. “Reconstruction of chromatographic peaks using the exponentially modified gaussian function”. *Journal of Chemometrics* **25**, 352–356 (2011).
- [56] Stephen J. Blundell and Katherine M. Blundell. “Concepts in Thermal Physics”. *Oxford University Press*. (2009).
- [57] Elizabeth Crosson and Samuel Slezak. “Classical simulation of high temperature quantum ising models” (2020). [arXiv:2002.02232](https://arxiv.org/abs/2002.02232).
- [58] Maxime Dupont, Nicolas Didier, Mark J. Hodson, Joel E. Moore, and Matthew J. Reagor. “Entanglement perspective on the quantum approximate optimization algorithm”. *Physical Review A* **106** (2022).
- [59] J M Deutsch. “Thermodynamic entropy of a many-body energy eigenstate”. *New Journal of Physics* **12**, 075021 (2010).
- [60] J. M. Deutsch, Haibin Li, and Auditya Sharma. “Microscopic origin of thermodynamic entropy in isolated systems”. *Phys. Rev. E* **87**, 042135 (2013).
- [61] Lea F. Santos, Anatoli Polkovnikov, and Marcos Rigol. “Entropy of isolated quantum systems after a quench”. *Phys. Rev. Lett.* **107**, 040601 (2011).
- [62] Michael A. Nielsen and Isaac L. Chuang. “Quantum computation and quantum information: 10th anniversary edition”. *Cambridge University Press*. (2010).
- [63] Edward Farhi, Jeffrey Goldstone, and Sam Gutmann. “A quantum approximate optimization algorithm” (2014). [arXiv:1411.4028](https://arxiv.org/abs/1411.4028).
- [64] Milena Grifoni and Peter Hänggi. “Driven quantum tunneling”. *Physics Reports* **304**, 229–354 (1998).
- [65] Masahito Ueda. “Quantum equilibration, thermalization and prethermalization in ultracold atoms”. *Nature Reviews Physics* **2**, 669–681 (2020).
- [66] Luca D’Alessio and Anatoli Polkovnikov. “Many-body energy localization transition in periodically driven systems”. *Annals of Physics* **333**, 19–33 (2013).
- [67] Luca D’Alessio and Marcos Rigol. “Long-time behavior of isolated periodically driven interacting lattice systems”. *Physical Review X* **4** (2014).
- [68] Achilleas Lazarides, Arnab Das, and Roderich Moessner. “Equilibrium states of generic quantum systems subject to periodic driving”. *Phys. Rev. E* **90**, 012110 (2014).
- [69] Keith R. Fratus and Mark Allen Srednicki. “Eigenstate thermalization and spontaneous symmetry breaking in the one-dimensional transverse-field ising model with power-law interactions” (2016). [arXiv:1611.03992](https://arxiv.org/abs/1611.03992).
- [70] Attila Felinger, Tamás Pap, and János Inczédy. “Curve fitting to asymmetrical chromatograms by the extended kalman filter in frequency domain”. *Talanta* **41**, 1119–1126 (1994).
- [71] K. F. Riley, M. P. Hobson, and S. J. Bence. “Mathematical methods for physics and engineering: A comprehensive guide”. *Cambridge University Press*. (2006). 3 edition.
- [72] Brian C. Hall. “An elementary introduction to groups and representations” (2000). [arXiv:math-ph/0005032](https://arxiv.org/abs/math-ph/0005032).
- [73] Michael M. Wolf, Frank Verstraete, Matthew B. Hastings, and J. Ignacio Cirac. “Area laws in quantum systems: Mutual information and correlations”. *Phys. Rev. Lett.* **100**, 070502 (2008).
- [74] Martin Kliesch and Arnau Riera. “Properties of thermal quantum states: Locality of temperature, decay of correlations, and more”. In *Fundamental Theories of Physics*. Pages 481–502. Springer International Publishing (2018).
- [75] S.H. Simon. “The oxford solid state basics”. OUP Oxford. (2013).

A Details for the short time limit

A.1 Calculating the torsion

Here we provide the details for the short time limit for CTQW. Firstly, we calculate the torsion, and the related quantity curvature. The curvature is a local measure of how much the quantum evolution deviates from a geodesic. These are conserved quantities valid for the whole evolution. We then show how this leads to short time predictions on $\langle H_p \rangle$ for these problems in Sec. A.2.

From Laba et al. [50] the curvature of the wave-function, for a time-independent Hamiltonian is given by:

$$\mathcal{C} = \langle \Delta H^4 \rangle - \langle \Delta H^2 \rangle^2, \quad (43)$$

where $\Delta H = H - \langle H \rangle$. The torsion is given by:

$$\mathcal{T} = \langle \Delta H^4 \rangle - \langle \Delta H^2 \rangle^2 - \frac{\langle \Delta H^3 \rangle^2}{\langle \Delta H^2 \rangle}, \quad (44)$$

Therefore, calculating the torsion and curvature reduces to evaluating $\langle \Delta H^j \rangle$, for $j = 2, 3, 4$. Since these are all conserved quantities, we can take the expectation with respect to the initial state, $|+\rangle$. The rest of this section details the rather tedious steps required to find these expectation values. Those willing to take the expression for \mathcal{T} at face-value are invited to skip to Sec. A.2.

The first step is to find $\langle H \rangle$:

$$\begin{aligned} \langle H \rangle &= \langle + | H_d + \gamma H_p | + \rangle \\ &= -n, \end{aligned}$$

since each $Z_i Z_j$ in H_p will flip qubits i and j to the minus state. Consequently, only H_d contributes to this expectation value.

Consider now:

$$\begin{aligned} \langle \Delta H^2 \rangle &= \langle H^2 \rangle - \langle H \rangle^2 \\ &= \langle H_d^2 \rangle + \gamma^2 \langle H_p^2 \rangle + \gamma \langle H_p H_d + H_d H_p \rangle - \langle H \rangle^2 \\ &= n^2 + \gamma^2 \kappa_2 + 0 + 0 - n^2 \\ &= \gamma^2 \kappa_2, \end{aligned}$$

where $\kappa_2 = |E|$ is the number of edges in the graph. It was pointed out by Gnatenko et al. [51], that $\langle H_p^2 \rangle$ is equal to the number of edges in the graph. To see why this is the case, we can write out this term explicitly to give:

$$\langle H_p^2 \rangle = \sum_{i_1, j_1, i_2, j_2} \langle + | Z_{i_1} Z_{j_1} Z_{i_2} Z_{j_2} | + \rangle,$$

where the sum is over the ordered pairs (i_k, j_k) , for $k = 1, 2$. The only possible way a term in the sum can be non-zero is if $i_1 = i_2$ and $j_1 = j_2$. Therefore the number of non-zero terms corresponds to the number of edges in the graph. The rest of the terms in $\langle \Delta H^2 \rangle$ are relatively trivial to calculate.

Extending the above logic, Gnatenko et al. [51] demonstrated that:

$$\langle H_p^3 \rangle = 6\kappa_3 \quad (45)$$

where κ_3 is the number of triangles in the graph, and

$$\langle H_p^4 \rangle = \kappa_2 + 3\kappa_2(\kappa_2 - 1) + 24\kappa_4, \quad (46)$$

where κ_4 is the number of squares in the graph.

Using the above results we can calculate $\langle \Delta H^3 \rangle$ and $\langle \Delta H^4 \rangle$. Starting with $\langle \Delta H^3 \rangle$:

$$\langle \Delta H^3 \rangle = \langle H^3 \rangle - 3\langle H^2 \rangle \langle H \rangle + 2\langle H \rangle^3. \quad (47)$$

Calculating $\langle H^3 \rangle$ gives:

$$\begin{aligned}\langle H^3 \rangle &= \langle (H_d + \gamma H_p)^3 \rangle \\ &= \langle H_d^3 + \gamma^2 H_d H_p^2 + \gamma^2 H_p^2 H_d + \gamma^3 H_p^3 + \gamma^2 H_p H_d H_p \rangle \\ &= -n^3 - 2n\gamma^2 \kappa_2 + 6\gamma^3 \kappa_3 + \gamma^2 \langle H_p H_d H_p \rangle,\end{aligned}$$

where we have neglected to write out terms that are trivially zero and made use of Eq. 45. To find $\langle \Delta H^3 \rangle$ it just remains to calculate $\langle H_p H_d H_p \rangle$:

$$\langle H_p H_d H_p \rangle = - \sum_{i_1, j_1, i_2, j_2, k} \langle + | Z_{i_1} Z_{j_2} X_k Z_{i_2} Z_{j_2} | + \rangle. \quad (48)$$

The X_k in each term is not going to change which terms are non-zero. It will only introduce a minus sign if k is i_2 or j_2 . Hence $\langle H_p H_d H_p \rangle$ is equal to:

$$\begin{aligned}&= - \sum_{i_2, j_2, k} \langle + | Z_{i_2} Z_{j_2} X_k Z_{i_2} Z_{j_2} | + \rangle \\ &= - \sum_{\substack{i_2, j_2 \\ k \neq i_2, j_2}} \langle + | Z_{i_2} Z_{j_2} Z_{i_2} Z_{j_2} | + \rangle + \sum_{j_2} \langle + | Z_k Z_{j_2} Z_k Z_{j_2} | + \rangle + \sum_{i_2} \langle + | Z_{i_2} Z_k Z_{i_2} Z_k | + \rangle \\ &= -(n-2)\kappa_2 + \kappa_2 + \kappa_2 \\ &= -(n-4)\kappa_2.\end{aligned}$$

Alternatively, we can argue as follows: in the absence of the X_k term, Eq. 48 would give $-n\kappa_2$. Accounting for the edge cases where $k = i_2$ or $k = j_2$ gives:

$$\langle H_p H_d H_p \rangle = -(n-4)\kappa_2. \quad (49)$$

Combining all of the above we are left with:

$$\langle \Delta H^3 \rangle = 2\gamma^2 (2\kappa_2 + 3\gamma\kappa_3). \quad (50)$$

To find \mathcal{C} and \mathcal{T} , it remains to find $\langle \Delta H^4 \rangle$.

$$\langle \Delta H^4 \rangle = \langle H^4 \rangle - 4\langle H \rangle \langle H^3 \rangle + 6\langle H \rangle^2 \langle H^2 \rangle - 4\langle H \rangle^3 \langle H \rangle + \langle H \rangle^4$$

Evaluating $\langle H^4 \rangle$ gives:

$$\begin{aligned}\langle H^4 \rangle &= \langle H_d^4 + \gamma^2 (H_d^2 H_p^2 + H_d H_p H_d H_p + H_d H_p^2 H_d + H_p H_d^2 H_p + H_p H_d H_p H_d + H_p^2 H_d^2) \\ &\quad + \gamma^3 (H_d H_p^3 + H_p^2 H_d H_p + H_p^3 H_d + H_p H_d H_p^2) + \gamma^4 H_p^4 \rangle.\end{aligned}$$

Again, we have neglected to write out the terms that evaluate to zero. Most of the terms in this expansion we have already evaluated, or are simple to evaluate. We are left with $\langle H_p H_d^2 H_p \rangle$, $\langle H_p H_d H_p^2 \rangle$, $\langle H_p^2 H_d H_p \rangle$, and $\langle H_p^4 \rangle$, the last of which can be evaluated with Eq. 46. Turning now to $\langle H_p H_d^2 H_p \rangle$:

$$\langle H_p H_d^2 H_p \rangle = \sum_{i_1, j_1, i_2, j_2, k_1, k_2} \langle + | Z_{i_1} Z_{j_1} X_{k_1} X_{k_2} Z_{i_2} Z_{j_2} | + \rangle.$$

Again the presence of X in the terms of the sum is not going to change which terms evaluate to be non-zero. Fixing the edge (i_2, j_2) , if either one of the X terms coincides with this edge and the other X term does not, it contributes -1 to the sum; there are $4(n-2)$ ways of this happening. If both X terms do not coincide with this edge then the term contributes 1; there are $(n-2)^2$ ways of this happening. Finally if both edges coincide, this edge contributes 1; there are four ways of this happening. The result is:

$$\langle H_p H_d^2 H_p \rangle = (n-4)^2 \kappa_2.$$

Thus, it remains to evaluate $\langle H_p H_d H_p^2 \rangle$ and $\langle H_p^2 H_d H_p \rangle$. Note that:

$$\langle H_p H_d H_p^2 \rangle^* = \langle H_p^2 H_d H_p \rangle.$$

To evaluate $\langle H_p H_d H_p^2 \rangle$, consider $\langle H_p^3 \rangle$. As before, the introduction of H_d is not going to change which terms in the sum are non-zero, if we were to write out a similar sum for this term, as in Eq. 48. Therefore, $\langle H_p H_d H_p^2 \rangle$ is going to depend on the number of triangles in the graph. The introduction of H_d is going to result in sign flips to some terms. The result is

$$\langle H_p H_d H_p^2 \rangle = -6\kappa_3(n-4). \quad (51)$$

Finally, we can assemble all the expectation values to find

$$\langle \Delta H^4 \rangle = \gamma^2 \left[-2\kappa_2(\gamma^2 - 8) + 3\gamma^2\kappa_2^2 + 24\gamma(2\kappa_3 + \gamma\kappa_4) \right]. \quad (52)$$

The resulting expression for the curvature is:

$$\mathcal{C} = 2\gamma^2 \left[-\kappa_2(\gamma^2 - 8) + \gamma^2\kappa_2^2 + 12\gamma(2\kappa_3 + \gamma\kappa_4) \right]. \quad (53)$$

and for the torsion:

$$\mathcal{T} = 2\gamma^4 \left[\kappa_2(\kappa_2 - 1) - \frac{18\kappa_3^2}{\kappa_2} + 12\kappa_4 \right]. \quad (54)$$

Note that the torsion is zero for a two-qubit system, with one edge, as expected from the spin-flip symmetry in the problem. The torsion is also zero for a ring of three qubits.

A.2 Short-time estimate for the performance

In the previous section we calculated the torsion. The torsion measures how much the wave-function deviates from a two-dimensional space in a time Δt . The error from deviating from this space is given by [51]:

$$\varepsilon_{2D} = \mathcal{T}\Delta t^4. \quad (55)$$

with $\varepsilon_{2D} = 0$ corresponding to remaining in the two-dimensional subspace and $\varepsilon_{2D} = 1$ corresponding to departing the subspace. If $\varepsilon_{2D} \ll 1$, then the two-dimensional approximation holds well. That is for times $t \ll \mathcal{T}^{-1/4}$, we can apply a two-level approximation.

At $t = 0$ the state of the system is $|\psi(t=0)\rangle = |+\rangle$. Integrating the system forward some infinitesimal time δt , gives $|\psi(t=\delta t)\rangle = (I - iH\delta t)|+\rangle$.

Applying the Gram-Schmidt procedure to the two vectors, $|\psi(t=0)\rangle$ and $|\psi(t=\delta t)\rangle$, gives the orthonormal basis for the subspace:

$$|e_0\rangle = |+\rangle \quad (56)$$

$$|e_1\rangle = \Delta H |+\rangle / \sqrt{\langle \Delta H^2 \rangle}, \quad (57)$$

where the expectation is with respect to the $|+\rangle$ state. Writing out H in this subspace is:

$$H^{(2D)} = \begin{pmatrix} \langle H \rangle & \frac{\langle H\Delta H \rangle}{\sqrt{\langle \Delta H^2 \rangle}} \\ \frac{\langle \Delta H H \rangle}{\sqrt{\langle \Delta H^2 \rangle}} & \frac{\langle \Delta H H \Delta H \rangle}{\langle \Delta H^2 \rangle} \end{pmatrix}, \quad (58)$$

which evaluates to

$$H^{(2D)} = \begin{pmatrix} -n & \sqrt{\gamma^2\kappa_2} \\ \sqrt{\gamma^2\kappa_2} & 2\left(2 + \frac{3\gamma\kappa_3}{\kappa_2}\right) - n \end{pmatrix} \quad (59)$$

$$(60)$$

$$= \sqrt{\gamma^2\kappa_2} X - \left(2 + \frac{3\gamma\kappa_3}{\kappa_2}\right) Z, \quad (61)$$

neglecting the uninteresting term proportional to the identity.

The problem Hamiltonian, within this space is:

$$H_p^{(2D)} = \begin{pmatrix} \langle H_p \rangle & \frac{\langle H_p \Delta H \rangle}{\sqrt{\langle \Delta H^2 \rangle}} \\ \frac{\langle \Delta H H_p \rangle}{\sqrt{\langle \Delta H^2 \rangle}} & \frac{\langle \Delta H H_p \Delta H \rangle}{\langle \Delta H^2 \rangle} \end{pmatrix}, \quad (62)$$

evaluating this gives:

$$H_p^{(2D)} = \begin{pmatrix} 0 & \text{sgn}(\gamma) \sqrt{\kappa_2} \\ \text{sgn}(\gamma) \sqrt{\kappa_2} & \frac{6\kappa_3}{\kappa_2} \end{pmatrix}, \quad (63)$$

where $\text{sgn}(\gamma)$ denotes the sign of γ .

Having now explicitly evaluated the relevant operators in this subspace, it remains to calculate the expectation of the problem Hamiltonian, which is:

$$\langle H_p^{(2D)}(t) \rangle = -4\gamma\kappa_2 \frac{\sin^2 \omega t}{\omega^2}, \quad (64)$$

where

$$\omega^2 = \gamma^2 \kappa_2 + \frac{(2\kappa_2 + 3\gamma\kappa_3)^2}{\kappa_2^2}. \quad (65)$$

B Spin-flip symmetry and the ETH

In Sec. 3 we pointed out MAX-CUT has a spin-flip symmetry. To recap the notation, given a CTQW for MAX-CUT with Hamiltonian H , then $[H, G] = 0$, where:

$$G = \prod_{i=1}^n X_i. \quad (66)$$

Typically, the system starts in the $|+\rangle$ state so the evolution is restricted to the plus one eigenspace of G for the entire evolution.

Despite this restriction, in Sec. 5, we used the ‘complete’ Hamiltonian as opposed to the Hamiltonian projected onto the correct symmetry sector to apply the ETH. This was numerically verified in Sec. 5.1. We justify this choice on the following grounds [18]:

1. G is a global operator and corresponds to a global symmetry.
2. We are interested in calculating the expectation values of local interactions only.

It is well established within the ETH literature that local symmetries can prevent thermalisation [18]. In contrast the spin-flip symmetry is a global symmetry. In general, there are a large number of globally conserved quantities for evolution under a constant Hamiltonian. If H belongs to a Hilbert Space \mathcal{H} with dimension $\dim \mathcal{H}$, then there are at least $\dim \mathcal{H}$ globally conserved quantities [18]. To see this, let the projector onto each eigenstate of H be denoted by P_k with $k = 1, \dots, \dim \mathcal{H}$, then [18]:

$$[H, P_k] = 0 \quad (67)$$

$$[P_k, P_j] = 0. \quad (68)$$

Hence $\langle P_k \rangle$ are globally conserved quantities. Despite the large number of conserved quantities, we generally do not expect these global symmetries to have much impact on the dynamics.

Though the above argument suggests we should be careful when considering the role of globally conserved quantities, it fails to consider the specific case of the spin-flip symmetry which we know

restricts dynamics to the positive eigenspace of G for a CTQW. Here we emphasise again that we are only interested in local quantities. A locally measurable consequence of the spin-flip symmetry is that:

$$\begin{aligned}\langle Z_i(t) \rangle &= \langle + | U^\dagger Z_i(t) U | + \rangle \\ &= \langle + | G U^\dagger Z_i(t) U G | + \rangle \\ &= \langle + | U^\dagger G Z_i(t) G U | + \rangle \\ &= - \langle + | U^\dagger G Z_i(t) G U | + \rangle \\ &= - \langle Z_i(t) \rangle,\end{aligned}$$

where we have used the initial state being an eigenstate of G and $[U, G] = 0$, consequently $\langle Z_i(t) \rangle = 0$.

Since $\langle Z_i(t) \rangle$ is a local observable, this should agree with the result if we were to replace the unitary evolution with a Gibbs distribution:

$$\langle Z_i(t) \rangle_\beta \propto \text{Tr} \left(Z_i e^{-\beta H} \right). \quad (69)$$

As $[H, G] = 0$ this is also zero. Thus the thermal-state recovers the correct expectation value for any local observable that flips sign under conjugation by G (i.e., $G O_L G = -O_L$).

That is not to say that the spin-flip symmetry has no role in calculating expectation values. Consider the case where the initial state

$$|\psi_i\rangle = \cos \theta/2 |\varphi_+\rangle + e^{-i\phi} \sin \theta/2 |\varphi_-\rangle, \quad (70)$$

is a linear superposition of two states, where $G |\varphi_+\rangle = |\varphi_+\rangle$ and $G |\varphi_-\rangle = -|\varphi_-\rangle$. Then

$$\begin{aligned}\langle \psi_i | U^\dagger O_L U | \psi_i \rangle &= \cos^2 \theta/2 \langle \varphi_+ | U^\dagger O_L U | \varphi_+ \rangle + \sin^2 \theta/2 \langle \varphi_- | U^\dagger O_L U | \varphi_- \rangle \\ &\quad + e^{-i\phi} \cos \theta/2 \sin \theta/2 \langle \varphi_+ | U^\dagger O_L U | \varphi_- \rangle + c.c.\end{aligned} \quad (71)$$

If O_L also commutes with G then (as is the case with H_p) it follows that:

$$\langle \varphi_+ | U^\dagger O_L U | \varphi_- \rangle = - \langle \varphi_+ | G U^\dagger O_L U G | \varphi_- \rangle \quad (72)$$

$$= - \langle \varphi_+ | U^\dagger O_L U | \varphi_- \rangle, \quad (73)$$

hence $\langle \varphi_+ | U^\dagger O_L U | \varphi_- \rangle = 0$ and

$$\langle \psi_i | U^\dagger O_L U | \psi_i \rangle = \cos^2 \theta/2 \langle \varphi_+ | U^\dagger O_L U | \varphi_+ \rangle + \sin^2 \theta/2 \langle \varphi_- | U^\dagger O_L U | \varphi_- \rangle. \quad (74)$$

However, we have conjectured each (and numerically investigated) that each of the above matrix elements can be represented by replacing each unitary evolution with a thermal state. That is to say,

$$\langle \psi_i | U^\dagger O_L U | \psi_i \rangle \approx \cos^2 \theta/2 \frac{\text{Tr} \{ O_L e^{-\beta_+ H} \}}{\text{Tr} \{ e^{-\beta_+ H} \}} + \sin^2 \theta/2 \frac{\text{Tr} \{ O_L e^{-\beta_- H} \}}{\text{Tr} \{ e^{-\beta_- H} \}}, \quad (75)$$

therefore the system would need to be assigned two temperatures for each symmetry sector.

Throughout this paper we have tacitly assumed we are in the paramagnetic phase, justified by the associated high temperatures. Work on spontaneous symmetry breaking and the ETH can be found in [30, 69].

C Calculating the moments of the density of states

To better understand the density of states (DOS) it is possible to calculate moments of the distribution. We can use these as fitting parameters to the model for the DOS. We are interested in calculating the

moments of the distribution produced by the eigenenergies of the CTQW Hamiltonian for the graph $G = (V, E)$:

$$H_{QW} = H_d + \gamma H_p, \quad (76)$$

where:

$$H_d = - \sum_{i=1}^n X_i \quad (77)$$

$$H_p = \sum_{(i,j) \in E} Z_i Z_j, \quad (78)$$

where $n = |V|$ is the number of nodes in the graph. Note that the Hamiltonian is constructed of Pauli matrices, which are traceless [62].

Denoting the eigenenergies of H_{QW} as E_k , then the mean of the distribution of eigenenergies is given by:

$$\mu = \frac{1}{2^n} \sum_k E_k = \text{Tr } H_{QW}, \quad (79)$$

and since the Hamiltonian is traceless, this evaluates to zero.

Repeating this approach for the variance gives:

$$\begin{aligned} \sigma^2 &= \frac{1}{2^n} \sum_k E_k^2 \\ &= \frac{1}{2^n} \text{Tr } H_{QW}^2 \\ &= \frac{1}{2^n} \text{Tr} \left(H_d^2 + 2\gamma H_d H_p + \gamma^2 H_p^2 \right) \\ &= n + 0 + \gamma^2 \kappa_2, \end{aligned}$$

where κ_2 is the number of edges in the graph. Between the penultimate and final line, we have used that only terms that are equal to the identity will contribute to the trace.

The same approach can be used to find the third moment:

$$\begin{aligned} \frac{1}{2^n} \sum_k E_k^3 &= \frac{1}{2^n} \text{Tr } H_{QW}^3 \\ &= \frac{1}{2^n} \text{Tr} \left(\gamma^3 H_p^3 \right) \\ &= 6\gamma^3 \kappa_3, \end{aligned}$$

where κ_3 is the number of triangles in the graph. Further details can be seen in [51] or Appendix A. The fourth moment also follows from the above logic:

$$\begin{aligned} \frac{1}{2^n} \sum_k E_k^4 &= \frac{1}{2^n} \text{Tr } H_{QW}^4 \\ &= \frac{1}{2^n} \text{Tr} \left(H_d^4 + 4\gamma^2 H_d^2 H_p^2 + 2\gamma^2 H_d H_p H_d H_p + \gamma^4 H_p^4 \right) \\ &= \left(n^2 + 2(n^2 - n) \right) + 4\gamma^2 n \kappa_2 + 2\gamma^2 (n - 4) \kappa_2 + \gamma^4 (\kappa_2 + 3\kappa_2(\kappa_2 - 1) + 24\kappa_4). \end{aligned}$$

From here it is straightforward to calculate the skewness and excess kurtosis. In theory, this process of moments could be continued to higher orders, incorporating loops of greater lengths. However, as shown above, even by the fourth order this becomes cumbersome.

D Estimating the temperature

D.1 Gaussian density of states

In Sec. 5.1 we numerically verified that CTQW were well approximated by a thermal state $\rho_{th}(\beta)$, with the temperature fixed by:

$$\text{Tr}\{H_{QW}\rho_{th}(\beta)\} = -n, \quad (80)$$

in agreement with the ETH. In this section we provide the details for estimating β and associated quantities, under the following assumptions:

1. **Continuity:** the energy levels can be modelled as a continuous variable.
2. **Normality:** the density of states is well modelled by a normal distribution.

Under these assumptions, we can take the density of states to be:

$$g(E)dE = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{E^2}{2\sigma^2}} dE, \quad (81)$$

where $\sigma^2 = \text{Tr}(H_{QW}^2)/2^n = n + \gamma^2\kappa_2$. The density of states has been normalised to one. Calculating the partition function gives:

$$\begin{aligned} \mathcal{Z} &= \int_{-\infty}^{\infty} e^{-\beta E} g(E) dE \\ &= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} e^{-\frac{E^2}{2\sigma^2}} e^{-\beta E} dE \\ &= e^{\beta^2\sigma^2/2}. \end{aligned}$$

Using $\langle E \rangle = -\partial \ln \mathcal{Z} / \partial \beta$, gives:

$$\langle E \rangle = -\beta\sigma^2. \quad (82)$$

Setting this equal to $-n$ gives

$$\beta = \frac{n}{n + \gamma^2\kappa_2}. \quad (83)$$

This is our estimate for β , in the thermodynamic limit. From here the entropy follows trivially from Eq. 26. Calculating $\langle H_p \rangle$ from Eq. 29 gives:

$$\langle H_p \rangle_{est} = -\beta\gamma\kappa_2. \quad (84)$$

Substituting Eq. 83 into the above equation gives Eq. 30. This can then be minimised to find the optimal choice of γ .

D.2 Exponentially modified Gaussian density of states

The steps for treating the EMG DOS are exactly the same as for the Gaussian distribution. Therefore in this section we only demonstrate how to fit the EMG and calculate the associated partition function.

The EMG distribution is a convolution of a Gaussian distribution with an exponential distribution [70]. Denoting the exponential distribution as:

$$f(x) = \begin{cases} 0 & \text{for } x < 0 \\ \lambda e^{-\lambda x} & \text{for } x \geq 0 \end{cases} \quad (85)$$

and the normal distribution as

$$g(x) = \frac{1}{\sqrt{2\pi\nu^2}} e^{-\frac{(x-m)^2}{2\nu^2}}, \quad (86)$$

then the exponentially modified Gaussian distribution is given by:

$$\begin{aligned}
h(x) &= f(x) * g(x) \\
&= \int_{-\infty}^{\infty} dy f(y)g(x-y) \\
&= \frac{\lambda}{2} e^{\frac{\lambda}{2}(2m+\lambda\nu^2-2x)} \operatorname{erfc}\left(\frac{m+\lambda\nu^2-x}{\sqrt{2\nu^2}}\right).
\end{aligned} \tag{87}$$

Through application of the convolution theorem [71] it is straightforward to write down the moment generating function (and hence partition function) from the moment generating functions of the exponential and Gaussian distributions. The resulting partition function in terms of the fitting parameters m , ν^2 and λ is:

$$\mathcal{Z}(\beta) = \left(1 + \frac{\beta}{\lambda}\right)^{-1} e^{-m\beta + \frac{1}{2}\nu^2\beta^2} \tag{88}$$

To find the fitting parameters we fit the moments of $h(x)$ to the moments associated with H_{QW} . Utilising the moment generating function to find the mean (μ), variance (σ^2) and skew (s) of $h(x)$ gives:

$$\mu = m + \frac{1}{\lambda} \tag{89}$$

$$\sigma^2 = \nu^2 + \frac{1}{\lambda^2} \tag{90}$$

$$s = \frac{2}{\nu^3\lambda^3} \left(1 + \frac{1}{\nu^2\lambda^2}\right)^{-3/2}. \tag{91}$$

The above equations can be inverted to find the fitting parameters in terms of the mean, variance and skew associated with H_{QW} . These properties can be extracted from H_{QW} by calculating $\operatorname{Tr}\{H_{QW}^k\}$ for $k = 1, 2, 3$. The method can be found in Appendix C. From here the same steps as Appendix D.1 can be followed to derive predictions for the temperature, entropy, $\langle H_p \rangle$ and optimal choices of γ

E Derivation of Eq. 29.

Using Eq. 29 removes the need for matrix exponentiation, hence here we discuss in greater detail the assumptions behind this equation.

We assume that the system is well modelled by a Gibbs state and consider the function

$$\Omega = e^{\beta H_{QW}} \partial_\gamma e^{-\beta H_{QW}}, \quad (92)$$

which we differentiate with respect to β , to get:

$$\begin{aligned} \partial_\beta \Omega &= e^{\beta H_{QW}} H_{QW} \partial_\gamma e^{-\beta H_{QW}} - e^{\beta H_{QW}} \partial_\gamma (H_{QW} e^{-\beta H_{QW}}) \\ &= -e^{\beta H_{QW}} \partial_\gamma (H_{QW}) e^{-\beta H_{QW}} \\ &= -e^{\beta H_{QW}} H_p e^{-\beta H_{QW}}. \end{aligned} \quad (93)$$

Now through successive differentiation or by simply applying the well-known result for writing the right-hand-side in terms of nested commutators [72] we conclude:

$$\Omega = -\beta H_p - \frac{\beta^2}{2} [H_{QW}, H_p] - \frac{\beta^3}{3!} [H_{QW}, [H_{QW}, H_p]] + \dots \quad (94)$$

where $[\cdot, \cdot]$ denotes the commutator. Tidying this expression up using the notation:

$$[H_{QW}^{(k)}, H_p] = \begin{cases} H_p & \text{if } k = 0 \\ \underbrace{[H_{QW}, \dots [H_{QW}, [H_{QW}, H_p]] \dots]}_{k \text{ times}} & \text{otherwise.} \end{cases}$$

gives:

$$\Omega = -\sum_{k=1}^{\infty} \frac{\beta^k}{k!} [H_{QW}^{(k-1)}, H_p]. \quad (95)$$

Acting on both sides with $e^{-\beta H_{QW}}$ and taking the trace gives:

$$\text{Tr}\{\partial_\gamma e^{-\beta H_{QW}}\} = -\sum_{k=1}^{\infty} \frac{\beta^k}{k!} \text{Tr}\{[H_{QW}^{(k-1)}, H_p] e^{-\beta H_{QW}}\}. \quad (96)$$

Evaluating the terms in the sum for which $k > 1$:

$$\begin{aligned} &\text{Tr}\{[H_{QW}^{(k-1)}, H_p] e^{-\beta H_{QW}}\} \\ &= \sum_j \langle E_j | [H_{QW}^{(k-1)}, H_p] e^{-\beta H_{QW}} | E_j \rangle \\ &= \sum_j \langle E_j | H_{QW}, [H_{QW}^{(k-2)}, H_p] | E_j \rangle e^{-\beta E_j} \\ &= \sum_j E_j \langle E_j | ([H_{QW}^{(k-2)}, H_p] - [H_{QW}^{(k-2)}, H_p]) | E_j \rangle e^{-\beta E_j} \\ &= 0, \end{aligned}$$

where $|E_j\rangle$ denotes the eigenvectors of H_{QW} . Therefore, once we assume the state is well modelled by a Gibbs distribution it follows that:

$$\text{Tr}\{\partial_\gamma e^{-\beta H_{QW}}\} = -\beta \text{Tr}\{H_p e^{-\beta H_{QW}}\}. \quad (97)$$

Commuting through the trace with the partial derivative and dividing both sides by the partition function gives Eq. 29.

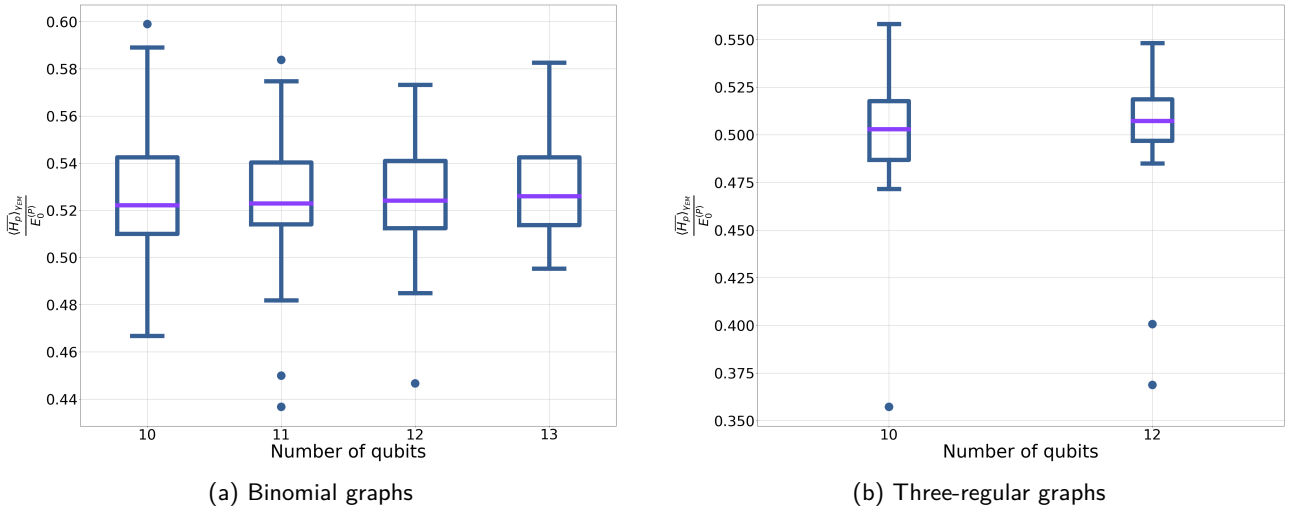


Figure 30: The performance (i.e., $\langle \bar{H}_p \rangle$) of the CTQW for the optimal γ predicted by the EMG DOS compared to the ground state energy of the problem Hamiltonian

F The performance of CTQWs for MAX-CUT

For the problem instances considered in Sec. 5 we consider how the value of $\langle \bar{H}_p \rangle$ evaluated at the γ predicted by the EMG DOS compares to the ground state energy of $\langle H_p \rangle$, (i.e. $E_0^{(P)}$). The results can be seen in Fig. 30. For the small problem sizes considered the performance seems largely independent of system size. This perhaps is unsurprising for a high temperature Gibbs state with a local Hamiltonian [73, 74]. We might expect different parts of the system to be weakly correlated, hence the CTQW is optimising locally and does not scale with problem size.

In practice one is perhaps less interested in how the average of the energy distribution of $\langle H_p \rangle$ compares with the absolute minimum, particularly if considering CTQWs as an exact solver. In such a case one might be more interested in the ground-state probability or time-to-solution. This requires some notion of run-time not discussed in this paper and will be the subject of future investigations.

G A brief introduction to Floquet systems

In this appendix we provide a very brief introduction to Floquet systems, with the aim of providing sufficient background to understand Sec. 7. This introduction draws from [64].

In this paper we have used Floquet system to refer to a closed quantum system with a time periodic Hamiltonian. A closed system will evolve under the Schrödinger equation:

$$i \frac{d}{dt} |\psi(t)\rangle = H(t) |\psi(t)\rangle. \quad (98)$$

We are interested in Hamiltonians that are time-periodic, such that:

$$H(t) = H(t + \tau), \quad (99)$$

where τ is the period of the Hamiltonian. The Floquet theorem then tells us that there exist solutions to the Schrödinger equation of the form

$$|\psi(t)\rangle = e^{-i\epsilon_\alpha t} |\phi_\alpha(t)\rangle, \quad (100)$$

where ϵ_α are termed quasi-energies and $|\phi_\alpha(t)\rangle$ the Floquet modes. The Floquet modes are time periodic, $|\phi_\alpha(t)\rangle = |\phi_\alpha(t + \tau)\rangle$. The quasi-energies are time-independent but are only uniquely defined up to multiples of $2\pi/\tau$. This is reminiscent of Bloch's theorem [75] for spatially periodic crystals.

Noting that

$$U(t + \tau, t) |\psi(t)\rangle = |\psi(t + \tau)\rangle \quad (101)$$

where $U(t + \tau, t)$ is the unitary associated with a single period. Substituting Eq. 100 into the above equation gives:

$$U(t + \tau, t) e^{-i\epsilon_\alpha t} |\phi_\alpha(t)\rangle = e^{-i\epsilon_\alpha(t+\tau)} |\phi_\alpha(t + \tau)\rangle \quad (102)$$

$$= e^{-i\epsilon_\alpha \tau} e^{-i\epsilon_\alpha t} |\phi_\alpha(t)\rangle. \quad (103)$$

Hence $|\phi_\alpha(t)\rangle$ is an eigenvector of a single period unitary with eigenvalue $e^{-i\epsilon_\alpha \tau}$. Therefore, by studying the single period unitary we can construct the wave function at all times, including the long-time average in Eq. 36.