A Solution to the Sign Problem Using a Sum-of-Few-Fermions Hamiltonian

David H. Wei*

Quantica Computing, LLC, San Jose, California

* Email: david.hq.wei@gmail.com

Abstract: A restricted path integral method is proposed to efficiently simulate a class of quantum systems and Hamiltonians called sum-of-few-fermions (SFF) by Monte Carlo on a classical computer without a numerical sign problem. Then a universality is proven to assert that any bounded-error quantum polynomial time (BQP) algorithm can be mapped to the ground state of an SFF Hamiltonian and becomes efficiently simulatable. Therefore, the computational complexity class of bounded-error probabilistic polynomial time (BPP) is precisely the same as the class of BQP.

The ability to simulate quantum systems efficiently on a programmable computer, classical [1,2] or quantum [1,3,4], is crucially important for fundamental sciences and practical applications. Of particular importance is to simulate the ground state of a many-body quantum system efficiently on a classical computer. Among many numerical methods, quantum Monte Carlo (QMC) [2] is uniquely advantageous as being based on first principles without uncontrolled systematic errors and using polynomially efficient importance sampling from an exponentially large Hilbert space, until its polynomial efficiency is spoiled by the notorious numerical sign problem [3,4]. On the other hand, a fundamental question in computational complexity theory is whether the complexity class of bounded-error probabilistic polynomial time (BPP) is the same as that of bounded-error quantum polynomial time (BQP) [5, 6], which will be answered affirmatively in this presentation by firstly identifying and characterizing a class of Hamiltonians called sum-of-few-fermions (SFF), whose Gibbs kernels and ground states are amenable to efficient Monte Carlo simulations on a classical computer, then demonstrating that a universal BQP algorithm can be mapped into an SFF Hamiltonian, therefore, QMC simulated efficiently [3].

Let a triple \((C, \mathcal{H}, \mathcal{B})\) represent a general quantum system [8], where \(C\) is a configuration space consisting of eigenvalues of a collection of dynamical variables associated with the quantum system, for example but not limited to spatial positions of particles, while \(\mathcal{H} \equiv \mathcal{H}(C) \subseteq L^2(C)\) is a Hilbert space of state vectors (i.e., wavefunctions) supported by \(C\), and \(\mathcal{B} \equiv B(\mathcal{H})\) is a Banach algebra of bounded operators acting on vectors in \(\mathcal{H}\), which contains a strongly continuous one-parameter semigroup of Gibbs operators \(\exp\{(-\tau H)\}_{\tau \in [0,\infty]}\), whose infinitesimal generator \(H\) is designated as the Hamiltonian for the quantum system. A Hamiltonian is the epitone of a general lower-bounded self-adjoint operator \(h\), called a partial Hamiltonian, whose lowest eigenvalue is denoted by \(\lambda_0(h) \in \mathbb{R}\), and the associated ground state is denoted by \(\psi_0(h)\), as either a non-degenerate wavefunction or an arbitrary representative of the equivalence class of vectors in the ground state subspace. It is assumed that all fundamental equations of physics, especially the Schrödinger and quantum field-theoretic equations, are nondimensionalized and written in a so-called natural unit system. It is without loss of generality (WLOG) to assume that \(C\) is a compact Riemannian manifold with a finite dimension \(\dim(C) \in \mathbb{N}\) and a finite diameter \(\text{diam}(C) \in \mathbb{R}\), thus a finite size \(C \equiv \dim(C) + \text{diam}(C)\), and all wavefunctions are real-valued, so the Hilbert spaces and Banach algebras are over \(\mathbb{R}\) [7,10].

Consider a many-fermion system (MFS) of a variable size, comprising a number \(S \in \mathbb{N}\) of fermion species, each species labeled by \(s \in [1, S]\) consisting of a number \(n_s \in \mathbb{N}\) of identical fermions moving on a low-dimensional Riemannian manifold \(\mathcal{X}_s\), where the total number of particles \(N_s = \sum_{s=1}^{S} n_s\) may go up unbounded, while both \(d_s = \dim(\mathcal{X}_s)\) and \(D_s = \text{diam}(\mathcal{X}_s)\), \(\forall s \in [1, S]\) are always bounded by a finite number independent of \(N_s\). Mathematically, the identical fermions of each species \(s \in [1, S]\) may be artificially labeled by an integer \(n \in [1, n_s]\), so that a configuration coordinate (point) \(q_s \equiv (q_{s1}, q_{s2}, \ldots, q_{sn_s}) \in \mathcal{X}_s^{n_s}\) can represent their spatial configuration. But physically, the indistinguishability among identical fermions dictates that all of the label-exchanged coordinates be equivalent and form an orbit \(G_s q_s \equiv \{\pi S q_s : \pi S \in G_s\}\) for any \(q_s \in \mathcal{X}_s^{n_s}\), where \(G_s\) is the symmetry group of permuting \(n_s\) labels, \(\pi S \in G_s\) is a typical permutation. Straightforwardly, the Cartesian product \(C \equiv \prod_{s=1}^{S} \mathcal{X}_s^{n_s}\) is a configuration space for the MFS, and the group direct product \(G_s \equiv \prod_{s=1}^{S} G_s\), called the exchange symmetry group of the MFS, acts on \(C\) and partitions it into disjoint
orbits. Clearly, every pair of permutations \( \pi \in G_s, s \in [1, S] \) and \( \pi' \in G_{s'}, s' \in [1, S] \) with \( s \neq s' \) commute, hence each \( G_s, s \in [1, S] \) is straightforwardly a normal subgroup of \( G_s \). All of the even permutations in \( G_s \) form a subgroup \( A_s \), called the exchange alternating group. It is an axiom of physics that any legitimate quantum state \( \psi \in \mathcal{H}(C) \) must be exchange-symmetric as \( [\pi \psi](q) \equiv \psi(\pi q), \forall q \in C, \forall \pi \in G_s \).

With respect to \( G_s \) and its actions on \( C \) and \( \mathcal{H}(C) \), an exchange symmetrization operator is defined as \( \mathcal{P} \equiv \sum \pi \in G_s (-1)^{\pi} \pi, \) up to a normalization constant.

For any partial Hamiltonian \( h \), and any \( (r, q, \tau) \in C^2 \times (0, \infty) \), let \( \langle r \rangle e^{-r \tau} |q \rangle \) represent an artificial, non-negative definite, boltzmannonic Gibbs transition amplitude from \( q \) to \( r \) in (imaginary) time \( \tau \) due to \( h \), which ignores the fermionic exchange symmetry and regards all particles distinguishable, let \( \langle r \rangle e^{-r \tau} |Pq \rangle \equiv \sum \pi \in G_s (-1)^{\pi} \tau h \pi |q \rangle \) or \( \langle r \rangle e^{-r \tau} |Pq \rangle \equiv \sum \pi \in G_s (-1)^{\pi} \tau h \pi |q \rangle \) denote a pre- or post-symmetrized fermionic Gibbs transition amplitude. The function \( |e^{-r \tau}| \) is called the boltzmannonic Gibbs kernel, and \( \langle e^{-r \tau} |P \rangle \in L^2(C^2) \) or \( \langle P |e^{-r \tau} \rangle \in L^2(C^2) \) is called the pre- or post-symmetrized fermionic Gibbs kernel, associated with a Gibbs operator \( \exp(-r h), \tau \in (0, \infty) \) generated by a partial Hamiltonian \( h \).

It is obvious that \( \langle e^{-r \tau} |P \rangle \equiv \langle P |e^{-r \tau} \rangle \), either one may be referred to as the fermionic Gibbs kernel.

Given an MFS governed by a Hamiltonian \( H \), a computationally important number is the (descriptive) size of \( H \), denoted by size(\( H \)), which is basically the minimum number of classical bits, up to a constant factor, that is needed to describe \( H \). All computational complexities and singular values of operators will be measured against size(\( H \)). Of great interests are the so-called (computationally) local Hamiltonians \([13]\) of the form \( H = \sum_{k=1}^K H_k, K \in \mathbb{N}, K = O(\text{poly}(N_s)), \) where the size of \( H \) is defined as size(\( H \)) \( \equiv \text{size}(C) + K, \) thus size(\( H \)) = \( O(\text{poly}(N_s)) \), each \( H_k \in [1, K] \) moves a small number of artificially labeled fermions. By an \( H_k, k \in [1, K] \) moving an \( n \)-th fermion of an \( s \)-th species, \( s \in [1, S], n \in [1, n_s], \) it is meant that there exist a \( \tau \in (0, \infty) \) and two configuration points \( q = (\cdots q_{sn} \cdots) \in C \) and \( r = (\cdots r_{sn} \cdots) \in C \), such that \( r_{sn} \neq q_{sn} \) while the boltzmannonic Gibbs transition amplitude \( \langle r \rangle e^{-\tau h_k} |q \rangle \neq 0 \).

One exemplary local Hamiltonian describes an MFS with a large number of identical particles, which is a conventional Schrödinger operator \( H = \frac{1}{2} \sum_n \frac{\partial^2}{\partial x_n^2} + V(x_1, \cdots, x_N), \) but rewritten into a local Hamiltonian as \( H = \sum_{n=1}^N H_n, \) with each partial Hamiltonian moving three artificially labeled fermions as \( H_n \equiv \frac{1}{2} \sum_{k=1}^{n-1} \frac{\partial^2}{\partial x_k^2}, + V(x_1, \cdots, x_N) / N_s, \forall n \in [1, N_s], \) where \( k \parallel N_s \) denotes the unique number in \([1, N_s] \) such that \( k \parallel N_s \equiv 0 \pmod{N_s}, \forall k \in \mathbb{N}. \)

Another exemplary local Hamiltonian \( H = \sum_{k=1}^K H_k \) describes an important class of MFS called a many-species fermionic system (MSFS), that comprises a large number \( S \in \mathbb{N} \) of fermion species, each of which has no more than a small constant of identical fermions, where each \( H_k, k \in [1, K] \) moves either all or none of the identical particles of any fermion species. For each \( k \in [1, K] \), the species and fermions moved by \( H_k \) constitute a subsystem called a few-fermion cluster, or just a few-fermion in short, which is associated with a factor subspace \( C_k \) and a factor subgroup \( G_k \leq G_s \), in the sense that, \( C_k \) and \( G_k \) are subspaces of \( C \) and \( G_s \), respectively, \( G_k \) is a normal subgroup of \( G_s \) and itself a direct product of a small number of normal subgroups from the list \( \{G_s\}_{s \in [1, s]} \).

**Definition 1.** Given an MSFS with a variable number \( S \in \mathbb{N} \) of fermion species, a form sum \( H = \sum_{k=1}^K H_k \), \( K = O(\text{poly}(S)) \) defining a local Hamiltonian is called a sum of few-fermions, when each \( H_k, k \in [1, K] \), called a few-fermion interaction (FFI), involves no more than a small constant number of fermion species, moves either all or none of the identical particles of each involved fermion species, and is invariant under any exchange of identical particles, namely, \( \pi^{-1} H_k \pi = H_k, \forall \pi \in G_s \).

Such MSFS and \( H \) are said to be sum-of-few-fermions (SFF), with sum-of-few-fermions (SFF) serving as an adjective. Although an SFF system may have a large total number of particles \( N_s \), each FFI \( H_k, k \in [1, K] \) corresponds to a few-fermion comprising a small number of fermions and species. Due to the small size of a few-fermion, it is always computationally easy, with the complexity bounded by a constant, to solve any \( H_k, k \in [1, k] \) or any associated Gibbs kernel or any related physics, either analytically or numerically.

**Definition 2.** Let \( H = \sum_{k=1}^K H_k \) be a form sum defining an SFF Hamiltonian. The form sum is called a Lie-Trotter-Kato (LTK) decomposition, and \( H \) is called LTK-decomposed, when \( \forall \epsilon > 0, \) there exists an \( m \in \mathbb{N}, m = O(\text{poly}(\text{size}(H) + e^{-1})), \) such that \( \langle r | \left( \prod_{k=1}^K e^{-H_k/m} \right)^m |Pq \rangle \in \langle r | e^{-H} |Pq \rangle (1 - \epsilon, 1 + \epsilon), \forall (r, q) \in C^2 \).
satisfying \(|\langle r | e^{-H} | pq \rangle | \geq \epsilon e^{-\lambda_0(H)}\). The same form sum is called a ground-state projection (GSP) decomposition, and \(H\) is called GSP-decomposed, when \(\forall \epsilon > 0\), there exists an \(m \in \mathbb{N}\), \(m = O(\text{poly}(\text{size}(H) + \epsilon^{-1}))\), such that \(\| \prod_{k=1}^n I_k \| < \epsilon\), where \(\| \cdot \|\) denotes the operator norm, \(I_k \equiv \lim_{\tau \to \infty} e^{-\tau[H-k\lambda_{\infty}(H)]}\) and \(I_k \equiv \lim_{\tau \to -\infty} e^{-\tau[H_k-k\lambda_{\infty}(H)]}\) are projections to the ground state subspaces of \(H\) and \(H_k\), \(k \in [1, K]\).

The definition of an LTK- or GSP-decomposed Hamiltonian is inspired by the LTK product formula \(e^{-\tau H} = \lim_{m \to \infty} \left\{ \prod_{k=1}^m e^{-\tau H_k/m} \right\} \), \(\tau \in (0, \infty)\), which suggests to divide \([0, \tau]\) into time intervals delimited by time instants \(\tau_n \equiv (n/K)\), \(\tau_n \in [0, N]\), \(\tau_n \equiv n/m\), \(N \equiv mK\), and break the Gibbs operator \(e^{-\tau H}\) down into a sequence of Gibbs operators \(G_n \equiv e^{-\delta\tau H_n|k}\) \(n \in [1, N]\), so to compute Gibbs kernels using the Feynman path integral, also known as the functional integration \([13,14]\), of equation (3), even though the whole group \(G\) of Feynman paths yields a non-negative definite, boltzmannonic Gibbs transition amplitude. On the right side of equation (3), even though the whole group \(G\) is used for the domain of exchange symmetry, it is really only those permutations in the subgroup \(G_{n1:n2}\) generated by the set of subgroups \(\{G_n = G_n|k|_n \}_{n1:n2}\) that are active and relevant, since \(\rho(q_{n1:n2}; \pi q_{n1:n2}; \pi_{n1:n2}) = 0\) for all \(\pi \notin G_{n1:n2}\).

A number of consecutive Feynman slabs with the corresponding sequence of Gibbs operators \(G_n\), \(n \in [n1, n2]\), \(0 < n1 \leq n2 \leq N\) constitute a Feynman stack with the two Feynman planes \((C_{n1}, \tau_{n1})\) and \((C_{n2}, \tau_{n2})\) forming its boundaries \([8]\). Let \(n0 \equiv n1 - 1, \forall n1 \in N\). Pick two points \(q_{n0} \in C_{n0}\) and \(q_{n2} \in C_{n2}\) on the two boundary Feynman planes, the set of all Feynman paths

\[
\Gamma(q_{n2}, \tau_{n2}; q_{n0}, \tau_{n0}) \equiv \{ \gamma(t) : t \in [\tau_{n0}, \tau_{n2}] \to \mathcal{C} \} \text{ such that } \gamma(\tau_{n0}) = q_{n0}, \gamma(\tau_{n2}) = q_{n2}
\]

(1)

together with the sequence of partial Hamiltonians \(\{H_n|k|\}_{n \in [n1, n2]}\) constitute a Feynman spindle \([8]\), which gives rise to two Gibbs transition amplitudes, one being bôttzmannonic and defined as

\[
\rho(q_{n2}, \tau_{n2}; q_{n0}, \tau_{n0}) \equiv \int \cdots \int \left\{ \Pi^{n2-1}_{n=n1} dq_n \right\} \left\{ \Pi^{n2-1}_{n=n1} \right\}.
\]

(2)

the other being fermionic and symmetrized as

\[
\rho(q_{n2}, \tau_{n2}; \mathcal{P}q_{n0}, \tau_{n0}) \equiv \sum_{\pi \in G_{\pi}} (-1)\pi \rho(q_{n2}, \tau_{n2}; \pi q_{n0}, \tau_{n0}).
\]

(3)

In equation (3), the fermionic exchange symmetry is ignored and the integration over all particle-distinguished Feynman paths yields a non-negative definite, bôttzmannonic Gibbs transition amplitude. On the right side of equation (3), even though the whole group \(G\) is used for the domain of exchange symmetry, it is really only those permutations in the subgroup \(G_{n1:n2}\) generated by the set of subgroups \(\{G_n = G_n|k|_n \}_{n1:n2}\) that are active and relevant, since \(\rho(q_{n1:n2}; \pi q_{n1:n2}, \pi_{n1:n2}) = 0\) for all \(\pi \notin G_{n1:n2}\).

For Feynman slabs or slices that are sufficiently thin, there are simple rules for Feynman flights \([8]\), which determine the associated Gibbs transition amplitude between two points \(q\) and \(r\) on two narrowly separated Feynman planes respectively \([13,14]\). It is almost always the case, here taken as an axiomatic premise, that all partial Hamiltonians in consideration are of the Schrödinger type, as a sum of an elliptic differential operator \(-\Delta\) and a bounded potential \(V\), so to substantiate the Hopf lemma and the strong Hopf extremum principle \([8,15,16]\), and ensure that the bôttzmannonic Gibbs transition amplitude of a Feynman flight becomes exponentially localized as \(|\langle r |e^{-\Delta+V}\rangle| \leq c_1 e^{-c_2 \text{dist}(r,q)^2/\delta r^3}\) when \(\delta r > 0\) approaches 0, where \(c_1, c_2, c_3\) are positive constants, \(\text{dist}(r,q)\) denotes the geodesic distance between \(q\) and \(r\) on the Riemannian manifold \(\mathcal{C}\). Said rules for Feynman flights induce a Wiener measure that assigns a non-negative amplitude \(W(\gamma) = e^{-U(\gamma)}\) to each Feynman path \(\gamma\). Importantly, the action functional \(U(\gamma)\) is linear with respect to path concatenation, in that \(U(\gamma) = U(\gamma_1) + U(\gamma_2)\) holds when two segments of Feynman paths \(\gamma_1\) and \(\gamma_2\) concatenate into a continuous Feynman path \(\gamma \equiv \gamma_2 \ast \gamma_1\), which starts at the start point of \(\gamma_1\), goes to the end point of \(\gamma_1\) that coincides with the start point of \(\gamma_2\), and continues till the end point of \(\gamma_2\).

The formulation of Feynman path integral is rightly suited for simulating a Gibbs kernel via Monte Carlo integration over a many- but finite-dimensional space. Path integral Monte Carlo would realize BPP simulations of quantum systems, were it not for the sign problem \([5,6]\) due to the presence of negative amplitudes, particularly in fermionic systems. QMC methods using restricted path integrals (RPIs) \([6,17,20]\) have been proposed and applied to avoid negative amplitudes. But previous RPIs are only approximate methods as

3
they rely on a priori approximations for the nodal surfaces of Gibbs kernels associated with the total Hamiltonian, which are unknown and hard to compute. Here I will show that for an SFF Hamiltonian, negative amplitudes can be avoided by restricting Feynman paths locally, with respect to the efficiently computable nodal surface of a Gibbs kernel associated with an individual FFI.

By Feynman’s rule of amplitude multiplication for events occurring in succession [13], also known as the Chapman-Kolmogorov equation in probability theory, in conjunction with the (anti)symmetry of Gibbs kernels under particle exchanges, the fermionic Gibbs transition amplitude due to a Feynman stack, from a start point \( q_{\text{no}} \in C_{n_0} \) at \( \tau_{\text{no}} \), \( n_0 = n_1 - 1 \) to an end point \( q_{n_2} \in C_{n_2} \) at \( \tau_{n_2} \), can be computed as

\[
\rho(q_{n_2}, \tau_{n_2}; P_{q_{n_0}, \tau_{n_0}}) = \int_{q_{1}} \rho(q_{n_2}, \tau_{n_2}; q_{n_1}, \tau_{n_1}) \rho(q_{n_1}, \tau_{n_1}; P_{q_{n_0}, \tau_{n_0}}) dq_{n_1}
\]

\[
= \int_{q_{1}} \rho(q_{n_2}, \tau_{n_2}; q_{n_1}, \tau_{n_1}) \rho(q_{n_1}, \tau_{n_1}; P_{q_{n_0}, \tau_{n_0}}) dq_{n_1}
\]

\[
\propto \int_{q_{1}} \rho(q_{n_2}, \tau_{n_2}; q_{n_1}, \tau_{n_1}) \rho(q_{n_1}, \tau_{n_1}; P_{q_{n_0}, \tau_{n_0}}) dq_{n_1}
\]

\[
\propto \int_{q_{1}} \rho(q_{n_2}, \tau_{n_2}; q_{n_1}, \tau_{n_1}) \rho(q_{n_1}, \tau_{n_1}; P_{q_{n_0}, \tau_{n_0}}) dq_{n_1}
\]

(4)

where \( \mathcal{R}_{n_1} \triangleq \mathcal{R}_{n_1}(q_{n_0}, \tau_{n_0}) \triangleq \{ q_{n_1} \in C_{n_1} : \rho(q_{n_1}, \tau_{n_1}; q_{n_0}, \tau_{n_0}) > 0 \} \subseteq C_{n_1} \) is called the boltzmannonic range of \( q_{n_0} \) from \( \tau_{n_0} \) to \( \tau_{n_1} \). “\( \propto \)” indicates an equality up to a universal (that is, (\( q_{n_2}; q_{n_0}, \tau_{n_0} \))-independent) constant, \( \mathcal{R}_{n_1}/G_s \) is any subset of \( \mathcal{R}_{n_1} \) that tiles up \( \mathcal{R}_{n_1} \) under the action of \( G_s \), i.e., \( \bigcup_{q_1 \in \mathcal{R}_{n_1}/G_s} \rho(q_1) \subseteq \mathcal{R}_{n_1} \). It suffices to choose \( \mathcal{R}_{n_1}/G_s = \mathcal{N}_{n_1}^+(q_{n_0}, \tau_{n_0}) \triangleq \{ q_{n_1} \in \mathcal{R}_{n_1} : \rho(q_{n_1}, \tau_{n_1}; P_{q_{n_0}, \tau_{n_0}}) > 0 \} \), which is the set union of all of the positive nodal cells of \( \rho(q_{n_1}, \tau_{n_1}; P_{q_{n_0}, \tau_{n_0}}) \) regarded as a function of \( q_{n_1} \in \mathcal{R}_{n_1} \), having \( q_{n_0}, \tau_{n_0}, \tau_{n_1} \) fixed. Similarly, let \( \mathcal{N}_{n_1}^-(q_{n_0}, \tau_{n_0}) \triangleq \{ q_{n_1} \in \mathcal{R}_{n_1} : \rho(q_{n_1}, \tau_{n_1}; P_{q_{n_0}, \tau_{n_0}}) < 0 \} \) denote the set union of all of the negative nodal cells. Clearly, any odd permutation \( \pi \in G_s \setminus A_s \) induces a one-to-one correspondence between the set of positive nodal cells and the set of negative nodal cells. Therefore,

\[
\rho(q_{n_2}, \tau_{n_2}; P_{q_{n_0}, \tau_{n_0}}) \propto \int_{q_{1}} \rho(q_{n_2}, \tau_{n_2}; q_{n_1}, \tau_{n_1}) \rho(q_{n_1}, \tau_{n_1}; P_{q_{n_0}, \tau_{n_0}}) dq_{n_1}
\]

(5)

In general, it has been rigorously proved that, under the action of the group \( G_s \), any nodal cell \( \mathcal{N}_{n_1} \) tiles up the configuration subspace \( \mathcal{R}_{n_1} \), thus, \( \mathcal{N}_{n_1}^+(q_{n_0}, \tau_{n_0}) = A_s \mathcal{N}_{n_1}(q_{n_0}, \tau_{n_0}) \triangleq \bigcup_{q_1} \rho \mathcal{N}_{n_1}(q_{n_0}, \tau_{n_0}) \), provided that the FFIs are of the Schrödinger type and substantiate the Hopf lemma and the strong Hopf extremum principle [8]. Moreover, any wavefunction or Gibbs kernel over \( \mathcal{R}_{n_1} \) is fully determined by and easily recovered from its restriction on \( \mathcal{N}_{n_1}^+(q_{n_0}, \tau_{n_0}) \).

Now consider \( \rho(q, \tau; P_{q_{n_0}, \tau_{n_0}}) \triangleq \langle q | e^{-iH_{\text{no}} \delta_1} K \Gamma | P_{q_{n_0}} \rangle \) as a \( (q, \tau) \)-jointly continuous function of \( (q, \tau) \in C \times (\tau_{n_0}, \tau_{n_1} + \delta_1) \), \( \delta_1 > 0 \). The preimage \( \{(q, \tau) : \rho(q, \tau; P_{q_{n_0}, \tau_{n_0}}) > 0 \} \) is an open set, the connected component of which containing the trivial node-free path \( \{ (q_{n_0}, \tau) : \tau \in (\tau_{n_0}, \tau_{n_1}) \} \) is called the nodal tube or the Ceperley reach of \( q_{n_0} \in C_{n_0} \) [8] which is denoted by \( \mathcal{T}_{n_1}(q_{n_0}, \tau_{n_0}) \). Let \( \mathcal{T}_n(q; q_{n_0}, \tau_{n_0}) \triangleq \mathcal{T}_n(q_{n_0}, \tau_{n_0}) \cap C \times \{ \tau \} \), \( \forall \tau \in (\tau_{n_0}, \tau_{n_1}) \), which is clearly a nodal cell of the fermionic Gibbs kernel \( \rho(q, \tau; P_{q_{n_0}, \tau_{n_0}}) \) as a function of \( q \in C \). It is well known that, inside the nodal tube \( \mathcal{T}_n(q_{n_0}, \tau_{n_0}) \), the fermionic kernel \( \rho(q, \tau; P_{q_{n_0}, \tau_{n_0}}) \) coincides with an RPI \( \rho_{\text{RPI}}(q, \tau; q_{n_0}, \tau_{n_0}) \), which is a boltzmannonic path integral over all particle-distinguished Feynman paths that lie fully within the nodal tube \( \mathcal{T}_n(q_{n_0}, \tau_{n_0}) \) [8][17][20].

In a hypothetical procedure of numerical integration, the fermionic Gibbs kernel \( \rho(q_{n_1}, \tau_{n_1}; P_{q_{n_0}, \tau_{n_0}}) \) as a function of \( q_{n_1} \in \mathcal{N}_{n_1}^+(q_{n_0}, \tau_{n_0}) \) may be sampled using two nested loops, where an outer loop moves the configuration coordinate \( q_{n_1} \in \mathcal{N}_{n_1}^+(q_{n_0}, \tau_{n_0}) \), while an inner loop regards \( q_{n_0} \) and \( q_{n_1} \) being fixed, samples from the set \( \Gamma(q_{n_1}, \tau_{n_1}; G_s, q_{n_0}, \tau_{n_0}) \triangleq \bigcup_{\pi \in G_s} \Gamma(q_{n_1}, \tau_{n_1}; \pi q_{n_0}, \tau_{n_0}) \) of Feynman paths that connect a start point in the orbit \( G_s q_{n_0} \) to the end point \( q_{n_1} \), and integrates the signed Wiener measure \( (-1)^{n_1} W(\gamma) \) for each Feynman path \( \gamma \) connecting a certain \( \pi q_{n_0}, \pi \in G_s \) to \( q_{n_1} \). It turns out that any Feynman path crossing or making contact with the boundary of the nodal tubes \( \{ \partial \mathcal{T}_n(\pi q_{n_0}, \tau_{n_0}) \} \pi \in A_s \) belongs to a \( G_s \)-orbit of sets of Feynman paths that integrate to a net zero amplitude, so all of which can be removed from path integration. Specifically, let \( \gamma_1 \) denote a segment of Feynman path connecting a point \( (q, \tau) \in \partial \mathcal{T}_n(q_{n_0}, \tau_{n_0}) \),
\( \pi \in (\tau_{n_0}, \tau_{n_1}) \) to the end point \( q_{n_1} \), then for each \( \pi \in G_* \), the set of concatenated Feynman paths

\[
\gamma_1 \ast \Gamma(q, \tau; \pi q_{n_0}, \tau_{n_0}) \overset{\text{def}}{=} \{ \gamma_1 \ast \gamma_0 : \gamma_0 \in \Gamma(q, \tau; \pi q_{n_0}, \tau_{n_0}) \},
\]

together with the relevant FFIs constitute a **post-tethered Feynman spindle**, which yields a Wiener measure \( \int_{\gamma_0 \in \Gamma(q, \tau; \pi q_{n_0}, \tau_{n_0})} W(\gamma_1) W(\gamma_0) d\gamma_0 = W(\gamma_1) \rho(q, \tau; \pi q_{n_0}, \tau_{n_0}) \). With \( \pi \) traversing the group \( G_* \), it enumerates a \( G_* \)-orbit of post-tethered Feynman spindles associated with \( \{ \gamma_1 \ast \Gamma(q, \tau; \pi q_{n_0}, \tau_{n_0}) \}_{\pi \in G_*} \), whose corresponding Wiener measures should be signed accordingly and summed up to yield

\[
\sum_{\pi \in G_*} W(\gamma_1) (-1)^\pi \rho(q, \tau; \pi q_{n_0}, \tau_{n_0}) = W(\gamma_1) \rho(q, \tau; \pi q_{n_0}, \tau_{n_0}) = W(\gamma_1) \rho(Pq, q; \tau_{n_0}, \tau_{n_0}),
\]

which makes a net zero contribution to \( \rho(q_{n_1}, \tau_{n_1}; \pi q_{n_0}, \tau_{n_0}) \), when \( (q, \tau) \in \partial T_{n_1}(q_{n_0}, \tau_{n_0}) \). Moreover, it is obvious that the same analysis applies to a more general \( G_* \)-orbit of post-tethered Feynman spindles associated with \( \{ \gamma'_1 \ast \Gamma(q, \tau; \pi q_{n_0}, \tau_{n_0}) \}_{\pi \in G_*} \), where \( \tau \in (\tau_{n_0}, \tau_{n_1}) \), and \( \gamma'_1 \) is a Feynman path connecting the point \( (q, \tau) \) to any point \( (q', \tau') \in \mathcal{C} \times (\tau_{n_1}, \tau_N) \), including \( (q', \tau') = (q_{N_1}, \tau_N) \), so to conclude that any Feynman path crossing or making contact with the boundary of the nodal tube \( \partial T_{n_1}(q_{n_0}, \tau_{n_0}) \) belongs to a \( G_* \)-orbit of sets of Feynman paths that make a net zero contribution to the Gibbs transition amplitude \( \rho(q', \tau'; Pq_{n_0}, \tau_{n_0}) \), thus can be safely removed from path integration. The only thing to note is that the calculation of the Wiener measure \( W(\gamma'_1) \) for the path \( \gamma'_1 \) may involve other FFIs than \( H_{n_1} \).

Still further, it is straightforward to write the fermionic Gibbs kernel in the post-symmetrized from as \( \rho(Pq_{n_1}, \tau_{n_1}; q_{n_0}, \tau_{n_0}) \) and apply exactly the same analyses on a \( G_* \)-orbit of **pre-tethered Feynman spindles** associated with

\[
\{ \Gamma(\pi q_{n_1}, \tau_{n_1}; q, \tau) \ast \gamma_0 \overset{\text{def}}{=} \{ \gamma_1 \ast \gamma_0 : \gamma_1 \in \Gamma(\pi q_{n_1}, \tau_{n_1}; q, \tau) \} \}_{\pi \in G_*},
\]

where \( \gamma_0 \) is a segment of Feynman path connecting \( (q_{n_0}, \tau_{n_0}) \) to \( (q, \tau) \), is \( \tau_{n_1} \) is a forward nodal tube for the Gibbs kernel \( \rho(Pq_{n_1}, \tau_{n_1}; q', \tau') \overset{\text{def}}{=} \langle Pq_{n_1} | e^{-\gamma_{n_1} - \gamma'} K \rangle_{H_{n_1}} \rho \rangle \) regarded as a function of \( (\rho', \tau') \in \mathcal{C} \times (\tau_{n_0} - \delta_0, \tau_{n_1}) \), \( \delta_0 > 0 \) with respect to the fixed \( q_{n_1} \in \mathcal{C}_{n_1} \) and \( \tau_{n_1} \), and where \( T_{n_1}^*(q_{n_1}, \tau_{n_1}) \) is the connected component of the open set \( \{ (q', \tau') : \rho(Pq_{n_1}, \tau_{n_1}; q', \tau') > 0 \} \) that contains the path \( \{ (q_{n_1}, \tau') : \tau' \in (\tau_{n_0}, \tau_{n_1}) \} \). It is clear that any Feynman path crossing or touching the boundary of a forward nodal tube can be safely removed from path integration, and similarly, the purge of nodal surface-touching Feynman paths can be extended to when \( \gamma_0 \) starts at a time earlier than \( \tau_{n_0} \).

Consider a practical Markov chain Monte Carlo (MCMC) procedure that samples a dual-symmetrized Gibbs kernel \( \rho(Pq, \tau; Pq_0, \tau_0) \) for the whole Feynman stack associated with the Gibbs operators \( \{ G_n \}_{n \in \mathbb{N}} \in \mathbb{N} \), where \( (q_{n_0}, q_0) \in (\mathcal{C} / G_* )^2 \), \( \mathcal{C} / G_* \) can be any subset that tiles up \( \mathcal{C} \) under the action of \( G_* \). It follows from the above analyses and method of RPI that

\[
\rho(Pq_{n_1}, \tau_{n_1}; Pq_0, \tau_0) = \int_{q_{n_1} \in \mathcal{C}_{n_1}} \int_{q_1 \in \mathcal{C}_1} \rho(Pq_{n_1}, \tau_{n_1}; q_{n_1} - 1, \tau_{n_1}) \rho(q_{n_1} - 1, \tau_{n_1}; q_1, \tau_1) \rho(q_1, \tau_1; Pq_0, \tau_0) dq_1 dq_{n_1} - 1
\]

\[
\times \int_{q_{n_1} \in \mathcal{N}_{n_1}(q_{n_1}, \tau_{n_1})} \int_{q_1 \in \mathcal{N}_{n_1}(q_0, \tau_0)} \sum_{(\pi_{n_1} - 1, \pi_1) \in G_{n_1}^2} \rho(Pq_{n_1}, \tau_{n_1}; \pi_{n_1} - 1 q_{n_1} - 1, \tau_{n_1} - 1) \rho(\pi_{n_1} - 1 q_{n_1} - 1, \tau_{n_1} - 1; \pi_1 q_1, \tau_1; Pq_0, \tau_0) dq_1 dq_{n_1} - 1
\]

\[
\times \int_{q_{n_1} \in \mathcal{N}_{n_1}(q_{n_1}, \tau_{n_1})} \int_{q_1 \in \mathcal{N}_{n_1}(q_0, \tau_0)} \rho(Pq_{n_1}, \tau_{n_1}; \pi_{n_1} - 1 q_{n_1} - 1, \tau_{n_1} - 1) \rho(\pi_{n_1} - 1 q_{n_1} - 1, \tau_{n_1} - 1; \pi_1 q_1, \tau_1; Pq_0, \tau_0) dq_1 dq_{n_1} - 1
\]

\[
\times \int_{q_{n_1} \in \mathcal{N}_{n_1}(q_{n_1}, \tau_{n_1})} \int_{q_1 \in \mathcal{N}_{n_1}(q_0, \tau_0)} \sum_{\pi_{n_1} - 1 \in A_{n_1}} \sum_{\tau_{n_1} - 1 \in \tau_{n_1} - 1} \rho_\tau(\pi_{n_1} - 1 q_{n_1} - 1, \tau_{n_1} - 1; \pi_1 q_1, \tau_1; Pq_0, \tau_0) dq_1 dq_{n_1} - 1 \forall (q_{n_1}, q_0) \in (\mathcal{C} / G_* )^2,
\]

where \( \mathcal{N}_1(q_0, \tau_0) \) is the nodal cell of \( \rho(\cdot, \tau_1; Pq_0, \tau_0) \) at \( \tau_1 \) enclosing \( q_0 \), and \( \mathcal{N}_n(q_n, \tau_n) \overset{\text{def}}{=} T_{n_1}(q_{n_1}, \tau_{n_1}) \cap (\mathcal{C} \times \{ \tau_{n_1} - 1 \}) \) is the backward nodal cell of \( \rho(Pq_{n_1}, \tau_{n_1}; q_{n_1} - 1, \tau_{n_1} - 1) \) at \( \tau_{n_1} - 1 \) enclosing \( q_{n_1} \), \( \mathcal{N}_{n_1}(q_{n_1}, \tau_{n_1}) \) and \( \mathcal{N}_{n_1}^*(q_{n_1}, \tau_{n_1}) \) are the corresponding set unions of nodal cells and backward nodal cells respectively, while \( \rho_\tau(\pi_1 q_1, \tau_1; Pq_0, \tau_0) \)
and \(\rho_\pi(q_N, \tau_N; q_{N-1}, \tau_{N-1})\) for any \(\pi \in A_r\) are non-negative definite transition amplitudes computed using RPIs within the first and the last Feynman slabs respectively. In other words, the first and the last Feynman slabs have been path rectified, with the fermionic exchange and sign considerations no longer necessary and relevant. The same method of RPI can be repeated to have all Feynman slabs path rectified, such that

\[
\rho(\mathcal{P} q_N, \tau_N; \mathcal{P} q_0, \tau_0) \propto \left\{ \int_{q_n \in N_n} \sum_{\tau_n \in A} \right\}^{N-1}_{n=1} \rho_\pi(\pi_{N-1} q_N, \tau_N; \pi_{N-1} q_{N-1}, \tau_{N-1})
\]

\[
\times \prod_{n=1}^{N-1} \rho_\pi(\pi_n q_n, \tau_n; \pi_n q_{n-1}, \tau_{n-1}) \prod_{n=1}^{N-1} dq_n,
\]

\(\forall(q_N, q_0) \in (C/G_n)^2\), where \(N_n \; d e f \; N_n(q_{n-1}, \tau_{n-1})\) for each \(n \in [1, \lfloor N/2 \rfloor], \lfloor N/2 \rfloor \; d e f \; \max\{n \in \mathbb{N} : 2n \leq N\}\) represents the nodal cell of the pre-symmetrized Gibbs kernel of the \(n\)-th Feynman slab with respect to \((q_{n-1}, \tau_{n-1})\), and \(N_n \; d e f \; N^*_n(q_{n+1}, \tau_{n+1})\) for each \(n \in [1+\lfloor N/2 \rfloor, N-1]\) denotes the backward nodal cell of the post-symmetrized Gibbs kernel of the \((n+1)\)-th Feynman slab with respect to \((q_{n+1}, \tau_{n+1})\). Due to the frequent insertion of \(\pi \in A_r\) permutations, the Feynman paths in a path integral such as equation (8) appear to undergo abrupt coordinate jumps in the space \(C \times (\tau_0, \tau_N)\), which do not represent actual physical discontinuities, since all points in any orbit \(A_q, q \in C\) are physically equivalent and represent the same physical reality. Indeed, such Feynman paths are actually continuous in the space \((C/A_r) \times (\tau_0, \tau_N)\), where \(C/A_r\) is an orbifold regarding each orbit \(A_q, q \in C\) as a single point. A practical and effective means to incorporate such exchange equivalence is to associate each point \((q_n, \tau_n)\), \(n \in [0, N]\) on a Feynman plane with two equivalent coordinates \(q_n^- \in A_q q_n\) and \(q_n^+ \in A_q q_n\), where \(q_n^+\) serves the Feynman stack or slice from \(\tau_{n-1}\) to \(\tau_n\), and \(q_n^-\) is used by the Feynman stack or slice from \(\tau_n\) to \(\tau_{n+1}\).

It is worth noting that all path rectifications are done on a per Feynman slab basis, requires a solution for the nodal surface of the specific FFI associated with each Feynman slab, which can be obtained at no more than a constant computational cost due to the bounded number of fermion species moved by each FFI, while each species has a bounded number of identical particles. It is also important to note that the order in which the Feynman slabs are path rectified in equations (7) and (8) apparently creates a sequence of conditional dependence among the nodal cells on the Feynman planes and integrations over the corresponding configuration coordinates. However, such conditional dependence is insubstantial, since the order of the Feynman slabs being path rectified can be arbitrary, and different choices of such an order yields the same result for \(\rho(\mathcal{P} q_N, \tau_N; \mathcal{P} q_0, \tau_0)\). Specifically, for any \(n \in [1, N-1]\), if desired, the method of RPI can be applied iteratively such that the first to the \((n-1)\)-th Feynman slabs are firstly path rectified in sequence, then the \((N-1)\)-th to the \((n+1)\)-th Feynman slabs are path rectified backward in order, leaving \(\rho(\mathcal{P} q_{n-1}, \tau_{n-1}; \mathcal{P} q_{n+1}, \tau_{n+1})\) as the only fermionic and signed factor in the LTK product of Gibbs kernels, which is finally path rectified by discarding any Feynman path from a given \(q_n^+ \in A_q q_n\) at \(\tau_{n-1}\) to a given \(q_{n+1}^- \in A_q q_{n+1}\) at \(\tau_{n+1}\) that ever crosses or touches a nodal surface of either \(\rho(q, \tau; \mathcal{P} q_{n-1}, \tau_{n-1})\), \(\tau \in (\tau_{n-1}, \tau_n)\) or \(\rho(\mathcal{P} q_{n+1}, \tau_{n+1}; q, \tau)\), \(\tau \in (\tau_n, \tau_{n+1})\) as a function of \(q \in C\), \(\forall \tau \in (\tau_{n-1}, \tau_{n+1})\). In particular, the fully rectified path integral integrates over the points \(q_n\) such that, both \(q_n^- \in A_q q_n\) is in the Ceperley reach of \(q_n^-\) with respect to \(\langle e^{-\delta \mathcal{H}_n | K} |\mathcal{P} q_{n-1}\rangle\), and \(q_n^+ \in A_q q_n\) is in the Ceperley reach of \(q_{n+1}^-\) backward with respect to \(\langle \mathcal{P} q_{n+1}^- | e^{-\delta \mathcal{H}_n | K} |\rangle\). This perspective justifies a Monte Carlo subroutine that wiggles a Feynman path by walking only one coordinate on a single Feynman plane at any given time, subject to the constraint of no nodal crossing or touching in the two bordering Feynman slabs or slices.

In one exemplary embodiment of numerical simulation, a random and warm start \(q_0 \in C\) is chosen, and the RPI of equation (8) is approximated by running an MCMC procedure that starts with an initial Feynman path having \(q_n = q_0\), \(\forall n \in [0, N]\), and repeats two nested loops for a polynomial number of times, where an outer loop walks each coordinate \(q_{iK}, i \in [0, m]\) in turn using a Metropolis-Hastings algorithm [21,22] or Gibbs sampling [23] based on a conditional probability \(\text{Pr}(q_{iK} | q_{i+1K}, q_{i-1K})\), with the index \(i = 0\) or \(i+1\) being void and the coordinate \(q_{i-1K}\) or \(q_{i+1K}\) being omitted when \(i = 0\) or \(i = m\), while an inner loop samples from all of the restricted Feynman paths running from a fixed \(q_{i-1K}^+ \in N_{i-1K}^+\), passing a proposed \(q_{iK} \sim q_{iK}^\ast \sim q_{iK} \in N_{iK}^+\), and finally reaching a fixed \(q_{i+1K} \in N_{i+1K}^+\), \(\forall i \in [0, m]\), so to compute an RPI approximating the probability

\[
\text{Pr}(q_{iK} | q_{i+1K}, q_{i-1K}) \propto \rho_\pi(q_{i+1K}, \tau_{i+1K}; q_{iK}, \tau_{iK}) \rho_\pi(q_{iK}, \tau_{iK}; q_{i-1K}, \tau_{i-1K}),
\]

which coincides with the dual-symmetrized Gibbs kernel \(\langle \mathcal{P} q_{i+1K} | e^{-\delta \mathcal{H}_n | K} | q_{iK} \rangle \langle q_{iK} | e^{-\delta \mathcal{H}_n | K} | \mathcal{P} q_{i-1K} \rangle\) as a func-
tion of $q_{iK} \in C_{iK}/A_*$ restricted to a positive nodal cell. At each instant of time during either the inner or the outer loop, a random permutation $\pi \in A_*$ and two random integers $(n_1, n_2) \in [0, N]^2$, $n_1 < n_2$ can be chosen, such that a segment $\gamma(n_1:n_2)$ of the instantaneous Feynman path that starts at $\tau_{n_1}$ and ends at $\tau_{n_2}$ can be transformed into $\pi \gamma(n_1:n_2)$, where in particular, the sequence of points $\{\{\pi q_n, \tau_n\} : n \in [n_1, n_2]\}$ as an element of the so-called cylinder set is transformed into $\{\{\pi q_n, \tau_n\} : n \in [n_1, n_2]\}$. The two nested loops create and operate at two timescales, where the inner loop is related to a short timescale, within which a random walk among the restricted Feynman paths mixes rapidly to yield a good estimate for $\Pr(q_{iK} | q_{iK+1}^+, q_{iK}^+)$, while the outer loop is associated with a long timescale, in which the coordinates $\{q_{iK} \} \in [0,m]$ walk randomly and mix at a slower but still polynomial rate. The slow dynamics of the outer loop always sees the fast dynamics of the inner loop in its equilibrium at any instant on the long timescale.

Specifically, for an SFF Hamiltonian that is LTK-decomposed, given a fixed $\tau > 0$, an $m \in \mathbb{N}$ can be chosen sufficiently large such that $\delta \tau = \tau/m$ is sufficiently small. The inner loop is essentially a multi-dimensional integration over the points $\{q_n : n \in [(i-1)K, (i+1)K]\} \in (C/A_*)^{2K+1}$ weighted by the Wiener measure $\prod_{n=\ldots}^{(i+1)K-1} [q_{n+1} | e^{-\delta \tau H_{(n+1)/K}}] [q_n]$, while subject to the constraint that $\langle \mathcal{P}_q_{n+1} | e^{-\delta \tau H_{(n+1)/K}} | \mathcal{P}_q_n \rangle$ does not change sign or vanish, $\forall n \in [(i-1)K, (i+1)K-1]$. The Wiener measure is substantially a product between a Gaussian measure representing Brownian motions of free particles and a multiplicative factor valued within the interval $[e^{-2\delta \tau K V_0}, e^{2\delta \tau K V_0}]$, the latter being due to the potential energy terms in the FFIs that are bounded within $[-V_0, V_0]$. A small $\delta \tau$ dictates that the coordinates $\{q_n : n \in [(i-1)K, (i+1)K]\}$ can not be far separated in the $C/A_i$ orbifold topology lest the Wiener measure be exponentially small, and each nodal surface can be well approximated locally as a hyperplane bisecting an Euclidean configuration space, for any configuration point $q_n, n \in [(i-1)K, (i+1)K]$ that ever gets close to such a nodal surface. Therefore, the multi-dimensional integration performed by the inner loop is mostly like an MCMC sampling from and integrating a logconcave distribution over a convex polyhedron defined by the hyperplanes locally approximating the nodal surfaces. The latter is proved to mix rapidly \cite{24} with a polynomial-sized Cheeger constant or conductance $CheegConst$ \cite{25,27}, so is the former rapidly-mixing with a Cheeger constant no smaller than $e^{-2\delta \tau K V_0} \times CheegConst$.

For each step of the outer loop a coordinate $q_{iK} \in C_{iK}/A_*$, $i \in [0, m]$ to a new position $r_{iK} \in C_{iK}/A_*$, with all of the other coordinates $\{q_{jK} : j \in [0, m], j \neq i\}$ being fixed, the Metropolis-Hastings algorithm or Gibbs sampling accepts or rejects $q_{iK} \rightarrow r_{iK}$ move with a probability that depends on a ratio between two conditional probabilities $\Pr(r_{iK} | q_{iK+1}^+, q_{iK}^+)/(r_{iK} | q_{iK+1}^-)$, subject to a constraint that both $q_{iK}$ and $r_{iK}$ can be reached by restricted Feynman paths from both $q_{iK+1}^+$ and $q_{iK-1}^-$. This is essentially a random walk over a nodal cell of $Pr(q_{iK} | q_{iK+1}^-)$ as a function of $q_{iK} \in C_{iK}/A_*$. The outer loop overall is to drive the distribution of coordinates from an initial $\phi(\{q_{iK} : i \in [0, m]\}) \in L^1(C/A_i^{m+1})$ to a stationary distribution $\phi_0(\{q_{iK} \} : i \in [0, m])$ at equilibrium. To see how fast the process converges, expand the multi-variable function $\phi(\cdot)$ in terms of products among the eigenfunctions $\{|\psi_n(\Omega)\rangle : n \geq 0\}$ listed in the order of increasing energies, such that

$$\phi(\{q_{iK} : i \in [0, m]\}) = \sum_{n_0n_1\ldots n_m} \epsilon_{n_0n_1\ldots n_m} c_{n_0n_1\ldots n_m} \psi_{n_0}(q_0) \psi_{n_1}(q_1) \cdots \psi_{n_m}(q_m).$$

Assume that the ground state $\psi_0(\Omega)$ is non-degenerate. Then it is obvious that the equilibrium distribution is precisely $\phi_0(\{q_{iK}\}) = \psi_0(q_0) \psi_0(q_1) \cdots \psi_0(q_m)$. Any monomial $\psi_{n_0}(q_0) \cdots \psi_{n_i}(q_i) \cdots \psi_{n_m}(q_m)$ in equation $\psi(\{q_{iK}\})$ that has an excited state $\psi_{n_i}, n_i \neq 0, i \in [0, m]$ as a factor will have its coefficient $c_{n_0n_1\ldots n_m}$ strictly reduced by a factor $e^{-\alpha(i)\tau \lambda_{n_0}(H) - \lambda_0(H)}$ after each step of the outer loop walking the coordinate $q_{iK}$, with $\alpha(0) = \alpha(m) = 1$ and $\alpha(i) = 2$ for all $i \in [1, m-1]$. Therefore, the mixing time $\tau\tau$ of the outer loop is upper-bounded by $O(m^2\Lambda \Delta \lambda_0(\Omega))$, with $\Lambda \Delta \equiv \lambda_0(\Omega)-\lambda_0(\Omega)$ being the spectral gap of $\Omega$.

**Theorem 1.** Let $H = \sum_{K=1}^{K-1} H_k$ be an SFF Hamiltonian that is LTK-decomposed, whose ground state is non-degenerate and separated from all excited states by an $\Omega(1/\poly(size(\Omega)))$ energy gap. Then the Gibbs kernel $\langle \mathcal{P}_r | e^{-\tau H} | \mathcal{P}_q \rangle, (r,q) \in C^2$ is BPP simulatable for all $\tau \in [0, \infty)$, namely, there is a Monte Carlo algorithm that runs for a time no longer than $\poly(size(\Omega), \epsilon^{-1})$ to generate a random sample of $(r, q) \in C^2$, whose probability distribution is no more than $\epsilon$ away from $\langle \mathcal{P}_r | e^{-\tau H} | \mathcal{P}_q \rangle$ in the total variance distance $\tau\tau$, for all $(r, q) \in C^2$, for any given $\epsilon > 0$.

**Proof.** By the exchange symmetry, it is sufficient to simulate the Gibbs kernel $\langle r | e^{-\tau H} | q \rangle$ for $q \in C$ and for $r \in N(q; \tau H)$, with $N(q; \tau H)$ denoting the nodal cell of $\{|e^{-\tau H} | q \rangle\}$ with respect to and containing a start
point \(q \in \mathcal{C}\). Due to the polynomial energy gap, it is WLOG to assume that \(\tau = O(\text{poly(size}(H)))\) bounded, because \(e^{-\tau[H-H_0(H)\lambda]}\) becomes essentially a projection to the ground state when \(\tau\) is sufficiently large. Employ an MCMC procedure as specified in the above exemplary embodiment, with an \(m\) chosen sufficiently large but still \(O(\text{poly(size}(H)))\) bounded, so that \(\delta \tau = \tau/m\) is sufficiently small and the inner loop converges well within \(O(\text{poly(size}(H)))\) iterations. Again due to the polynomial energy gap and the uniqueness of \(\psi_0(H)\), the outer loop needs to iterate only \(O(\text{size}(H), \epsilon^{-1})\) times to produce a random sample \((r, q) \in \mathcal{C}^2\) that is distributed \(\epsilon\)-close to \(\langle r | e^{-\tau H} | p q \rangle\). The overall runtime is clearly \(O(\text{poly(size}(H), \epsilon^{-1}))\). □

Running such an MCMC procedure for a polynomial number of times generates a polynomial number of random samples of \((r, q) \in \mathcal{C}^2\) according to the Gibbs kernel \(\langle r | e^{-\tau H} | p q \rangle\), \(q \in \mathcal{C}, r \in \mathcal{N}(q; r \mathcal{H})\). In particular, by enforcing a periodic boundary condition \(r = q_n = q_0 = q\) and running the MCMC for a polynomial number of times, enough samples can be generated to estimate the partition function \(Z = \int \langle q | e^{-\tau H} | p q \rangle dq\) to within a polynomial accuracy. In general, for any observable \(Q\) whose matrix elements \(\langle q | Q | r \rangle\), \((q, r) \in \mathcal{C}\) can be efficiently computed, a Feynman slab implementing \(\langle q | Q | r \rangle\) can be added to extend a Feynman stack that computes \(\langle r | e^{-\tau H} | p q \rangle\), \((q, r) \in \mathcal{C}\), then the extended Feynman stack can be MCMC simulated, again with the periodic boundary condition enforced, to estimate the expectation value \(Z^{-1} \int \langle q | Q | r \rangle \langle r | e^{-\tau H} | p q \rangle dq\) to within a polynomial accuracy at the cost of a polynomial runtime. For many applications, there is a contextual prescription to choose a guaranteed warm start \(q \in \mathcal{C}\). Alternatively, almost all configuration points are a warm start so long as the Hamiltonian \(H\) is well-behaved analytically, e.g., when \(H = -\Delta + V\) is of the Schrödinger type with a bounded potential \(V\).

For an SFF Hamiltonian \(H = \sum_{k=1}^{K} H_k\) that is GSP-decomposed, it is convenient to adopt the perspective and method of reptation QMC [29][31], where a sequence of Gibbs operators \(\{G_k\}_{k \in [1, K]}\) is applied to a known initial quantum state \(|\Psi\rangle\) in turn and repeated for \(m \in \mathbb{N}\), \(m = O(\text{poly(size}(H)))\) times, with \(G_k \equiv e^{-\delta \tau[H_k - \lambda_0(H_k)]}, \forall k \in [1, K]\), \(\delta \tau\) being no longer small but sufficiently large such that \(G_k\) is essentially the same as \(P_k = \lim_{\tau \to \infty} e^{-\delta \tau[H_k - \lambda_0(H_k)]}\) up to an error that is exponentially small, \(\forall k \in [1, K]\), and \(m\) sufficiently large such that the projected state \(|\Psi\rangle \approx (\prod_{k=1}^{K} G_k)^m|\Psi\rangle\) is essentially the same as \(\psi_0(H)\) up to an error this is \(O(1/\text{poly}(m))\). For any measurement operator \(Q\) of interest, the expectation value \(\langle \Psi | Q | \Psi \rangle = \langle \Phi | (\prod_{k=1}^{K} G_{K-k+1})^m Q (\prod_{k=1}^{K} G_k)^m | \Phi \rangle\) can be estimated by RPI with a Feynman stack consisting of \(2mK\) Feynman slabs, each of which corresponds to a Gibbs operator \(G_k, k \in [1, K]\) and is path rectified in accordance with the nodal structure of the Gibbs kernel \(\langle |G_k| \rangle\), which is efficiently computable and readily available in relation to the associated FFI \(\mathcal{H}_k, \forall k \in [1, K]\). Either \(Q\) is coordinate-diagonal, so that \(\langle \Psi | Q | \Psi \rangle\) can be computed straightforwardly from a polynomial number of Feynman path samples, or else \(Q\) involves a small number of fermion species, and the total Hamiltonian is modified into \(H + \eta Q, \eta \in \mathbb{R}\), which is still SFF, so that an RPI-based reptation QMC is performed for the Hamiltonian \(H + \eta Q\), from which \(\langle \Psi | Q | \Psi \rangle\) is derived by invoking the Feynman-Hellman theorem [29][31]. In order to ensure a warm start, it is helpful to employ a sequence of SFF Hamiltonians \(H(t) = \sum_{k=1}^{K} H_k(t), t \in [0, 1]\), which evolves adiabatically from an easily solvable \(H(0)\) having \(|\Phi\rangle\) as its unique ground state to \(H(1) = H\), the target Hamiltonian, such that the Feynman stack associated with the nested sequence of FFIs \(\{\{\mathcal{H}_k(t/m)\}\}_{k \in [1, K]}\) projects \(|\Phi\rangle\) at the start into \(|\Psi\rangle = \psi_0(H(1))\) in the middle [8].

The above method of reptation QMC uses path integral to compute an inner product \(\langle \Psi | Q | \Psi \rangle\) for the wavefunction \(|\Psi\rangle\) using the method of convolution \(\sum_{n=1}^{mK} G_n|\Phi\rangle\), with \(G_n \equiv G_{n|K}, \forall n \in [1, mK]\). Alternatively, a time-inhomogeneous MCMC can be constructed to compute or sample from the outer (dyadic) product \(\langle \Psi | \Psi \rangle\), that is a density matrix or operator, also called a Gibbs wavefunction of the second order. Such time-inhomogeneous MCMC repeats an iterative loop comprising \(mK+1\) steps, wherein the zeroth step picks a random sample \(|r_0\rangle\langle q_0|\) from a positive nodal cell of \(|\Phi\rangle\langle \Phi|\) according to the density function \(\Phi(r_0)\Phi(q_0)\), and stores a triple \((r_0, q_0, p_0)\) with \(p_0 \equiv \Phi(r_0)\Phi(q_0)\) as the initial value of the time variable for the time-inhomogeneous Markov chain: Next, for each \(n \in [1, mK]\), the \(n\)-th step fixes \((r_{n-1}, q_{n-1})\) and samples a random dyadic \(|r_n\rangle\langle q_n|\) from a positive nodal cell of \(G_n|r_{n-1}\rangle\langle q_{n-1}|G_n^\dagger\) according to the density function \(\langle r_n| G_n|r_{n-1}\rangle\langle q_{n-1}|G_n^\dagger|q_n\rangle\), then updates the value of the state variable from \((r_{n-1}, q_{n-1}, p_{n-1})\) to \((r_n, q_n, p_n)\), with \(p_n \equiv \langle r_n| G_n|r_{n-1}\rangle\langle q_{n-1}|G_n^\dagger|q_n\rangle\); In each said \(n\)-th step, \(n \in [1, mK]\), the density function is efficiently computable and easily path rectified since each \(G_n, n \in [1, mK]\) is associated with an FFI; Finally, the \(mK\)-th step produces a value \((r_{mK}, q_{mK}, p_{mK})\) of the state variable, which represents a random sample \(|r_{mK}\rangle\langle q_{mK}|\) from \(|\Psi\rangle\langle \Psi|\) with a probability weight \(p_{mK}\). Repeating said iterative loop for a polynomial number of times generates an ensemble of samples to
represent $|\Psi\rangle\langle\Psi|$, from which a physical quantity of interest can be derived.

Not only SFF Hamiltonians can be simulated efficiently, but also they are universal for many-body physics and quantum computing. The following will show that any BQP algorithm given as an ordered sequence of controlled R-gates \[ \{U_t \equiv I \otimes Z_t^+ + R_t \otimes Z_t^- \}_{t \in [1, T]}, T \in \mathbb{N} \] on a quantum computer of \( n \in \mathbb{N} \) qubits can be mapped to an LTK- or GSP-decomposed SFF Hamiltonian, where \( R_t \equiv X \sin(\theta(t)) + Z \cos(\theta(t)) \), \( \theta(t) \in [-\pi, \pi) \) is a self-inverse gate applied to a qubit indexed by an \( i(t) \in [1, n] \), and \( Z_t^\pm \equiv (\pm Z)/2 \) operators applied to a control qubit indexed by a \( j(t) \in [1, |T|] \), \( \forall t \in [1, T] \). \( Z, Z_t^\pm \) are the usual Pauli matrices acting on a single qubit as the simplest quantum system (\( \mathcal{C}_0, \mathcal{H}_0, \mathcal{B}_0 \)), with \( \mathcal{C}_0 \equiv \{(0,0), (1,1) : \alpha, \beta \in \mathbb{R} \} \), \( \mathcal{B}_0 \) being the Banach algebra of \( 2 \times 2 \) real matrices. Such a BQP algorithm, or its associated quantum circuit, is said to have a computational size \( T + n \).

**Definition 3.** A homomorphism \( \mathfrak{M} : (\mathcal{C}, \mathcal{H}, \mathcal{B}) \to (\mathcal{C}', \mathcal{H}', \mathcal{B}') \) between two quantum systems with Hamiltonians \( H \in \mathcal{B} \) and \( H' \in \mathcal{B}' \) is an injective mapping that sends any subset \( D \subseteq C \) to a unique \( D' \equiv \mathfrak{M}(D) \subseteq C' \), maps any \( \psi \in \mathcal{H} \) to a unique \( \psi' \equiv \mathfrak{M}(\psi) \in \mathcal{H}' \), and sends any \( Q \in \mathcal{B} \) to a unique \( Q' \equiv \mathfrak{M}(Q) \in \mathcal{B}' \), such that \( C \supseteq D \implies \mathfrak{M}(D) \subseteq C' \) embeds the Boolean algebra of subsets \( \mathfrak{B}(C) \subset \mathfrak{B}(C) \) into the Boolean algebra of subsets \( \mathfrak{B}(C') \subset \mathfrak{B}(C') \); \( 2) \mathcal{H} \supseteq \psi \implies \mathfrak{M}(\psi) \in \mathcal{H}' \) embeds the Hilbert space \( \mathcal{H} \) into \( \mathcal{H}' \); \( 3) \mathcal{B} \ni Q \implies \mathfrak{M}(Q) \in \mathcal{B}' \) embeds the Banach algebra \( B \) into \( B' \); \( 4) \) there exists a constant \( c > 0 \), \( c^{-1} = O(\text{poly(size}(\mathcal{H}))) \), with which \( \langle \mathfrak{M}(\psi)|\mathfrak{M}(Q)|\mathfrak{M}(\phi) \rangle = c\langle \psi|Q|\phi \rangle \text{ holds } \forall \psi, \phi \in \mathcal{H}, \forall Q \in \mathcal{B} \); \( 5) \text{size}(\mathcal{H}) = O(\text{poly(size}(\mathcal{H}'))) \) and \( \text{size}(\mathcal{H}') = O(\text{poly(size}(\mathcal{H}'))) \). An homomorphism \( \mathfrak{M} \) is called an isomorphism when the mapping \( \mathfrak{M} \) is also surjective.

Firstly, it is useful to construct a bi-fermion system \( (\mathcal{C}_1, \mathcal{H}_1, \mathcal{B}_1) \) consisting of two non-interacting identical fermions moving on a circle \( T \equiv \mathbb{R}/2\mathbb{Z} \), governed by a single-particle Hamiltonian \( -(1/2)\partial^2/\partial x^2 + V(x) \), \( x \in \mathbb{T} \), with an external potential \( V(x) = V_0 \left[ d(x, 0) - a_0 \right] \), \( V_0 > 0 \), \( a_0 > 0 \) being a constant, where \( d(x, y) \) denotes the geodesic distance between \( x, y \in \mathbb{T} \). If the potential is sufficiently large, the potential well and barrier become essentially Dirac deltas, \( V(x) \equiv \gamma_0 \delta(x + 1) - \gamma_0 \delta(x), \) \( x \in [-1, 1] \) (mod 2), such that a bi-fermion under a nominal Hamiltonian \( \mathfrak{H}_0 = (\gamma_0^2 - \pi^2)/2 + \sum_{i=1}^2 [-(1/2)\partial^2/\partial x_i^2 + V(x_i)] \), \( (x_1, x_2) \in \mathbb{T}^2 \) behaves like a rebit with two low-energy states

\[
\begin{align*}
\psi_+(x_1, x_2) &= (1 - \pi_{12}) \sin \pi d(x_1, x_2) \left[ d(x_1, 0) - a_0 \right] e^{-\gamma_0 d(x_2, 0)}, \\
\psi_-(x_1, x_2) &= (1 - \pi_{12}) \sin \pi x_1 e^{-\gamma_0 d(x_2, 0)},
\end{align*}
\]

(11) (12)

that are degenerate at \( E_0 = 0 \), where \( \pi_{12} \) is the fermion exchange operator swapping the particle labels 1 and 2. Choose \( \alpha_0 = 2\gamma_0^{-1} \log \gamma_0 \), then for all \( x \) such that \( d(x, 0) > \alpha_0 \), the amplitude of the single-particle ground state \( e^{-\gamma_0 d(x, 0)} < \gamma_0^{-2} \), which is rather small. Construct potential functions

\[
\begin{align*}
X(x_1, x_2) &= \gamma_0 (1 + \pi_{12}) \left[ d(x_1, 0) - a_0 \right] \left[ d(x_2, 0) - a_0 \right] - (\pi^2/4\gamma_0^2), \\
Z^+(x_1, x_2) &= (1 + \pi_{12}) \left[ d(x_1, +1/2) - a_0 \right] \left[ d(x_1, 0) - a_0 \right], \\
Z^-(x_1, x_2) &= (1 + \pi_{12}) \left[ d(x_1, -1/2) - a_0 \right] \left[ d(x_1, 0) - a_0 \right],
\end{align*}
\]

(13) (14) (15)

\( \forall (x_1, x_2) \in [-1, 1]^2 \). It is clear that a bi-fermion implements a rebit via a homomorphism \( \mathcal{M}_1 : (\mathcal{C}_0, \mathcal{H}_0, \mathcal{B}_0) \to (\mathcal{C}_1, \mathcal{H}_1, \mathcal{B}_1) \) such that, with \( |\pm\rangle \equiv ((0 \pm 1))/\sqrt{2} \),

\[
\begin{align*}
\mathcal{M}_1(|\pm\rangle &\in \mathcal{H}_0) = \psi_{\pm}(x_1, x_2) \in \mathcal{H}_1, \\
\mathcal{M}_1(X \in \mathcal{B}_0) = (2\gamma_0^2/\pi^2) [\mathfrak{H}_0 + X(x_1, x_2)] \in \mathcal{B}_1, \\
\mathcal{M}_1(Z^\pm \in \mathcal{B}_0) = \gamma_0 \mathfrak{H}_0 + Z^\pm(x_1, x_2) \in \mathcal{B}_1.
\end{align*}
\]

(16) (17) (18)

Via linear combinations, the operators \( \mathcal{M}_1(X), \mathcal{M}_1(Z^+), \mathcal{M}_1(Z^-) \) generate all partial Hamiltonians that are of interest for quantum computing on a single bi-fermion, because \( \text{span}\{X, Z^+, Z^-\} \) contains all Hermitian elements in \( \mathcal{B}_0 \). It is noted in passing that, although it is preferred for the single-particle potential \( V(x) \), \( x \in \mathbb{T} \) to have a narrow and deep potential well around \( x = 0 \), approximating a fairly strong Dirac delta
to localize one of the two fermions in a small neighborhood of \(x = 0\), there is no practical necessity other than convenience of mathematical analysis, to require a steep potential barrier around \(x = \pm 1\). Rather, it is perfectly fine to place a relatively wide and low potential barrier, as long as its width and height are chosen properly to be commensurate with the Delta-like potential well around \(x = 0\), such that the nominal bi-fermion Hamiltonian \(H_{BF}\) defines a degenerate two-state Hilbert space implementing a reb bit.

Next, it is also straightforward to construct a homophysics \(\mathcal{M}_2 : (C_1^2, \mathcal{H}_1^2, B_1^2) \rightarrow (C_2^2, \mathcal{H}_2^2, B_2^2)\), with \(C_2^a \equiv C_i \times \mathcal{H}_2^i \equiv \mathcal{H}_i \times \mathcal{H}_2, B_2^a \equiv B_i \otimes B_2, \forall i \in \{0, 1\}\), so to implement a pair of interacting rebits using two bi-fermions conditioned and interacting through the following partial Hamiltonians,

\[
\mathcal{M}_2 (X_1 \otimes Z_{21}^\pm) = \frac{2\gamma_0^2}{\pi^2} [H_{BF,1} + H_{BF,2} + X(x_{11}, x_{12}) Z^\pm(x_{21}, x_{22})],
\]

\[
\mathcal{M}_2 (Z_1^\pm \otimes Z_{22}^\pm) = \gamma_0 H_{BF,1} + \gamma_0 H_{BF,2} + Z^\pm(x_{11}, x_{12}) Z^\pm(x_{21}, x_{22}),
\]

where \(\forall i \in \{1, 2\}\), \(X_i, Z_i^\pm, Z_i = Z_i^+ - Z_i^-\) are the X- and Z-gates on the i-th reb bit, \(H_{BF,i}\) is the nominal Hamiltonian of the i-th bi-fermion, \((x_{i1}, x_{i2}) \in T^2\) is the two-fermion configuration of the i-th bi-fermion. \(X_1 \otimes Z_{21}^\pm\) and \(Z_1 \otimes Z_{22}^\pm\) are called "single-rebit-controlled gates", whose linear combinations include all single-rebit-controlled \(R\) gates, which are already universal for ground state quantum computing (GSQC) in the sense that, using the so-called perturbative gadgets, up to an error tolerance \(\epsilon > 0\), the low-energy physics of any system of \(n \in \mathbb{N}\) rebits under a computationally k-local Hamiltonian, \(k \in \mathbb{N}\) being a fixed number, can be homophysically mapped to the low-energy physics of another system of \(poly(n, \epsilon^{-1})\) rebits under a Hamiltonian that involves only one-body and two-body interactions, especially the controlled \(R\)-gates, whose operator norms are upper-bounded by \(poly(\epsilon^{-1})\) [10][34][36]. In particular, the “XX from XZ gadget” of Biamonte and Love [10] can be employed to effect homophysically an \(X \otimes X\) interaction between a first and a second rebits through \(X \otimes Z\) interactions with a zeroth reb bit,

\[
I - X_1 \otimes X_2 \xrightarrow{\mathcal{M}} \gamma_0 (I - X_0) + (I - X_1 \otimes X_2),
\]

where \(\mathcal{M}\) reads and stands for “is homophysically mapped to”, \(\gamma_0 \gg 1\) is a large constant. Then the linear combinations of \(X \otimes X\) and \(Z \otimes X\) include all two-rebit interactions of the form \(R(\theta) \otimes X\), with \(R(\theta) \equiv X \sin \theta + Z \cos \theta, \theta \in [-\pi, \pi]\). Alternatively, there is a special class of few-rebit interactions called "multi-rebit-controlled gates" of the form \(R_i(\theta) \otimes \prod_{j \in J} Z_j\), with \(\theta \in [-\pi, \pi]\), \(i\) indexing a rebit being operated upon, \(J\) being a set indexing a fixed number of control rebits. Such a multi-rebit-controlled \(R\)-gate does not require a decomposition into two-rebit couplings, but can be implemented through a linear combination of the following homophysics,

\[
\mathcal{M} (X_1 \otimes \prod_{j \in J} Z_j^\pm) = \frac{2\gamma_0^2}{\pi^2} [H_{BF,i} + \sum_{j \in J} H_{BF,j} + X(x_{i1}, x_{i2}) \prod_{j \in J} Z^\pm(x_{j1}, x_{j2})],
\]

\[
\mathcal{M} (Z_1^\pm \otimes \prod_{j \in J} Z_j^\pm) = \gamma_0 H_{BF,i} + \gamma_0 \sum_{j \in J} H_{BF,j} + Z^\pm(x_{i1}, x_{i2}) \prod_{j \in J} Z^\pm(x_{j1}, x_{j2}).
\]

At any rate, it has been established that any computationally k-local Hamiltonian \(H\) involving an n-rebit \(H_{BF}\) inhomophysically implemented as an SFF Hamiltonian \(\mathcal{M}(H)\) involving no more than \(poly(n, \epsilon^{-1})\) bi-fermions, such that the low-energy physics of \(H\) and \(\mathcal{M}(H)\) are homophysically up to a variable error tolerance \(\epsilon > 0\), where each FFI in \(\mathcal{M}(H)\) moves no more than \(k' \in \mathbb{N}\) bi-fermions, with \(k'\) being another fixed number, and has an operator norm that is upper-bounded by \(poly(\epsilon^{-1})\), while all bi-fermions are mutually distinguishable entities.

Given a universal BQP algorithm \(\{U_t \equiv I \otimes Z_t^+ + R_t \otimes Z_t^-\}_{t \in [1, T]}\), \(T \in \mathbb{N}\), with each controlled \(R\)-gate \(U_t\) operating on an i\((t)\)-th, \(i(t) \in [1, n]\) and a j\((t)\)-th, \(j(t) \in [1, n]\) rebits in an n-rebit logic register represented by \((C, \mathcal{H}_C, B_C)\) as a quantum subsystem, each \(R_t\) being self-inverse, \(\forall t \in [1, T]\), where the successive applications of the controlled-\(R\) gates are meant to generate a series of quantum states \(|\phi_t\rangle_L \equiv U_t |\phi_{t-1}\rangle_L\), \(t \in [1, T]\), from a given initial state \(|\phi_0\rangle_L\) till a computational result \(|\phi_T\rangle_L = (\prod_{t=1}^T U_t) |\phi_0\rangle_L\) at the end, the celebrated Feynman-Kitaev construct [3][8][11][12] introduces a clock register represented by \((C_C, \mathcal{H}_C, B_C)\) as a
quantum subsystem to support clock states \( \{ |t\rangle_c \}_{t \in [0, T]} \subseteq \mathcal{H}_c \), so that the clock and logic registers constitute a GSQC system represented by \((\mathcal{C}_c \times \mathcal{L}_c, \mathcal{H}_c \otimes \mathcal{H}_L, B_\mathcal{C} \otimes B_\mathcal{L})\), on which the product states \( \{ |t\rangle_c |\phi\rangle_L \}_{t \in [0, T]} \subseteq \mathcal{H}_c \otimes \mathcal{H}_L \) map and encode the entire computational history of the BQP algorithm. Then Feynman’s clocked Hamiltonians \( H_{\text{Feyn}, t} \equiv |t\rangle_c ((t-1)|c\rangle \otimes U_t + (t-1)|c\rangle \otimes U_t, t \in [1, T] \) ensure that the associated quantum gates \( U_t, t \in [1, T] \) are applied to the logic register in the correct order when the clock register undergoes transitions between what he called the program counter sites (namely, the clock states) \( |t\rangle_c, t \in [1, T] \). Finally, Kitaev’s GSQC Hamiltonian (also called the Feynman-Kitaev Hamiltonian) \( H_{\text{FK}} \equiv H_{\text{clock}} + H_{\text{init}} + H_{\text{prop}} \) enforces computational constraints via energy penalties, with \( H_{\text{clock}} \) restricting the clock register to the manifold of span(\{ |t\rangle_c : t \in [0, T] \}), \( H_{\text{init}} \) setting the initial state, while \( H_{\text{prop}} \) performing the quantum computation as Feynman suggested, such that the ground state \( \psi_0(H_{\text{FK}}) = (T+1)^{-1/2} \sum_{t=0}^T |t\rangle_c |\phi_t\rangle_L \) is unique and polynomially gapped \([11, 12]\).

There are several choices of encoding a clock register for the clock states \( \{ |t\rangle_c \}_{t \in [0, T]} \) \([37, 38]\). Take Kitaev’s domain wall clock for example, which has a clock register consisting of \( T + 2 \) rebits indexed by integers within \([0, T + 1]\), and uses \( |t\rangle_c \equiv |1\rangle^{\otimes (t+1)} |0\rangle^{\otimes (T-t+1)}, t \in [0, T] \), such that \([8, 11, 12]\)

\[
\begin{align*}
H_{\text{FK}} & \equiv H_{\text{clock}} + H_{\text{init}} + \sum_{t=1}^T H_{\text{prop}, t}, \\
H_{\text{clock}} & \equiv Z_{c, 0}^+ + Z_{c, t+1}^- + \sum_{t=1}^T Z_{c, t-1}^- \otimes Z_{c, t}^+, \\
H_{\text{init}} & \equiv Z_{c, 1}^- \otimes (I - |\phi_0\rangle_L \langle \phi_0|_L), \\
H_{\text{prop}, t} & \equiv Z_{c, t-1}^- \otimes Z_{c, t+1}^+ \otimes (I - X_{c, t} \otimes U_t), \ \forall t \in [1, T],
\end{align*}
\]

where \( Q_{c, t}^{\delta} \) means to apply a single-rebit operator \( Q^\delta \) to the \( t \)-th rebit of the clock register, \( \forall \delta \in \{+, -, \text{void}\}, \ \forall t \in [0, T + 1] \). The partial Hamiltonians \( \{ H_{\text{prop}, t} \}_{t \in [1, T]} \) are called the Feynman-Kitaev propagators. It is clear that size(\( H_{\text{FK}} \)) = \( O(T + n) \). It is WLOG to assume that the initial state \( |\phi_0\rangle_L \langle \phi_0|_L \) is \( \mathcal{C}_L \)-diagonal, because, otherwise, \( |\phi_0\rangle_L \) must be preparable from a \( \mathcal{C}_L \)-coordinate eigenstate by another BQP algorithm. It is straightforward to implement such an \( H_{\text{FK}} \) into an \( \mathcal{M}(H_{\text{FK}}) \) for a system of \( 2T + n + 2 \) bi-fermions, where each of the \( T + 2 \) clock rebits and \( n \) logic rebits corresponds to one unique bi-fermion, each of the operator monomials in equations \([25, 26]\) is mapped straightforwardly to an interaction among the corresponding bi-fermions as in equation \([23]\), while the remaining \( T \) bi-fermions supply enough auxiliary rebits for perturbative gadgets to implement logic gates of the form \( R(\theta) \otimes X, \ \theta \in [-\pi, \pi] \) that may appear in an operator monomial of equation \([27]\).

**Theorem 2.** A homophysics \( \mathcal{M} \) exists, which maps the Feynman-Kitaev Hamiltonian \( H_{\text{FK}} \) as defined in equations \([24, 27]\) to an SFF Hamiltonian \( \mathcal{M}(H_{\text{FK}}) \) that is both LTK- and GSP-decomposed.

**Proof.** An \( \mathcal{M}(H_{\text{FK}}) \) as constructed above is obviously SFF, with each FFI corresponding to one of the operator monomials in equations \([25, 27]\), each FFI involving no more than 6 bi-fermions and moving even less. Such an SFF Hamiltonian is clearly LTK-decomposed, which also happens to be GSP-decomposed, because \( H_{\text{FK}} \) is frustration-free \([8, 11, 12, 39, 40]\), so is \( \mathcal{M}(H_{\text{FK}}) \).

**Theorem 3.** \( \text{BQP} \subseteq \text{BPP}, \text{ therefore } \text{BPP} = \text{BQP}, \text{ as } \text{BPP} \subseteq \text{BQP} \text{ is well known.}

**Proof.** By Theorem 2, any BQP algorithm of size \( N \in \mathbb{N} \) can be mapped to an SFF Hamiltonian \( \mathcal{M}(H_{\text{FK}}) \) with size(\( \mathcal{M}(H_{\text{FK}}) \)) = \( O(\text{poly}(N)) \), which is both LTK- and GSP-decomposed, as well as polynomially energy gapped. Theorem 1 says that such an \( \mathcal{M}(H_{\text{FK}}) \) can be efficiently simulated via Monte Carlo on a classical computer, hence \( \text{BQP} \subseteq \text{BPP} \).

In conclusion, it has been proved that BPP and BQP are exactly the same computational complexity class. As a consequence, any quantum system can be efficiently simulated via Monte Carlo on a classical computer, by just constructing a BQP algorithm simulating the quantum system, then mapping the BQP algorithm to an efficient MCMC through a Feynman-Kitaev construct.

Finally, it is useful to note that the analyses, algorithms, and methods presented supra can be extended straightforwardly to physical and computational systems over a discrete or continuous-discrete product configuration space \([8]\), only that the nodal restriction or path rectification may need to invoke the so-called lever rule \([8, 11, 12]\) using efficiently solved nodal structures of FFIs or their associated Gibbs kernels.
