Gauge-invariant quantum minimally entangled typical thermal states algorithm with mutually unbiased physical bases for $\mathbb{Z}_2$ lattice gauge theories at finite temperature and density

Reita Maeno *

Department of Physics, Graduate School of Science, The University of Tokyo, Tokyo 113-0033, Japan
and Yukawa Institute for Theoretical Physics, Kyoto University, Kitashirakawa Oiwakecho, Sakyo-ku, Kyoto 606-8502, Japan

(Received 25 May 2026; accepted 13 August 2026; published 14 September 2026)

In quantum computations of gauge theories at finite temperature and finite density, enforcing Gauss's law for all states contributing to the thermal ensemble is a nontrivial challenge. In this work, we adapt the quantum minimally entangled typical thermal states (QMETTS) algorithm to $\mathbb{Z}_2$ gauge-constrained systems and propose a method for computing finite-temperature and finite-density expectation values without eliminating redundant gauge-field degrees of freedom. In QMETTS, the thermal ensemble is sampled via a Markov chain of pure states generated by imaginary-time evolution and projective measurements. To preserve gauge invariance while maintaining efficient sampling, we introduce measurement bases that are gauge invariant and mutually unbiased within the physical subspace. We show that such measurement bases can be constructed efficiently for $\mathbb{Z}_2$ lattice gauge theories in arbitrary spatial dimensions and arbitrary boundary conditions by exploiting the correspondence between $\mathbb{Z}_2$ lattice gauge theories and the stabilizer formalism. Furthermore, since expectation-value estimation on quantum hardware is inherently affected by shot noise, we explicitly incorporate shot noise into the analysis. By formulating a finite-shot version of QMETTS, we show that the resulting estimator remains unbiased and that using one observable-measurement shot per sampled state is nearly optimal in terms of variance when the total number of circuit executions used to generate the Markov chain and measure observables is fixed. This result indicates that it is often more efficient to generate more QMETTS samples than to accurately estimate the expectation value for each individual pure state. We validate the proposed method numerically in a $(1+1)$-dimensional $\mathbb{Z}_2$ lattice gauge theory coupled to staggered fermions. Our results provide a gauge-invariant sampling framework for finite-temperature and finite-density quantum algorithms for lattice gauge theories.

DOI: 10.1103/qyj2-97tw

I. INTRODUCTION

Understanding the properties of equilibrium systems at finite temperature and finite density in gauge theories is a central issue in modern theoretical physics [1–3]. *Ab initio* numerical studies of gauge theories are among the most powerful tools for investigating such phenomena, and lattice gauge theory (LGT) provides a natural framework for this purpose through lattice regularization [4,5]. Conventional Monte Carlo studies of lattice gauge theories based on the Euclidean Lagrangian formalism have achieved remarkable success in reproducing a wide range of nonperturbative phenomena [6]. However, at finite density, these approaches suffer from the notorious sign problem, which severely limits their applicability [7].

Recently, quantum computing and quantum simulation have attracted considerable attention as promising routes to overcome this limitation. For finite-temperature and finite-density equilibrium calculations, however, one must access thermal ensembles rather than a single pure state. Although various quantum algorithms have been proposed for this purpose, preparing thermal states on quantum devices remains challenging [8–11]. One attractive alternative is the quantum minimally entangled typical thermal states (QMETTS) algorithm [12]. QMETTS estimates finite-temperature and finite-density observables by sampling an ensemble of pure states through a Markov chain, avoiding directly preparing a full thermal mixed state, which can require substantial quantum resources.

In addition to these general challenges of finite-temperature calculations, gauge theories are subject to constraints imposed by Gauss's law. Physical thermal states are supported on the physical Hilbert space, namely, the subspace spanned by states satisfying the gauge constraints. This constraint makes the design of quantum algorithms for finite-temperature gauge theories nontrivial. One common strategy is to explicitly solve the gauge constraints and work directly within the physical Hilbert space [13–17]. While this removes unphysical degrees of freedom, it can make the Hamiltonian more complicated. In addition, retaining the redundant gauge-field degrees of freedom can itself be advantageous, since gauge constraints provide useful diagnostics and can be exploited for error mitigation [18–22]. Therefore, it is important to develop methods that evaluate finite-temperature

*Contact author: maeno-reita916@g.ecc.u-tokyo.ac.jp

expectation values without explicitly eliminating the redundant Hilbert-space structure [23–25].

In this work, we develop a QMETTS-based framework for computing gauge-invariant expectation values while retaining the redundant Hilbert-space structure. QMETTS consists mainly of two steps: imaginary-time evolution (ITE) and projective measurements that generate the next pure state in the Markov chain. The imaginary-time evolution step preserves gauge symmetry as long as the Hamiltonian commutes with the Gauss's law operators. The projective-measurement step, on the other hand, is more subtle. In conventional many-body systems, efficient QMETTS sampling is often achieved by alternating mutually unbiased bases (MUBs), such as the Z and X bases [see Eq. (8) for the definition of mutually unbiasedness], whereas in gauge theories these standard bases generally violate the Gauss's law constraints. Therefore, realizing QMETTS for gauge theories requires measurement bases that satisfy two nontrivial requirements simultaneously: They must preserve gauge symmetry and retain mutual unbiasedness needed for efficient sampling.

To address this issue, we introduce mutually unbiased physical bases (MUPBs) for $\mathbb{Z}_2$ gauge theories, which satisfy the above two conditions. By exploiting the correspondence between Gauss's law constraints and the stabilizer formalism, which is well developed in the context of quantum error correction [26,27], we present a general construction of such bases in arbitrary dimensions and with arbitrary boundary conditions. The resulting circuits require $\mathcal{O}(N_q^2)$ gates and have depth $\mathcal{O}(\log N_q)$, where N_q is the number of qubits. Although related tensor-network studies have developed symmetry-adapted measurement bases for METTS, mainly for condensed-matter systems with global symmetries [28,29], extending these ideas to local gauge constraints and implementing the resulting measurements as scalable, circuit-efficient quantum procedures is nontrivial. Our construction is tailored to the quantum-computing setting and provides explicit quantum circuits for gauge-preserving mutually unbiased measurements. A more detailed comparison with these tensor-network approaches is given in Sec. IV A.

In addition to constructing MUPBs, we analyze the effect of shot noise on expectation-value estimation in QMETTS. Although shot noise is unavoidable on quantum hardware, its role in QMETTS has not been fully clarified. We formulate a finite-shot version of QMETTS and analyze how finite measurement statistics affect the mean and variance of the estimator. In this analysis, we show that the estimator remains unbiased under the stationary Markov chain. We then derive the optimal number of observable-measurement shots for each sampled pure state under a fixed total number of circuit executions used to generate the Markov chain and measure observables. While this optimal shot number is generally difficult to know *a priori*, we show that the single-shot strategy is near optimal: Its variance is at most twice the optimum.

As a numerical demonstration, we apply our algorithm to the $(1+1)$-dimensional $\mathbb{Z}_2$ lattice gauge theory coupled to staggered fermions. We investigate the performance of the proposed method and verify gauge-invariant sampling, improved mixing, and accurate finite-temperature and finite-density expectation values. Furthermore, to assess the practical performance of our framework, we investigate our algorithm in combination with the quantum imaginary-time evolution (QITE) method proposed by Motta *et al.* [12], which provides one approach to implementing imaginary-time evolution on quantum computers.

The rest of this paper is organized as follows. Section II introduces the $(1+1)$-dimensional $\mathbb{Z}_2$ lattice gauge theory as a concrete example. Section III briefly reviews the QMETTS algorithm. In Sec. IV, we present our approach to evaluating gauge-invariant expectation values at finite temperature and finite density within the QMETTS framework. In Sec. V, we study statistical errors in the estimation of expectation values and optimize the number of shots. Based on Secs. IV and V, Sec. VI presents numerical demonstrations for the $\mathbb{Z}_2$ lattice gauge theory at finite temperature and finite density. Finally, Sec. VII summarizes our results and discusses future directions.

II. MODEL

For concreteness, we first focus on $(1+1)$-dimensional $\mathbb{Z}_2$ LGT coupled to staggered fermions, which is the simplest discrete realization of the Schwinger model [30]. Despite its minimal local Hilbert space consisting of two-level systems, this model exhibits nontrivial properties such as confinement and has been widely used as a benchmark for quantum simulations of lattice gauge theories [31,32]. In the staggered formulation, staggered fermions at site $2n-1$ and $2n$, χ_{2n-1} and χ_{2n}, represent the two components of a Dirac spinor at site n, ψ_n. Explicitly, a Dirac fermion ψ_n is represented by

$$\frac{\psi_n}{\sqrt{a}} = \begin{pmatrix} \chi_{2n-1} \\ \chi_{2n} \end{pmatrix},$$

where a denotes the lattice spacing. In this formulation, staggered fermions reside on lattice sites and gauge fields live on the links between neighboring sites. An annihilation (creation) operator of the staggered fermion $\chi_n(\chi_n^\dagger)$ satisfies the canonical anticommutation relation $\{\chi_n, \chi_m^\dagger\} = \delta_{n,m}$. The Hamiltonian is given by

$$H = \frac{1}{2a}\sum_{n=1}^{L_{\mathrm{KS}}-1}\left[\chi_n^\dagger \sigma_{n,n+1}^Z \chi_{n+1} + \mathrm{H.c.}\right] + ag^2\sum_{n=1}^{L_{\mathrm{KS}}-1}\left[1-\sigma_{n,n+1}^X\right] + m\sum_{n=1}^{L_{\mathrm{KS}}}(-1)^{n+1}\chi_n^\dagger\chi_n, \quad (1)$$

where g and m are the gauge coupling and the fermion mass, respectively. $\sigma_{n,n+1}^Z$ and $\sigma_{n,n+1}^X$ are Pauli operators in the $\mathbb{Z}_2$ link and electric field operators, respectively. As noted above, the number of Dirac-fermion sites, L_D, is half the number of staggered-fermion sites, L_{KS}, i.e., $L_D = L_{\mathrm{KS}}/2$.

After mapping fermionic degrees of freedom into spin operators via Jordan-Wigner transformation [33], the qubit description of the above Hamiltonian (1) is given in the following form:

$$H = -\frac{1}{4a}\sum_{n=1}^{L_{\mathrm{KS}}-1}\left[X_n\sigma_{n,n+1}^Z X_{n+1} + Y_n\sigma_{n,n+1}^Z Y_{n+1}\right] + ag^2\sum_{n=1}^{L_{\mathrm{KS}}-1}\left[1-\sigma_{n,n+1}^X\right] + \frac{m}{2}\sum_{n=1}^{L_{\mathrm{KS}}}(-1)^{n+1}Z_n,$$

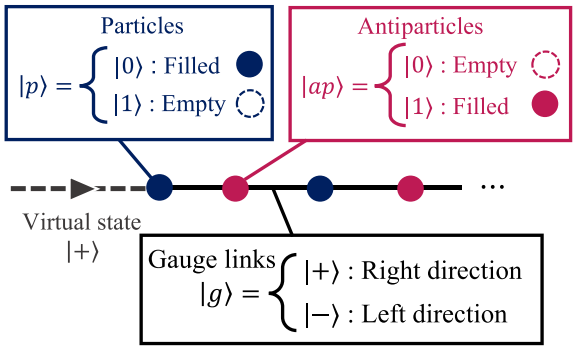

FIG. 1. Schematic setup for $(1+1)$-dimensional $\mathbb{Z}_2$ LGT coupled to staggered fermions. Occupied sites correspond to $|0\rangle$ ($|1\rangle$) on particle (antiparticle) sites, while empty sites correspond to $|1\rangle$ ($|0\rangle$), respectively. For gauge links, $|+\rangle$ represents right-directional flux and $|-\rangle$ represents left-directional flux. We impose open boundary conditions, with the left boundary fixed by a virtual link in the $|+\rangle$ state.

where X_n, Y_n, and Z_n are usual Pauli matrices acting on the qubit at site n. In this representation, as shown in Fig. 1, the qubit basis is defined as follows. On odd sites, which correspond to particle states, $|0\rangle$ ($|1\rangle$) corresponds to an occupied (empty) fermionic state, whereas on even (antiparticle) sites, $|1\rangle$ ($|0\rangle$) represents an occupied (empty) state. For gauge links, $|+\rangle$ ($|-\rangle$) represents right (left) directional flux. We impose open boundary conditions, with the left boundary fixed by a virtual link in the $|+\rangle$ state.

In addition, gauge theories are subject to local constraints associated with gauge symmetry, known as Gauss's law constraints. In our setup, Gauss's law operators are represented by

$$G_n = (-1)^n \sigma^X_{n-1,n} Z_n \sigma^X_{n,n+1} (1 \leqslant n \leqslant L_{\mathrm{KS}} - 1). \tag{2}$$

Due to the boundary condition, we fix $\sigma^X_{0,1} = 1$. Their eigenvalues are $g_n = \pm 1$. They commute with each other, $[G_n, G_m] = 0, \ \forall n, m$ and with the Hamiltonian $[H, G_n] = 0, \ \forall n$. In this work, we choose a sector without background, called physical sector, defined by $g_n = +1, \ \forall n$. Once we restrict to the physical sector, the states contributing to expectation values are only physical states $|\psi\rangle$ satisfying $G_n|\psi\rangle = +|\psi\rangle, \ \forall n$. We denote the subspace spanned by physical states by $\mathcal{H}_{\mathrm{phys}}$ and total Hilbert space by $\mathcal{H}_{\mathrm{tot}}$. Enforcing this constraint poses a nontrivial challenge for the computation of gauge theories.

To study thermal-equilibrium properties of the theory, we consider the grand-canonical expectation values of an observable $\mathcal{O}$ under inverse temperature $\beta = 1/T$ and chemical potential μ:

$$\langle \mathcal{O} \rangle_{\beta,\mu} := \mathrm{Tr}_{\mathcal{H}_{\mathrm{phys}}} \left[\mathcal{O} \frac{e^{-\beta(H-\mu N)}}{Z} \right], \tag{3}$$

where $Z = \mathrm{Tr}_{\mathcal{H}_{\mathrm{phys}}}[e^{-\beta(H-\mu N)}]$ is the partition function and N is the number operator given by

$$N := \sum_{n=1}^{L_{\mathrm{KS}}} \chi_n^\dagger \chi_n - \frac{L_{\mathrm{KS}}}{2} I = \frac{1}{2} \sum_{n=1}^{L_{\mathrm{KS}}} Z_n. \tag{4}$$

Here, we have dropped the additive constant that appears in the Jordan-Wigner mapping of the total staggered occupation number, since it only shifts $H - \mu N$ by a constant and therefore does not affect normalized thermal expectation values. Note that the trace $\mathrm{Tr}_{\mathcal{H}_{\mathrm{phys}}}[\cdot]$ in Eq. (3) is taken over the physical subspace $\mathcal{H}_{\mathrm{phys}}$. Since the number operator N in Eq. (4) commutes with all Gauss's law operators G_n in Eq. (2), the grand-canonical ensemble is compatible with gauge invariance.

For $\mathbb{Z}_2$ LGTs in a general setup, as in the $(1+1)$-dimensional case, the Gauss's law operators can be described by products of Pauli operators and mutually commute. Even in the case of the presence of fermionic degrees of freedom, we can express them as the Pauli products by appropriate qubit transformation such as Jordan-Wigner transformation. The physical Hilbert space is defined as the common $+1$ eigenspace of these operators, $G_n|\psi\rangle = +|\psi\rangle, \quad \forall n$, where n is the index of vertices in general lattice setup. Thus, all Gauss's law operators of $\mathbb{Z}_2$ LGTs in general dimensions and any boundary condition can also be represented as a set of commuting Pauli operators.

III. QMETTS ALGORITHM

In this section, we review the QMETTS algorithm. QMETTS is the quantum extension of the minimally entangled typical thermal states (METTS) algorithm, which was originally proposed by White and Stoudenmire [34,35] in the context of tensor-network methods. Motta *et al.* [12] adapted this framework to quantum computing and introduced the QMETTS algorithm. It provides a method for estimating expectation values of observables in finite-temperature and finite-density quantum many-body systems. In this section, we review the standard formulation of QMETTS without gauge constraints or shot noise. Gauge constraints will be incorporated into our modified scheme in Sec. IV, and shot noise will be discussed separately in Sec. V.

To explain the QMETTS algorithm, we rewrite the expectation value $\langle \mathcal{O} \rangle_{\beta,\mu} = \mathrm{Tr}[\mathcal{O} e^{-\beta(H-\mu N)}/Z]$ in terms of an orthonormal basis $\{|n\rangle\}$:

$$\begin{aligned} \langle \mathcal{O} \rangle_{\beta,\mu} &= \sum_{|i\rangle \in \{|n\rangle\}} \frac{1}{Z} \langle i | e^{-\beta(H-\mu N)/2} \mathcal{O} e^{-\beta(H-\mu N)/2} | i \rangle \\ &= \sum_{|i\rangle \in \{|n\rangle\}} \mathrm{Prob}_i \times \langle \phi_i | \mathcal{O} | \phi_i \rangle, \end{aligned} \tag{5}$$

where $|\phi_i\rangle = e^{-\beta(H-\mu N)/2}|i\rangle / \sqrt{\langle i | e^{-\beta(H-\mu N)} | i \rangle}$ and $\mathrm{Prob}_i = \langle i | e^{-\beta(H-\mu N)} | i \rangle / Z$. Since $|\phi_i\rangle$ is normalized, it represents a pure quantum state, and Prob_i forms a probability distribution satisfying $\mathrm{Prob}_i \geqslant 0$ and $\sum_i \mathrm{Prob}_i = 1$. When the basis $\{|n\rangle\}$ is chosen to be a product basis, the states $|\phi_i\rangle$ are often expected to have relatively low entanglement. For this reason, the states $|\phi_i\rangle$ are referred to as METTS following Ref. [34]. Therefore, as described in Eq. (5), the thermal expectation value can be interpreted as an average of the expectation values under each METTS $\langle \phi_i | \mathcal{O} | \phi_i \rangle$ weighted by Prob_i.

The goal of the QMETTS algorithm is to efficiently sample these contributions by generating a Markov chain of basis states $|i\rangle$ whose stationary distribution is Prob_i. The procedure is as follows. First, we sample one state $|i\rangle$ from the mea-

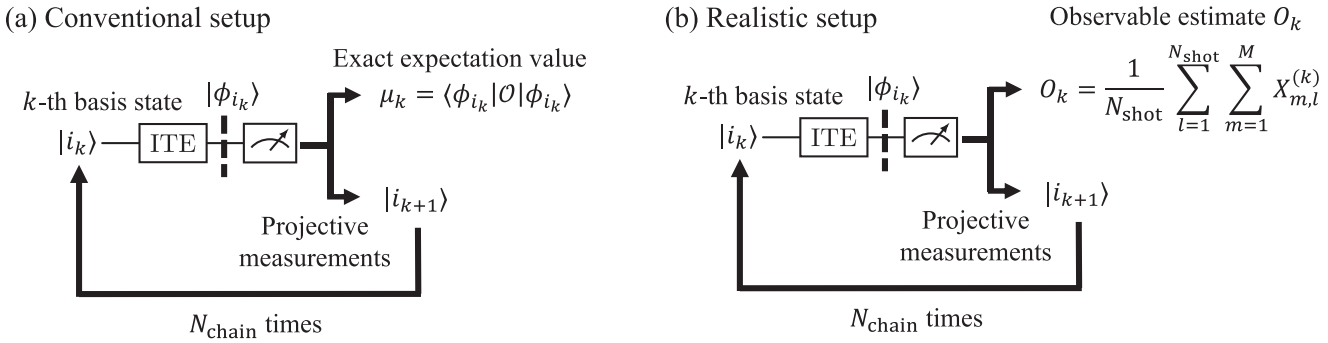

FIG. 2. Schematic illustration of two QMETTS measurement schemes. Many previous works assume an idealized setting in which shot noise is neglected and expectation values are evaluated exactly [panel (a)]. In contrast, on quantum hardware, expectation values must be estimated from a finite number of measurements. We therefore consider a realistic setting [panel (b)], where the observable $\mathcal{O} = \sum_{m=1}^{M} \mathcal{O}_m$ is decomposed into M components, each of which can be estimated using a single measurement circuit. For each of the N_{chain} Markov-chain samples, we perform N_{shot} measurements for each component, resulting in $N_{\text{shot}}M$ circuit executions for observable estimation, and one additional projective measurement to generate the next basis state. The corresponding total number of circuit executions is $N_{\text{tot}} = (N_{\text{shot}}M + 1)N_{\text{chain}}$. In the case of alternating measurements, $|i_k\rangle$ and $|i_{k+1}\rangle$ are replaced by $|i_k^{(r)}\rangle$ and $|i_{k+1}^{(s)}\rangle$, respectively, where $(r, s) = (1, 2)$ or $(2, 1)$.

surement basis $\{|n\rangle\}$ and apply the ITE operator $e^{-\beta(H-\mu N)/2}$ to $|i\rangle$, preparing the normalized imaginary-time-evolved state $|i\rangle \mapsto |\phi_i\rangle = e^{-\beta(H-\mu N)/2}|i\rangle/\sqrt{\langle i|e^{-\beta(H-\mu N)}|i\rangle}$ on a quantum computer. Next, we perform a projective measurement for the METTS $|\phi_i\rangle$ with the measurement basis $\{|n\rangle\}$ and obtain a new collapse state $|i'\rangle$. By repeating the processes of the ITE and the projective measurement, one can generate the Markov chain with the stationary distribution Prob_i because the transition probability from $|i\rangle$ to $|i'\rangle$, $T_{i\to i'} := |\langle i'|\phi_i\rangle|^2$, satisfies the detailed balance condition:

$$\frac{T_{i\to i'}}{T_{i'\to i}} = \frac{\langle i'|e^{-\beta(H-\mu N)}|i'\rangle}{\langle i|e^{-\beta(H-\mu N)}|i\rangle} = \frac{\text{Prob}_{i'}}{\text{Prob}_i}.$$

We generate a Markov chain of length N_{chain}, which produces a sequence of sampled METTS $\{|\phi_{i_k}\rangle\}_{k=1}^{N_{\text{chain}}}$. For the METTS $|\phi_{i_k}\rangle$ generated at the kth step, we compute the observable estimate $\mu_k = \langle\phi_{i_k}|\mathcal{O}|\phi_{i_k}\rangle$ (see the left of Fig. 2). Defining the Monte Carlo average $\bar{\mu} = (1/N_{\text{chain}})\sum_{k=1}^{N_{\text{chain}}} \mu_k$, this estimator satisfies

$$\mathbb{E}[\bar{\mu}] = \frac{1}{N_{\text{chain}}}\sum_{k=1}^{N_{\text{chain}}} \mathbb{E}[\mu_k] = \langle\mathcal{O}\rangle_{\beta,\mu}.$$

Thus, $\bar{\mu}$ is an unbiased estimator of $\langle\mathcal{O}\rangle_{\beta,\mu}$.

To quantify the statistical error of the Markov chain, one must account for the autocorrelations between successive samples. Let τ_μ denote the integrated autocorrelation time for the sequence of $\{\mu_k\}_{k=1}^{N_{\text{chain}}}$. For a sufficiently long Markov chain, the variance of the Monte Carlo average $\bar{\mu}$ is then given by [36–38]

$$\text{Var}[\bar{\mu}] = \frac{\tau_\mu}{N_{\text{chain}}}\sigma_\mu^2, \tag{6}$$

where $\sigma_\mu^2 = \sum_i \text{Prob}_i \times \langle\phi_i|\mathcal{O}|\phi_i\rangle^2 - (\sum_i \text{Prob}_i \times \langle\phi_i|\mathcal{O}|\phi_i\rangle)^2$ represents the variance of $\langle\phi_i|\mathcal{O}|\phi_i\rangle$ with respect to the METTS distribution Prob_i. Equation (6) shows that the Markov chain effectively yields $N_{\text{eff}} = N_{\text{chain}}/\tau_\mu$ independent samples. The integrated autocorrelation time τ_μ is defined by

$$\tau_\mu = 1 + 2\sum_{t=1}^{\infty} \rho_\mu(t), \qquad \rho_\mu(t) = \frac{\text{Cov}(\mu_k, \mu_{k+t})}{\sigma_\mu^2}, \tag{7}$$

where $\rho_\mu(t)$ is the normalized autocorrelation function. This function quantifies the degree of correlation between samples. For completely independent samples, $\rho_\mu(t) = 0$ for all $t \geqslant 1$, resulting in $\tau_\mu = 1$.

In practice, however, using only a single measurement basis cannot be efficient for two distinct reasons. First, if the chosen measurement basis consists of states with definite values of a conserved charge Q satisfying $[Q, H] = 0$, transitions between different charge sectors are forbidden. For example, if $Q|i\rangle = q_i|i\rangle$, $Q|j\rangle = q_j|j\rangle$, and $q_i \neq q_j$, then $|\phi_i\rangle \propto e^{-\beta H/2}|i\rangle$ remains in the q_i sector, so $|\langle j|\phi_i\rangle|^2 = 0$. Hence, the chain can remain trapped in a fixed sector, and ergodicity is not guaranteed. Second, near the high-temperature limit, where the imaginary-time evolution approaches the identity $e^{-\beta H} \approx I$, using a single measurement basis can repeatedly yield the same collapse state, producing strongly correlated consecutive samples and hence long autocorrelation times (see Refs. [28,34,39] for details).

To make the algorithm efficient, one needs to use two measurement bases and alternate between them. Efficient mixing is promoted when the two bases $\{|n^{(1)}\rangle\}$ and $\{|m^{(2)}\rangle\}$ satisfy

$$|\langle i^{(1)}|j^{(2)}\rangle|^2 = \frac{1}{d}, \tag{8}$$

for all $|i^{(1)}\rangle \in \{|n^{(1)}\rangle\}$ and $|j^{(2)}\rangle \in \{|m^{(2)}\rangle\}$, where d denotes the Hilbert-space dimension [28,39]. Such bases address both issues by allowing transitions between conserved sectors and suppressing this high-temperature autocorrelation effect. The most representative bases that satisfy this relation are X and Z (i.e., computational) bases. In the language of quantum information, two orthonormal bases satisfying Eq. (8) are said to be mutually unbiased and are referred to as MUBs [40]. In summary, the efficient algorithm is as follows:

(1) Choose one state $|i^{(1)}\rangle$ from a measurement basis $\{|n^{(1)}\rangle\}$.

(2) Generate the METTS by the ITE, $|\phi_i^{(1)}\rangle = e^{-\beta(H-\mu N)/2}|i^{(1)}\rangle / \sqrt{\langle i^{(1)}|e^{-\beta(H-\mu N)}|i^{(1)}\rangle}$.

(3) Compute the observable estimate of $\mathcal{O}$ under the METTS $|\phi_i^{(1)}\rangle$; $\langle\phi_i^{(1)}|\mathcal{O}|\phi_i^{(1)}\rangle$.

(4) Measure $|\phi_i^{(1)}\rangle$ with another measurement basis $\{|m^{(2)}\rangle\}$ satisfying the condition of the MUB with $\{|n^{(1)}\rangle\}$, leading to a collapse into a basis state $|j^{(2)}\rangle \in \{|m^{(2)}\rangle\}$ with probability $|\langle j^{(2)}|\phi_i^{(1)}\rangle|^2$.

(5) Repeat the same step from 2 to 4 switching the role of $\{|n^{(1)}\rangle\}$ and $\{|m^{(2)}\rangle\}$ for N_{chain} times.

(6) Average the samples $\langle\phi_i^{(a)}|\mathcal{O}|\phi_i^{(a)}\rangle$ $(a = 1, 2)$.

If all these steps are simulated classically, one recovers the original METTS algorithm. However, we formulate the construction in the quantum-circuit setting, because the gauge-preserving measurement bases introduced below are naturally implemented as quantum circuits and need not be simple product bases.

In the subsequent section, we extend this approach to gauge theories. The above strategy works in generic quantum many-body systems [34]. In gauge theories, however, we need to preserve gauge invariance. Starting from an initial state in physical subspace, the imaginary-time evolved states preserve gauge invariance due to $[H, G_n] = 0$. Nevertheless, generic projective measurements do not respect the gauge constraints and may collapse a physical state into an unphysical one. In the next section, we propose measurement bases preserving gauge symmetries while satisfying Eq. (8) within the physical sector.

Moreover, the above review and most previous works do not take into account shot noise in the estimation of expectation values, although such noise is unavoidable in quantum computation. In Sec. V, we extend the above discussion to incorporate shot noise and optimize the total number of circuit runs.

IV. EFFICIENT CONSTRUCTION OF MUPB IN $\mathbb{Z}_2$ LATTICE GAUGE THEORIES

In this section, we develop a QMETTS-based method for estimating the expectation value $\langle\mathcal{O}\rangle_{\beta,\mu}$ in $\mathbb{Z}_2$ LGTs. The central task of this section is to construct measurement bases that (1) preserve gauge invariance and (2) remain mutually unbiased within the physical subspace, while still being efficiently implementable. First, we refer to such a pair of bases as the MUPB, defined as follows.

Definition 1 (Mutually unbiased physical bases). Two measurement bases $\{|n^{(1)}\rangle\}$ and $\{|m^{(2)}\rangle\}$ are called mutually unbiased physical bases if they satisfy the following properties:

(1) Every state in the measurement bases $\{|n^{(1)}\rangle\}$ and $\{|m^{(2)}\rangle\}$ is an eigenstate of all Gauss's law operators:

$$G_n|i^{(a)}\rangle = g_n|i^{(a)}\rangle, \quad g_n \in \{\pm 1\}, \quad a = 1, 2.$$

(2) Any two physical states $|i^{(1)}\rangle, |j^{(2)}\rangle \in \mathcal{H}_{\text{phys}}$ satisfy the mutually unbiased condition within physical subspace $\mathcal{H}_{\text{phys}}$:

$$|\langle i^{(1)}|j^{(2)}\rangle|^2 = \frac{1}{d_{\text{phys}}},$$

where d_{phys} denotes the dimension of the physical subspace $\mathcal{H}_{\text{phys}}$.

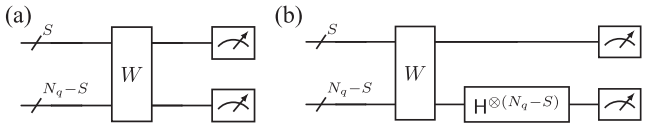

FIG. 3. Quantum circuits for implementing the MUPB. W is a classically computable Clifford unitary that can be implemented using $\mathcal{O}(N_q^2)$ Clifford gates with depth $\mathcal{O}(\log N_q)$. Circuit (a) implements the first measurement basis by applying W before computational-basis measurements. Circuit (b) implements the second measurement basis by applying W, followed by Hadamard gates on the remaining $N_q - S$ qubits, before computational-basis measurements.

The first condition ensures compatibility with the gauge constraints, while the second promotes mixing within the physical subspace. Because each basis state has definite Gauss's law eigenvalues, projective measurements in MUPB do not mix different gauge sectors. Therefore, if the initial state is physical, the resulting Markov chain remains entirely within the physical sector and never collapses into a gauge-violating state.

Since measurements in quantum computation are typically performed in the computational basis, MUPB must be implemented by designing appropriate quantum circuits. However, the definition of MUPB may appear complicated, and it is not obvious how to construct such bases efficiently using quantum circuits. To address this problem, we use the stabilizer formalism, a mathematical framework originally developed in the context of quantum error correction [26,27]. In this formalism, logical states are characterized as simultaneous $+1$ eigenstates of a set of commuting Pauli-product operators called stabilizers $\mathcal{S} = \{S_j\}$,

$$S_j|\psi\rangle = +|\psi\rangle, \quad [S_j, S_k] = 0, \quad \text{for } \forall S_j, S_k \in \mathcal{S}.$$

In a similar way, as mentioned in Sec. II, for $\mathbb{Z}_2$ LGTs in arbitrary spatial dimensions and with arbitrary boundary conditions, the gauge constraints can be viewed as commuting Pauli-product operators that define the physical subspace. Thus, physical states are precisely the states stabilized by these gauge constraint operators.

Based on this perspective, we use the canonical form of stabilizer groups discussed in Ref. [41]. This result implies that any stabilizer group can be transformed by a Clifford unitary into a canonical form generated solely by Pauli-Z operators. Applying this theorem to the gauge constraints, for Gauss's law operators $G_n (n = 1, \ldots, S)$ in $\mathbb{Z}_2$ LGTs in arbitrary spatial dimensions and with arbitrary boundary conditions, we obtain

$$W G_n W^\dagger = Z_n, \quad n = 1, \ldots, S,$$

where W is a classically computable Clifford unitary (see Appendix A). Furthermore, W can be implemented using $\mathcal{O}(N_q^2)$ Clifford gates with depth $\mathcal{O}(\log N_q)$, where N_q denotes the number of qubits. This transformation allows us to construct the MUPBs shown in Fig. 3 and to prove that they satisfy the defining conditions of the MUPB. Thus, the Clifford circuit required to realize MUPB has gate complexity $\mathcal{O}(N_q^2)$ and depth $\mathcal{O}(\log N_q)$, implying that the construction is efficient.

A proof that the bases shown in Fig. 3 satisfy the MUPB conditions is given in Appendix A.

Regarding the circuit complexity discussed here, we assume all-to-all connectivity and parallel application of Clifford gates. Whether the general MUPB construction described above can be implemented efficiently under nearest-neighbor connectivity remains unclear. Nevertheless, for the specific geometry introduced in Sec. II, we identify a specialized realization that uses only nearest-neighbor gates and has constant depth. This geometry-adapted realization is obtained heuristically using the same stabilizer-based perspective and satisfies the same MUPB conditions. It is described and validated in Sec. VI, and is used in the numerical simulations below.

The above stabilizer-based viewpoint also clarifies why the present construction is particularly natural in QMETTS compared with METTS. Since Clifford circuits contain entangling gates such as CNOT, the corresponding measurement bases need not be product bases. In the QMETTS setting, such gauge-preserving bases can nevertheless be implemented directly as measurement circuits.

Contribution and relation to prior work

The calculation of thermal equilibrium properties while preserving symmetries has been extensively studied within the framework of the METTS algorithm, particularly for systems with one or a few global conserved quantities by exploiting symmetry-adapted measurement bases in the context of condensed matter.

In Ref. [28], efficient mixing bases are achieved within the METTS framework by employing Fourier-transformed bases and Haar-random bases that respect global symmetries. While these constructions lead to good ergodic properties and short autocorrelations, they typically require high complexity.

Another related approach was proposed in Ref. [29], where nontrivial measurement bases are generated via time evolution under commuting operators with generators corresponding to symmetries within the METTS framework. This method does not generally guarantee efficient mixing of the Markov chain. Moreover, it remains unclear how this construction can be extended to accommodate local gauge constraints, which impose much more stringent conditions than global symmetries.

In contrast, our MUPB approach realizes mutually unbiased measurements within the physical subspace while preserving gauge invariance, and the corresponding basis changes can be implemented using shallow Clifford circuits. This provides a controlled way to enhance mixing within the physical subspace without sacrificing the redundant Hilbert-space structure.

V. SHOT-NOISE ANALYSIS

Unlike tensor-network methods, quantum computation inherently involves shot noise when estimating expectation values of observables. In this sense, the discussion in Sec. III corresponds to an idealized setting. Nevertheless, many previous analyses of the QMETTS algorithm assume exact expectation-value estimation, in close analogy with the ideal METTS framework. In this section, we extend the QMETTS analysis to include shot noise and investigate efficient strategies for observable estimation in realistic quantum-computing settings. Here, we address the following two questions: In the presence of shot noise, does the QMETTS estimator reproduce finite-temperature and finite-density expectation values, and what is the optimal number of observable-measurement shots per sampled collapse state when the total number of circuit executions used to measure observables and to generate the next collapse states in the Markov chain is fixed? Note that the analysis presented in this section is not limited to gauge systems and can be easily extended to the cases with alternating measurements. Throughout this analysis, we assume that the Markov chain is initialized from the stationary distribution. We also neglect gate noise, state-preparation errors, and ITE errors.

To address the questions, we consider the following setup. An observable $\mathcal{O}$ is decomposed as $\mathcal{O} = \sum_{m=1}^{M} \mathcal{O}_m$, where each $\mathcal{O}_m$ denotes a group of Pauli terms that are measured in the same measurement setting. Here, possible coefficients of Pauli strings are absorbed into the definition of $\mathcal{O}_m$. For a given METTS sample $|\phi_{i_k}\rangle$, we estimate each $\mathcal{O}_m$ using N_{shot} independent measurements on independent copies of the same state. We denote by $X_{m,l}^{(k)}$ the lth single-shot measurement outcome for $\mathcal{O}_m$ under the kth METTS sample, where $l = 1, \ldots, N_{\text{shot}}$. See the right illustration in Fig. 2. Thus, conditioned on the METTS sample $|\phi_{i_k}\rangle$, the measurement outcomes for distinct pairs (m, l) are statistically independent. The estimator for the observable under the kth METTS sample is then given by

$$O_k = \frac{1}{N_{\text{shot}}} \sum_{l=1}^{N_{\text{shot}}} \sum_{m=1}^{M} X_{m,l}^{(k)}.$$

Repeating this procedure along the METTS Markov chain yields a sequence of noisy observations $\{O_k\}_{k=1}^{N_{\text{chain}}}$. The corresponding Markov-chain estimator is defined as

$$\bar{O} = \frac{1}{N_{\text{chain}}} \sum_{k=1}^{N_{\text{chain}}} O_k.$$

Thus, the number of circuit executions used for observable estimation is $N_{\text{est}} = N_{\text{shot}} N_{\text{chain}} M$. Since one additional projective measurement is required to generate each subsequent METTS sample, the total number of circuit executions is

$$N_{\text{tot}} = N_{\text{est}} + N_{\text{chain}} = (N_{\text{shot}} M + 1) N_{\text{chain}}. \tag{9}$$

Now, we are in a position to answer the first question: whether the QMETTS estimator is unbiased. The sequence $\{O_k\}_k$ contains two sources of statistical fluctuation: the METTS-sampling fluctuations and the finite-shot measurement noise, making the analysis nontrivial. To separate these two sources, we first average over the shot noise while keeping the METTS samples fixed. Conditioned on the kth METTS sample $|\phi_{i_k}\rangle$, the expectation value of the single-shot outcome $X_{m,l}^{(k)}$ is given by

$$\mathbb{E}_{\text{shot}}\left[X_{m,l}^{(k)} \mid \phi_{i_k}\right] = \langle \phi_{i_k} | \mathcal{O}_m | \phi_{i_k} \rangle,$$

where $\mathbb{E}_{\text{shot}}[\cdot|\phi_{i_k}]$ denotes the expectation over the measurement outcomes conditioned on the fixed METTS sample $|\phi_{i_k}\rangle$.

Therefore,

$$\mathbb{E}_{\text{shot}}[O_k \mid \phi_{i_k}] = \sum_{m=1}^{M} \langle\phi_{i_k}|\mathcal{O}_m|\phi_{i_k}\rangle = \langle\phi_{i_k}|\mathcal{O}|\phi_{i_k}\rangle.$$

Using the law of total expectation, we then obtain

$$\begin{aligned}\mathbb{E}[\bar{O}] &= \mathbb{E}_{\text{chain}}\left[\mathbb{E}_{\text{shot}}\left[\bar{O} \mid \{\phi_{i_k}\}_{k=1}^{N_{\text{chain}}}\right]\right] \\ &= \sum_i \text{Prob}_i \langle\phi_i|\mathcal{O}|\phi_i\rangle = \langle\mathcal{O}\rangle_{\beta,\mu},\end{aligned}$$

where $\mathbb{E}_{\text{chain}}[\cdot]$ denotes the expectation over the METTS Markov chain. Thus, under the stationary distribution, the estimator remains unbiased for any N_{shot} and N_{chain}.

We next evaluate the variance of the Markov-chain estimator $\bar{O}$. As in the expectation-value calculation, we first consider the shot-noise variance for fixed METTS samples. For a fixed METTS sample $|\phi_{i_k}\rangle$, we define the conditional shot-noise variance $\sigma^2_{\text{shot}|\phi_{i_k}}$ by

$$\sigma^2_{\text{shot}|\phi_{i_k}} := \sum_{m=1}^{M} \left[\langle\phi_{i_k}|\mathcal{O}_m^2|\phi_{i_k}\rangle - (\langle\phi_{i_k}|\mathcal{O}_m|\phi_{i_k}\rangle)^2\right].$$

Then, the shot-noise variance of O_k conditioned on METTS $|\phi_{i_k}\rangle$ is given by

$$\text{Var}_{\text{shot}}[O_k|\phi_{i_k}] = \frac{1}{N_{\text{shot}}^2} \sum_{l=1}^{N_{\text{shot}}} \sigma^2_{\text{shot}|\phi_{i_k}} = \frac{\sigma^2_{\text{shot}|\phi_{i_k}}}{N_{\text{shot}}}.$$

Since the conditional single-shot variance $\sigma^2_{\text{shot}|\phi_{i_k}}$ depends on the METTS sample, we define its average over the METTS distribution as $\sigma^2_{\text{shot}} := \sum_i \text{Prob}_i \sigma^2_{\text{shot}|\phi_i}$. Using the law of the total variance, the variance of $\bar{O}$ with N_{shot} measurements is given by

$$\begin{aligned}\text{Var}[\bar{O}](N_{\text{shot}}) &= \text{Var}_{\text{chain}}\left[\mathbb{E}_{\text{shot}}\left[\bar{O} \mid \{\phi_{i_k}\}_{k=1}^{N_{\text{chain}}}\right]\right] \\ &\quad + \mathbb{E}_{\text{chain}}\left[\text{Var}_{\text{shot}}\left[\bar{O} \mid \{\phi_{i_k}\}_{k=1}^{N_{\text{chain}}}\right]\right] \\ &= \frac{\tau(N_{\text{shot}})}{N_{\text{chain}}}\left(\sigma_\mu^2 + \frac{\sigma^2_{\text{shot}}}{N_{\text{shot}}}\right), \qquad (10)\end{aligned}$$

where $\tau(N_{\text{shot}})$ is the integrated autocorrelation time of the observed sequence $\{O_k\}_{k=1}^{N_{\text{chain}}}$,

$$\begin{aligned}\tau(N_{\text{shot}}) &= 1 + 2\sum_{t=1}^{\infty} \rho_O(t, N_{\text{shot}}), \\ \rho_O(t, N_{\text{shot}}) &= \frac{\text{Cov}(O_k, O_{k+t})}{\sigma_\mu^2 + (1/N_{\text{shot}})\sigma^2_{\text{shot}}}. \qquad (11)\end{aligned}$$

When N_{shot} is sufficiently large that the shot noise is negligible, $\tau(N_{\text{shot}})$ is well approximated by the autocorrelation time of the original Markov chain τ_μ in Eq. (7).

This variance formula allows us to address the second question: What is the optimal choice of N_{shot} at fixed N_{tot}? Using Eq. (10) together with the budget constraint in Eq. (9), we minimize the estimator variance with respect to N_{shot}. We find

$$N^*_{\text{shot}} \sim \max\left(1, \sqrt{\frac{\sigma^2_{\text{shot}}}{M\sigma_\mu^2\tau_\mu}}\right), \qquad (12)$$

as derived in Appendix B. This expression shows that fewer shots are favored when the autocorrelation is strong, when the METTS-sampling variance σ_μ^2 is large, or when the number of measurement groups M is large. Conversely, when the uncertainty originating from shot noise σ^2_{shot} is large, it is advantageous to increase N_{shot}.

In practice, however, the quantities σ_μ^2, σ^2_{shot}, and τ_μ are generally not known in advance. Nevertheless, we find that the variance ratio between the single-shot strategy and the optimal-shot strategy satisfies

$$1 \leqslant \frac{\text{Var}[\bar{O}](1)}{\text{Var}[\bar{O}](N^*_{\text{shot}})} lt; 1 + \frac{1}{M} \leqslant 2, \qquad (13)$$

for any N^*_{shot} (see Appendix B). Thus, even without prior knowledge of the optimal parameters, the single-shot strategy provides a robust choice, with a variance at most twice the optimum.

It is also useful to compare the single-shot strategy with direct independent measurements on the Gibbs state. Although the general expression is more involved, it takes a simple form for $M = 1$ (see Appendix B for $M \neq 1$ cases). In this case, the total number of circuit executions is $N_{\text{tot}} = 2N_{\text{chain}}$, and the variance of the single-shot QMETTS estimator can be written as

$$\text{Var}[\bar{O}](1) = 2\tau(1)\text{Var}_{\text{Gibbs}}\left[\frac{1}{N_{\text{tot}}}\sum_{l=1}^{N_{\text{tot}}} Y_l\right], \qquad (14)$$

where Y_l denotes the lth independent shot measurement outcome of $\mathcal{O}$ on the Gibbs state. Here, $\text{Var}_{\text{Gibbs}}[\cdot]$ denotes the variance with respect to independent measurements performed on the Gibbs state, so that $\text{Var}_{\text{Gibbs}}[(1/N_{\text{tot}})\sum_{l=1}^{N_{\text{tot}}} Y_l]$ is the variance of the estimator obtained from N_{tot} independent Gibbs-state measurements. Thus, for $M = 1$, the single-shot QMETTS estimator differs from the corresponding Gibbs-state estimator by a factor of $2\tau(1)$. In particular, if the observable sequence exhibits sufficiently strong negative correlations such that $\tau(1) < 0.5$, the QMETTS variance can be smaller than that of direct Gibbs-state measurements. However, since the integrated autocorrelation time is generally not known before performing numerical simulations, it is difficult to design a strategy that is guaranteed in advance to outperform direct Gibbs-state measurements. We emphasize that this should not be regarded as a direct comparison of resource costs: It assumes access to independent measurements on the Gibbs state, even though the Gibbs state is mixed and cannot be prepared from a pure initial state by a unitary acting on the system alone, whereas QMETTS avoids this requirement by sampling an ensemble of pure states.

VI. NUMERICAL SIMULATION

In this section, we investigate the numerical performance of our algorithm, namely, QMETTS with MUPB and the single-shot strategy, by applying it to the $\mathbb{Z}_2$ LGT setup introduced in Sec. II. We first provide an explicit construction of the MUPB for the model considered here and next describe the numerical setup, including the simulation parameters and measurement protocol. We then present the results obtained using classically computed exact ITE. Finally, toward the

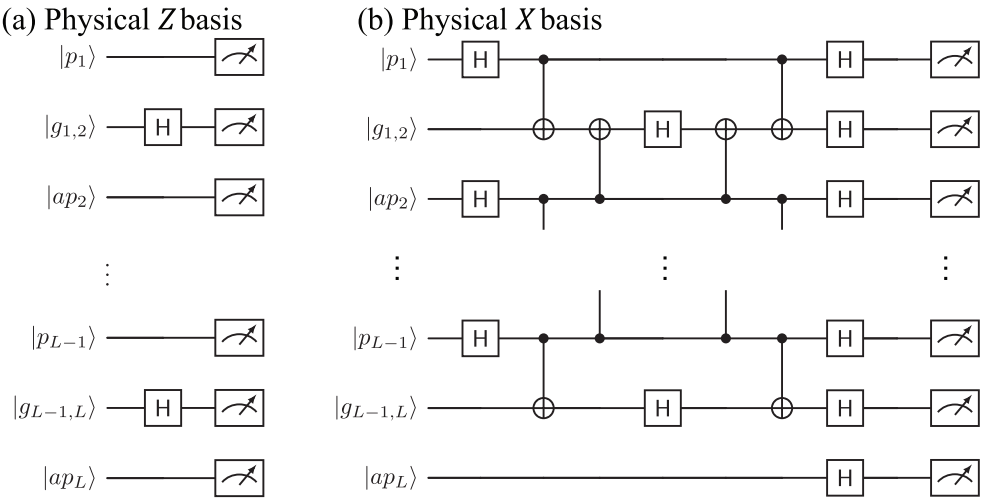

FIG. 4. Quantum circuits for implementing the MUPB in the $\mathbb{Z}_2$ LGT setup introduced in Sec. II. The left and right circuits correspond to measurements in the physical Z basis and the physical X basis, respectively. Each circuit shows the measurement setup for the local degrees of freedom, including the gauge-link field $|g_{n-1,n}\rangle$ and the matter fields $|p_n\rangle$ and $|ap_n\rangle$. The right circuit implements the transformation to the complementary physical basis using CNOT and Hadamard gates.

practical implementation of ITE on quantum hardware, we combine our measurement scheme with the QITE algorithm proposed by Motta *et al.* and heuristically assess its performance.

A. Concrete MUPB construction in the present setup

We explicitly construct a pair of MUPBs for the $\mathbb{Z}_2$ LGT setup in Sec. II. Here, using the same stabilizer-based framework as in Sec. IV, we identify a geometry-adapted realization that satisfies the same MUPB conditions but admits a substantially simpler nearest-neighbor, constant-depth implementation, thereby retaining efficient scaling to larger systems, shown in Fig. 4. The numerical simulations below use this specialized realization. We refer to these bases as the physical Z basis (left in Fig. 4) and the physical X basis (right in Fig. 4), and denote their basis states by $\{|e_i^{(Z_{\rm phys})}\rangle\}$ and $\{|e_i^{(X_{\rm phys})}\rangle\}$, respectively. The circuit for the physical Z basis consists of Hadamard gates acting on the gauge links. The physical X basis is obtained by supplementing these Hadamard gates with a unitary operator U, defined by

$$U := V^\dagger \mathsf{H}_{L_{\rm KS}} \times \left(\prod_{m=1}^{L_{\rm KS}-1} \sigma^{\mathsf{H}}_{m,m+1}\right) V, \tag{15}$$

$$V := \left(\prod_{m=1}^{L_{\rm KS}-2} \mathrm{CNOT}_{f_{m+1},g_{m,m+1}}\right) \times \left(\prod_{m=1}^{L_{\rm KS}-2} \mathrm{CNOT}_{f_m,g_{m,m+1}}\right) \times \left(\prod_{m=1}^{L_{\rm KS}-1} \mathsf{H}_m\right), \tag{16}$$

where $\mathrm{CNOT}_{f_p,g_{q,q+1}}$ denotes a CNOT gate with control qubit corresponding to pth fermionic site f_p and target qubit corresponding to $(q, q+1)$ gauge link $g_{q,q+1}$. H_p and $\sigma^{\mathsf{H}}_{p,p+1}$ denote Hadamard gates acting on site p and link $(p, p+1)$, respectively. The operator U plays a central role in generating mixing within the physical subspace while preserving the local gauge symmetries. A detailed proof that the physical Z and X bases satisfy the MUPB conditions is given in Appendix C.

B. Simulation and measurement setup

First, we describe the simulation parameters used in our calculations. In our simulations, all parameters are rescaled by the gauge coupling g, with $ag = 0.25$ and $m/g = 0.05$. All numerical simulations were performed using QISKIT.

Next, we describe the measurement setup used in the following numerical simulations. We focus on three observables: the energy density $\langle h\rangle_{\beta,\mu} := \langle H\rangle_{\beta,\mu}/L_D$, the chiral condensate $\langle\bar\psi\psi\rangle_{\beta,\mu}$, and the fermion number density $\langle n\rangle_{\beta,\mu} := \langle N\rangle_{\beta,\mu}/L_D$. The chiral condensate is represented in the qubit basis as

$$\bar\psi\psi := \frac{1}{L_D}\sum_{n=1}^{L_{\rm KS}}(-1)^{n+1}\chi_n^\dagger\chi_n = \frac{1}{L_{\rm KS}}\sum_{n=1}^{L_{\rm KS}}(-1)^{n+1}Z_n.$$

It serves as a useful indicator of the finite-density transition in this model. The fermion number density measures the fermion number relative to the staggered vacuum.

To measure the energy density, we decompose the Hamiltonian into three parts

$$H_X = -\frac{1}{4a}\sum_{n=1}^{L_{\rm KS}-1} X_n\sigma^Z_{n,n+1}X_{n+1},$$

$$H_Y = -\frac{1}{4a}\sum_{n=1}^{L_{\rm KS}-1} Y_n\sigma^Z_{n,n+1}Y_{n+1},$$

$$H_D = ag^2\sum_{n=1}^{L_{\rm KS}-1}\left(1-\sigma^X_{n,n+1}\right) + \frac{m}{2}\sum_{n=1}^{L_{\rm KS}}(-1)^{n+1}Z_n,$$

so that $H = H_X + H_Y + H_D$. The Pauli terms within each of H_A $(A = X, Y, D)$ are qubitwise commuting and can therefore be measured using a single measurement setting, respectively. With this grouping, we measure the energy using $M = 3$

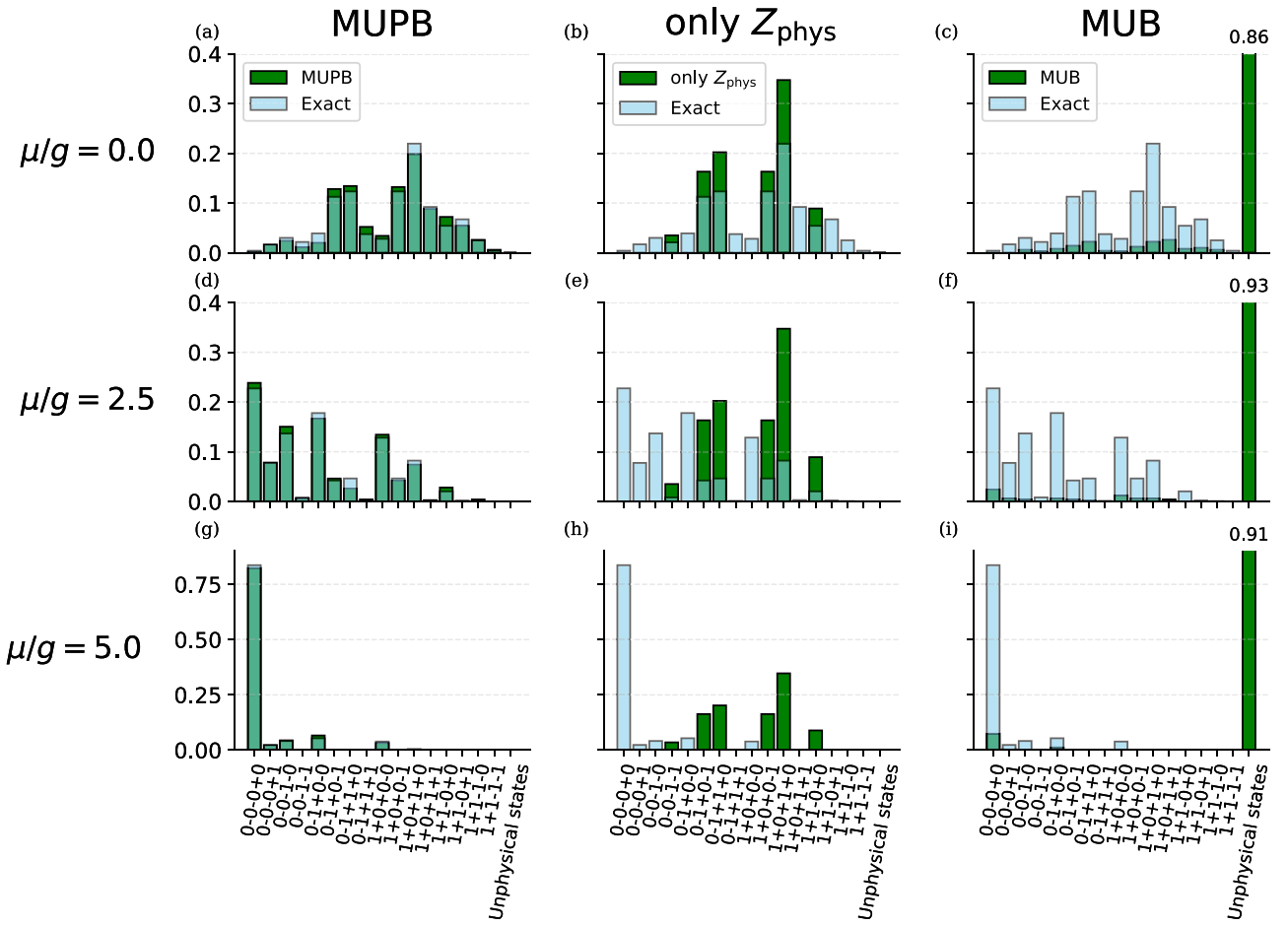

FIG. 5. Sampling distributions of collapse states obtained from projective measurements during QMETTS sampling (green) and the exact probabilities (light blue) for $L_{\text{KS}} = 4$. The ITE is implemented exactly using classical computation. The panels are arranged by chemical potential (rows) and measurement scheme (columns). The top panels (a)–(c), middle panels (d)–(f), and bottom panels (g)–(i) correspond to $\mu/g = 0.0$, 2.5, and 5.0, respectively, while the left, middle, and right columns correspond to MUPB, only Z_{phys}, and an MUB pair consisting of the physical Z basis and its mutually unbiased counterpart, respectively. The inverse temperature is fixed at $\beta g = 1.0$. For the MUPB and MUB runs, we plot only the collapse states obtained at the physical Z-basis measurement steps of the alternating Markov chain. All gauge-violating collapse states are grouped into "unphysical states." The exact probabilities are calculated as $\text{Prob}_i = \langle i^{(Z_{\text{phys}})}|e^{-\beta(H-\mu N)}|i^{(Z_{\text{phys}})}\rangle/Z$. For all panels, the initial state is fixed to $|1+0+1+0\rangle$, which corresponds to the trivial no-excitation state in the qubit description introduced in Sec. II. All QMETTS results are obtained with $N_{\text{chain}} = 1000$.

measurement groups. In contrast, the chiral condensate and the fermion number density contain only Z operators, and we measure each of them using a single measurement setting, i.e., $M = 1$.

C. Numerical results with exact ITE

In this section, we present the numerical results obtained using our proposed algorithm. To isolate the performance of our sampling scheme from errors in ITE, we first implement the ITE exactly by classical computation, where the imaginary-time evolution operator $e^{-\beta(H-\mu N)/2}$ is applied directly without any approximations. We first consider QMETTS with three different measurement schemes: the constructed MUPB, only the physical Z basis, and a pair of mutually unbiased bases, one of which is the physical Z basis. The additional basis can be constructed by applying the quantum circuit $\prod_{m=1}^{L_{\text{KS}}} \mathsf{H}_m$, which applies Hadamard gates to the fermionic matter qubits, before measurement in the computational basis. Together with the physical Z basis, this additional basis forms a pair of mutually unbiased bases, i.e., satisfying Eq. (8). Unlike the MUPB, however, the pair is not restricted to the physical subspace. We refer to this measurement scheme as the MUB scheme below.

We compare these three measurement schemes by examining the sampling distributions of the collapse states for a system size $L_{\text{KS}} = 4$, corresponding to $N_q = 7$ qubits. We fix the Markov-chain length at $N_{\text{chain}} = 1000$ and choose the initial state $|1+0+1+0\rangle$ in the qubit representation introduced in Sec. II, which corresponds to the trivial configuration without excitations. Figure 5 shows the results obtained using the MUPB scheme (left column), only the physical Z basis (middle column), and the MUB scheme (right column), for three representative chemical potentials, $\mu/g = 0.0$, 2.5, and 5.0, with $\beta g = 1.0$. For the MUPB and MUB schemes, we show only the collapse states obtained at the physical Z-basis measurement steps of the alternating Markov chains, so that they can be directly compared with the exact stationary distribution in the physical Z basis, $\text{Prob}_i =$

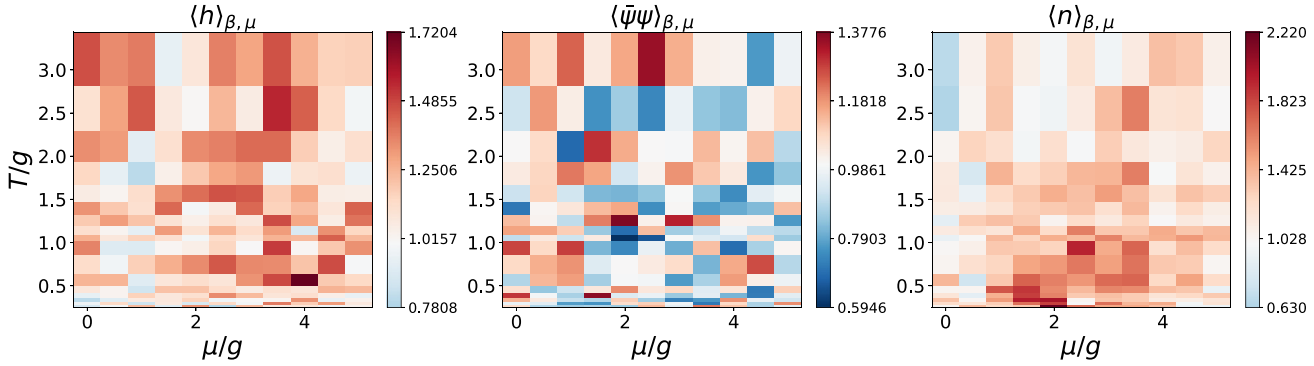

FIG. 6. Integrated autocorrelation times $\tau(1)$ for the energy density $\langle h \rangle_{\beta,\mu}$ (left), chiral condensate $\langle \bar{\psi}\psi \rangle_{\beta,\mu}$(middle), and fermion number density $\langle n \rangle_{\beta,\mu}$ (right), shown over the $(T/g, \mu/g)$ parameter space. The system size is $L_{\mathrm{KS}} = 8$. The autocorrelation sum is truncated using the windowing procedure. The window parameter $w = 3$. $\tau(1) = 1$ corresponds to independent sampling.

$\langle i^{(Z_{\mathrm{phys}})}|e^{-\beta(H-\mu N)}|i^{(Z_{\mathrm{phys}})}\rangle/Z$. In Fig. 5, all gauge-violating states are grouped into "unphysical states."

First, the MUB scheme fails to reproduce the exact stationary distribution. In particular, it predominantly samples gauge-violating states because the states in the additional measurement basis are not eigenstates of the Gauss's law operators. On the other hand, the MUPB scheme and the scheme using only the physical Z basis restrict the sampling to the physical sector because their measurement bases are compatible with gauge invariance. However, the approach using only the physical Z basis exhibits significant deviations from the exact stationary distribution. This indicates that using only the physical Z basis does not ensure ergodicity. This lack of mixing can be understood from the conservation of the number operator N. The initial state $|1+0+1+0\rangle$ belongs to the $N = 0$ sector, and the imaginary-time evolution preserves this sector because $[N, e^{-\beta(H-\mu N)}] = 0$. Moreover, measurements in the physical Z basis collapse the state onto eigenstates of N. Thus, when only the physical Z basis is used, the Markov chain remains trapped in the fixed-N sector in the present setup. In contrast, the MUPB-based QMETTS accurately reproduces the exact stationary distribution, consistent with the improved mixing expected from the MUPB.

In Appendix E, we further tested several distinct initial states using the MUPBs and found that the sampled collapse-state distributions after $N_{\mathrm{chain}} = 1000$ steps are all consistent with the exact stationary distribution within finite-sampling fluctuations. This observation indicates that, in the present finite-size benchmark with the MUPB sampling scheme, the effect of burn-in is not practically significant at the chain lengths considered [42].

Next, we increase the system size to $L_{\mathrm{KS}} = 8$, corresponding to $N_q = 15$ qubits, and investigate the autocorrelations in the Markov chains generated using the MUPB scheme. For the three observables specified above (the energy density, the chiral condensate, and the fermion number density), we estimate the integrated autocorrelation time $\tau(1)$. For each observable, we use $N_{\mathrm{chain}} = 1000$ METTS samples and one shot per measurement group. The number of measurement groups depends on the observable: $M = 3$ for the Hamiltonian, whereas $M = 1$ for the chiral condensate and the fermion number density. Therefore, the corresponding total numbers of circuit executions are not identical among the observables, although the autocorrelation analysis is performed using Markov chains of the same length. In practice, the finite-chain estimate of the autocorrelation sum in Eq. (11) is susceptible to statistical fluctuations, particularly at large time lags. To suppress these influences, we introduce a window parameter w and truncate the sum following the standard windowing procedure [36]. The integrated autocorrelation time is then estimated as $\tau(1) = 1 + 2\sum_{t=1}^{w} \rho_O(t, 1)$. We choose the window parameter $w = 3$ based on the analysis of the autocorrelation functions by investigating the time where statistically uncorrelated fluctuations become dominant at representative parameters. See Appendix F for details.

Figure 6 shows the resulting integrated autocorrelation times. As shown in Fig. 6, the autocorrelation behavior is observable dependent: The chiral condensate exhibits negative autocorrelations for many parameter choices, whereas the fermion number density tends to show positive autocorrelations as the temperature decreases. Despite these differences, the autocorrelation times remain moderate throughout the parameter range studied here. These results suggest that MUPB, combined with the single-shot measurement strategy, achieves efficient mixing in the present finite-size benchmark. In Appendix G, we further investigate the system-size dependence of the integrated autocorrelation time for $L_{\mathrm{KS}} = 2, 4, 6$, and 8. Within this range of system sizes, we observe no clear systematic system-size dependence.

Based on these observations, we examine the β dependence of the energy density $\langle h \rangle_{\beta,\mu}$ shown in Fig. 7 for $L_{\mathrm{KS}} = 8$. To illustrate the practical advantage of the single-shot strategy under a fixed total number of circuit executions, we compare $N_{\mathrm{shot}} = 1$ and 33 while keeping the total number of circuit executions fixed at $N_{\mathrm{tot}} = 4000$. The corresponding chain lengths are $N_{\mathrm{chain}} = 1000$ for $N_{\mathrm{shot}} = 1$ and $N_{\mathrm{chain}} = 40$ for $N_{\mathrm{shot}} = 33$ by Eq. (9) since $M = 3$ for the Hamiltonian. For both shot settings, the numerical results fluctuate around the exact-diagonalization result shown by the black solid line. While the difference between the two approaches is small in the low-temperature regime, the $N_{\mathrm{shot}} = 1$ strategy yields smaller error bars in the high-temperature regime. At high temperature, more states contribute appreciably, which increases the METTS-sampling variance σ_μ^2. According to the estimate in Eq. (12), an increase in σ_μ^2 drives the optimal num-

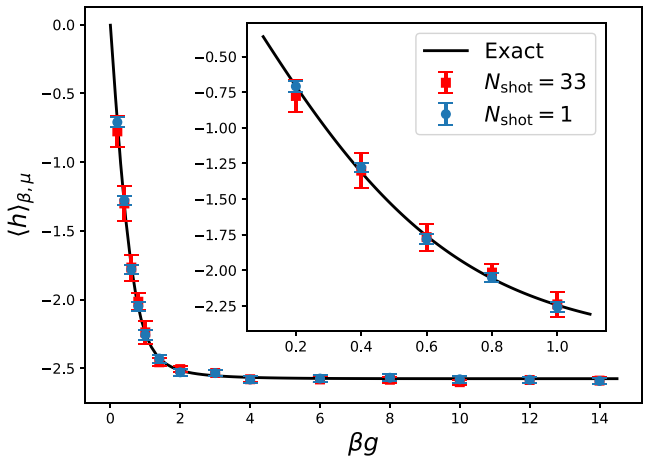

FIG. 7. Temperature dependence of the energy density $\langle h \rangle_{\beta,\mu}$ as a function of inverse temperature βg for $L_{\text{KS}} = 8$ at zero chemical potential ($\mu/g = 0$). The QMETTS results for $N_{\text{shot}} = 1$ (blue dots with error bars) and $N_{\text{shot}} = 33$ (red dots with error bars) are compared with the exact-diagonalization result (black solid line). In both settings, the total number of circuit executions is fixed at $N_{\text{tot}} = 4000$, so the corresponding chain lengths are $N_{\text{chain}} = 1000$ and 40, respectively. The inset highlights the high-temperature region $\beta g \in [0, 1.0]$.

ber of shots toward smaller values. Therefore, the behavior observed in the high-temperature regime of Fig. 7 is consistent with our analytical prediction that the single-shot strategy becomes favorable in this regime. Figure 8 shows the μ dependence of the chiral condensate $\langle \bar{\psi}\psi \rangle_{\beta,\mu}$ and the fermion number density $\langle n \rangle_{\beta,\mu}$ at three different temperatures for $L_{\text{KS}} = 8$ [43]. We compare the results obtained with $N_{\text{shot}} = 1$ and $N_{\text{shot}} = 49$ under the fixed total circuit budget $N_{\text{tot}} = 2000$. The corresponding chain lengths are $N_{\text{chain}} = 1000$ for $N_{\text{shot}} = 1$ and $N_{\text{chain}} = 40$ for $N_{\text{shot}} = 49$ because $M = 1$ for chiral condensate and fermion number density. Overall, both shot settings reproduce the exact-diagonalization results (black curves) within statistical fluctuations, and the single-shot strategy remains competitive across the parameter range studied. This trend is particularly clear at the highest temperatures, $\beta g = 0.4$, where the $N_{\text{shot}} = 1$ results perform comparably to, or better than, the $N_{\text{shot}} = 49$ results, as in the energy-density case. On the other hand, in the low-temperature regime $\beta g = 14.0$, the $N_{\text{shot}} = 49$ results sometimes reproduce the exact values more accurately in the small-μ region, $\mu/g \lesssim 1.5$, especially for the chiral condensate. This trend is also qualitatively consistent with the general shot-allocation analysis in Sec. V, which shows that a larger N_{shot} can become favorable when the shot noise is relatively more important than the METTS-sampling fluctuations, as the latter decrease at lower temperatures.

D. Numerical results with QITE

In the previous subsection, we investigated the performance of our algorithm using exact ITE. On a quantum computer, however, nonunitary ITE cannot be implemented directly without additional computational overhead. Here, we approximate the imaginary-time evolution using the QITE algorithm proposed by Motta *et al.* [12] and heuristically assess the performance of our algorithm when combined with QITE. We first briefly review the QITE algorithm.

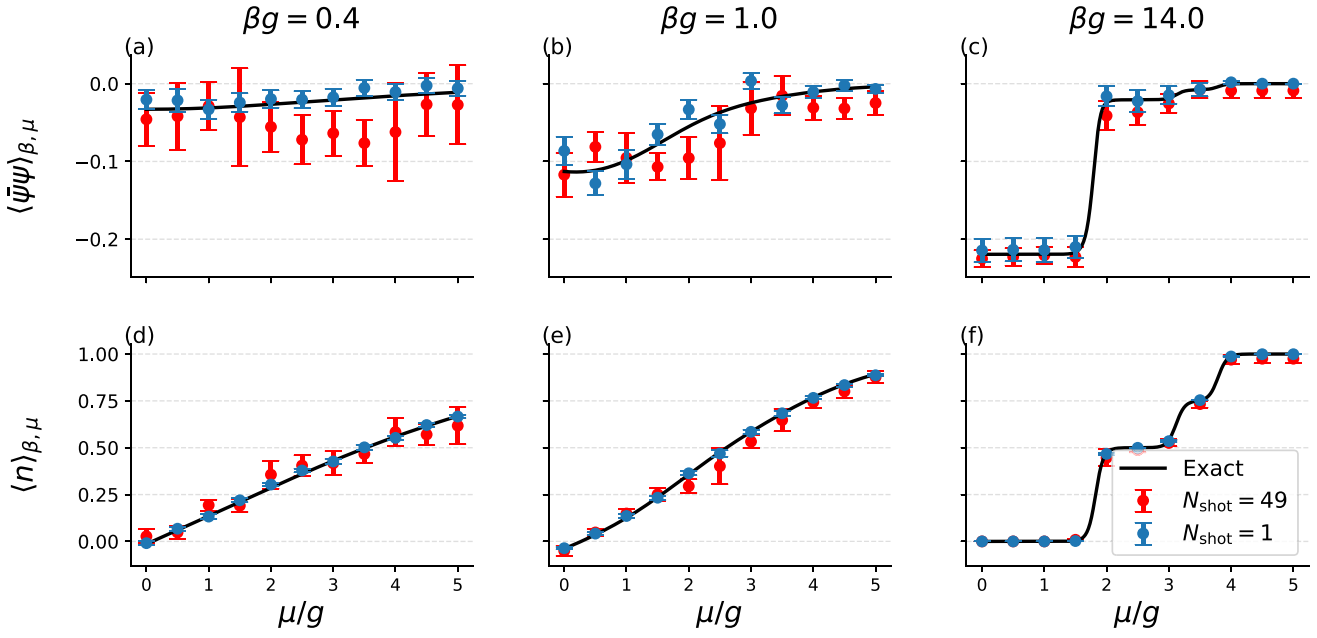

FIG. 8. Dependence of chiral condensate $\langle \bar{\psi}\psi \rangle_{\beta,\mu}$ and fermion number density $\langle n \rangle_{\beta,\mu}$ on the chemical potential μ/g for $L_{\text{KS}} = 8$. Top panels (a)–(c) show the chiral condensate, while bottom panels (d)–(f) show the fermion number density. From left to right, the columns correspond to inverse temperatures $\beta g = 0.4$, 1.0, and 14.0, respectively. Both observables are measured by a single measurement group setting, $M = 1$. The QMETTS results for $N_{\text{shot}} = 1$ (blue dots with error bars) and $N_{\text{shot}} = 49$ (red dots with error bars) are compared with the exact-diagonalization result (black solid line). In both settings, the total number of circuit executions is fixed at $N_{\text{tot}} = 2000$, so the corresponding chain lengths are $N_{\text{chain}} = 1000$ and 40, respectively.

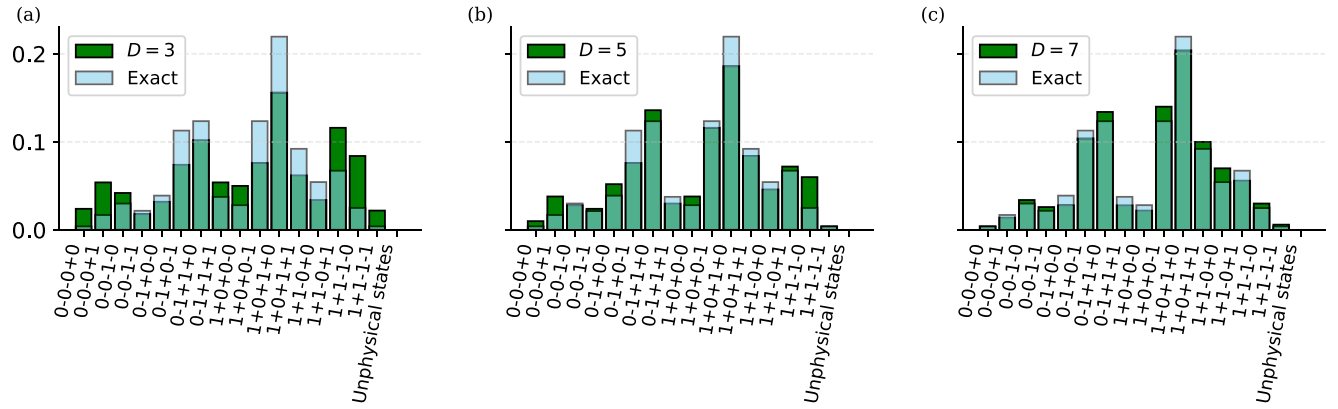

FIG. 9. Sampling distributions of collapse states obtained from projective measurements during QMETTS sampling using the MUPB scheme (green), together with the exact probabilities (light blue), for $L_{\mathrm{KS}} = 4$ with QITE. We plot only the collapse states obtained at the physical Z-basis measurement steps of the alternating Markov chain. All gauge-violating collapse states are grouped into "unphysical states." The exact probabilities are calculated as $\mathrm{Prob}_i = \langle i^{(Z_{\mathrm{phys}})}|e^{-\beta(H-\mu N)}|i^{(Z_{\mathrm{phys}})}\rangle/Z$. The inverse temperature and chemical potential are fixed at $\beta g = 1.0$ and $\mu/g = 0.0$, respectively. The domain sizes are $D = 3$, 5, and 7 in panels (a)–(c), respectively. All QMETTS results are obtained with $N_{\mathrm{chain}} = 1000$. For all panels, the initial state is fixed to $|1+0+1+0\rangle$, which corresponds to the trivial no-excitation state in the qubit description introduced in Sec. II.

In QITE, after Trotter decomposition of ITE, the action of each local imaginary-time evolution operator is approximated by a unitary transformation as

$$\frac{e^{-\Delta\tau h}|\Psi\rangle}{\|e^{-\Delta\tau h}|\Psi\rangle\|} \simeq e^{-i\Delta\tau\sum_I a_I\sigma_I}|\Psi\rangle, \tag{17}$$

where $|\Psi\rangle$ denotes the state before the imaginary-time step. The sum over I runs over the set $\mathcal{P}_D$ of Pauli strings acting on a domain of D qubits surrounding the support of h. The coefficients a_I are determined by optimizing the norm $\|(e^{-i\Delta\tau\sum_I a_I\sigma_I} - \mathcal{N}e^{-\Delta\tau h})|\Psi\rangle\|$ on classical computers, where $\mathcal{N} = \|e^{-\Delta\tau h}|\Psi\rangle\|^{-1}$. Since the Pauli basis on a D-qubit domain contains $|\mathcal{P}_D| = 4^D$ operators, the computational costs increase exponentially with D. For further details, see Ref. [12].

For larger systems, one therefore generally needs to restrict the domain size to fit the available computational resources. This finite-domain approximation introduces a trade-off between the accuracy of the imaginary-time evolution and its computational cost. The effectiveness of the domain truncation depends on the correlations between the degrees of freedom inside and outside the chosen domain. In general quantum many-body systems with local Hamiltonians, correlations of thermal states tend to remain short ranged in the high-temperature regime [44], which motivates the expectation that finite-domain approximations of the ITE operator can be more effective in this regime. However, if the gauge degrees of freedom are eliminated by explicitly solving the Gauss-law constraints, the resulting Hamiltonian may become nonlocal, which can reduce the effectiveness of such approximations. In contrast, our formulation retains the gauge degrees of freedom and preserves the locality of the Hamiltonian. Moreover, the geometry-adapted MUPB construction introduced above can be implemented using only local gates. These features avoid introducing additional nonlocality into either the Hamiltonian or the collapse states and can therefore be expected to be favorable for QITE with truncated domain sizes.

To examine this expectation numerically, we investigate the dependence on the QITE domain size D for $L_{\mathrm{KS}} = 4$. We choose this system size because the full-domain case, $D = N_q = 7$, remains computationally tractable, allowing us to directly compare truncated-domain calculations with the full-domain result. Further details of the QITE implementation in our setup are provided in Appendix D. We first examine how the domain truncation affects the sampling distribution. In Fig. 9, we plot the sampling distributions of the collapse states at $\mu/g = 0.0$ and $\beta g = 1.0$ using the constructed MUPB for domain sizes $D = 3$, 5, and 7. As shown in Fig. 9, the QMETTS sampling reproduces the exact stationary distribution more accurately as the domain size increases. Nevertheless, even for $D = 3$, we observe no gauge-violating collapse states. This result suggests that, within the present setup and numerical precision, introducing a finite-domain cutoff in QITE does not cause appreciable leakage from the physical sector.

Next, in Fig. 10, we compute the energy density at $\mu/g = 0.0$ by changing domain sizes from $D = 3$ to $D = 7$. The results obtained by exact diagonalization are shown by the black curve. As shown in Fig. 10, the full-domain calculation with $D = 7$ (green dots) accurately reproduces the exact result, whereas the result for $D = 3$ (blue dots) deviates substantially from it. For $D = 5$ (orange dots), the QITE result follows the exact behavior in the high-temperature regime, $\beta g \lesssim 1.0$, as shown in the inset. These results indicate that, for this model, domain truncation can retain reasonable accuracy in the high-temperature regime while reducing the computational cost as we discussed. In addition, although deviations appear at intermediate temperatures, the result approaches the exact value again as βg becomes large. This behavior can be interpreted as an effective delay of the cooling process induced by the finite-domain approximation, with the truncated-QITE result eventually approaching the exact low-temperature value at large β.

Finally, we turn to the finite-temperature and finite-density behavior of the model. To separate these physical features

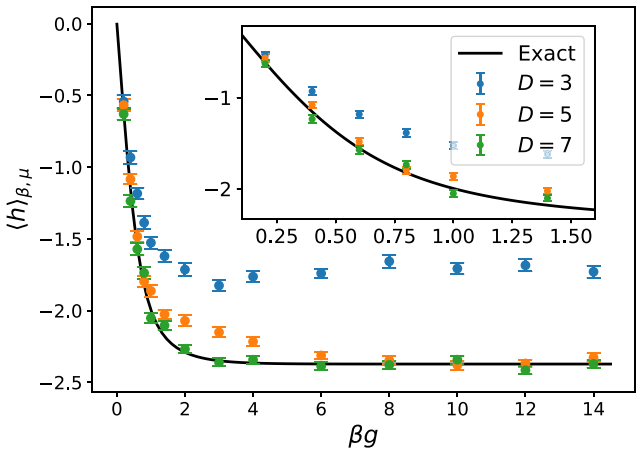

FIG. 10. Temperature dependence of the energy density $\langle h\rangle_{\beta,\mu}$ obtained using QMETTS with MUPB and the single-shot strategy combined with QITE for $L_{\mathrm{KS}} = 4$ at zero chemical potential ($\mu/g = 0$). The results for domain sizes $D = 3, 5$, and 7 are represented by blue, orange, and green dots with error bars, respectively, and are compared with the exact-diagonalization result (black solid line). The energy-density measurement requires $M = 3$ measurement groups. For all domain sizes, the total number of circuit executions is fixed at $N_{\mathrm{tot}} = 4000$, corresponding to a Markov-chain length of $N_{\mathrm{chain}} = 1000$. The inset highlights the high-temperature region $\beta g \in [0, 1.6]$.

from domain-truncation effects, we use the full-domain calculation with $D = 7$. Figure 11 summarizes the behavior of the chiral condensate and the fermion number density over the $(\mu/g, T/g)$ plane. The behavior of the chiral condensate is consistent with the expected tendency toward chiral-symmetry restoration as the temperature or chemical potential increases. The fermion number density further indicates that the high-density region appears in the same parameter regime where the chiral condensate is suppressed. These results demonstrate that the proposed method captures the expected finite-temperature and finite-density structure of the model within a Hamiltonian formalism that does not rely on Euclidean Monte Carlo sampling.

VII. CONCLUSION

In this work, we have developed a $\mathbb{Z}_2$ gauge-invariant framework for computing finite-temperature and finite-density expectation values by introducing MUPB in the QMETTS algorithm. Projective measurements within the MUPB ensure that the system never collapses into gauge-violating states and maintain ergodicity within the physical subspace. By leveraging the mathematical equivalence between $\mathbb{Z}_2$ LGT and the stabilizer formalism, we have identified an efficient construction for MUPBs using shallow circuits. This construction is applicable across arbitrary dimensions and boundary conditions. Furthermore, we analyzed the effect of shot noise on expectation-value estimation in QMETTS. We showed that the finite-shot QMETTS estimator remains unbiased for finite-temperature and finite-density expectation values, and that, under a fixed total number of circuit executions, the single-shot strategy provides a robust near-optimal choice. Its variance is at most a factor of two above the optimum, while requiring no prior knowledge of the relevant variances or autocorrelation times. To validate our methodology, we applied the proposed approach to a $(1 + 1)$-dimensional $\mathbb{Z}_2$ LGT with fermionic matter. First, using exact ITE, we investigated the performance of our approach and numerically demonstrated that it accurately reproduces thermal properties over a wide range of temperatures and chemical potentials. We then combined our framework with QITE to investigate its performance in a more practical implementation of imaginary-time evolution. By varying the QITE domain size, we found that domain truncation can reproduce the exact results reasonably well in the high-temperature regime while reducing the computational cost associated with the ITE step. While these results demonstrate the potential usefulness of domain truncation in the present finite-size benchmark, they do not determine how the domain size required for a given accuracy scales with the system size. Establishing this scaling for larger systems remains an important direction for future work.

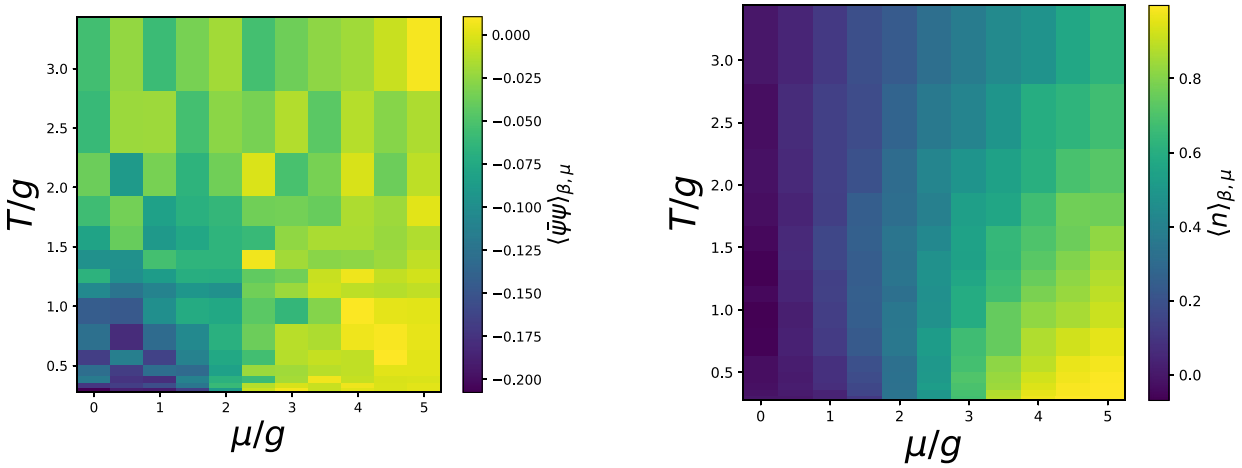

FIG. 11. Finite-size phase diagrams in the $(\mu/g, T/g)$ plane for the chiral condensate $\langle\bar{\psi}\psi\rangle_{\beta,\mu}$ (left) and the fermion number density $\langle n\rangle_{\beta,\mu}$ (right). The system size is $L_{\mathrm{KS}} = 4$. The QMETTS results are obtained with $N_{\mathrm{chain}} = 1000$ using the single-shot strategy. Both observables require a single measurement setting, $M = 1$. The color scales indicate the expectation values estimated by the proposed method. The horizontal axis represents the chemical potential μ/g, and the vertical axis represents the temperature T/g.

Our stabilizer-based construction may also be applicable to a broader class of gauge theories that admit stabilizer formulations. First, $\mathbb{Z}_d$ lattice gauge theories with prime d possess an analogous stabilizer structure [45–47]. For $\mathbb{Z}_d$ lattice gauge theories with d equal to a power of 2, it has also been shown that the theory can be reduced to a stabilizer formulation [48]. For continuous gauge groups, certain truncated formulations of U(1) and SU(2) lattice gauge theories can likewise be mapped to qubit- or qudit-based stabilizer descriptions [20,49–51], although the applicability depends on the truncation scheme and boundary conditions. Importantly, our MUPB construction relies on the stabilizer formalism rather than on properties specific to $\mathbb{Z}_2$ gauge theories. Therefore, whenever a gauge theory admits a compatible stabilizer formulation, our construction can be extended to that setting straightforwardly without modifying its basic structure. Extending our algorithm to more general continuous gauge groups remains a promising direction for future work.

Our main contribution is independent of the particular implementation of ITE, although the practical efficiency of the overall framework depends on the underlying ITE scheme. In the present work, we employ the QITE algorithm proposed by Motta *et al.* and heuristically investigate its performance by varying the domain size. As the domain size increases, the accuracy improves, while the computational cost grows exponentially. Recent studies have shown, however, that exploiting gauge symmetries can substantially reduce the cost of the ITE component in gauge-theory simulations [52]. In addition, combining our sampling framework with alternative ITE schemes, such as probabilistic implementations of ITE [53–55], adiabatic QITE [56], and double-bracket QITE [57], provides an important direction for future work. It would also be interesting to explore connections with sampling-based finite-temperature algorithms that use series expansions of imaginary-time evolution [58].

Finally, our proposed framework is naturally compatible with emerging techniques that leverage gauge symmetry for symmetry verification [18–22]. In such error-mitigation schemes, quantum errors are detected by monitoring violations of gauge constraints, a process that inherently relies on the redundancy of the Hilbert space. Unlike standard approaches that may trace out or eliminate these degrees of freedom, our method explicitly preserves this redundancy, allowing for straightforward integration with symmetry-based error suppression. By combining our MUPB-based QMETTS with these verification protocols, our approach provides a promising route toward more robust quantum simulations of lattice gauge theories.

ACKNOWLEDGMENTS

We thank Etsuko Itou, Koji Terashi, and Lento Nagano for fruitful discussions and/or comments. We thank Yutaro Iiyama for assistance with numerical calculations. This work was supported by JST SPRING, Grant No. JPMJSP2108.

DATA AVAILABILITY

There are no publicly available research data or software supporting this manuscript. Requests for further information or data should be sent to the authors.

APPENDIX A: EFFICIENT CONSTRUCTION OF MUPB FOR $\mathbb{Z}_2$ LGTs IN GENERAL DIMENSIONS AND WITH ARBITRARY BOUNDARY CONDITIONS

In this appendix, we provide a formal proof that the quantum circuits illustrated in Fig. 3 satisfy the conditions for MUPB defined through Definition 1. Our proof leverages the stabilizer formalism, which is particularly well suited for describing subspaces defined by symmetry constraints, such as gauge invariance. First, we provide a brief review of the stabilizer formalism. Subsequently, we apply this framework to the $\mathbb{Z}_2$ lattice gauge theory to rigorously prove our statement.

1. Stabilizer formalism

Let G be a group. If a subset $H \subseteq G$ can generate all elements of G through the multiplication of its elements, we refer to H as a generating set of G and denote the group as $G = \langle H \rangle$. The elements of H are called the generators of G. The N_q-qubit Pauli group is defined as $\mathcal{P}_{N_q} := \{\pm 1, \pm i\} \otimes \{I, X, Y, Z\}^{\otimes N_q}$. Stabilizer groups, operators, and states are defined using the Pauli group as follows.

Definition A1 (Stabilizer groups, operators, and states). A subgroup $\mathcal{S}_{N_q} \subseteq \mathcal{P}_{N_q}$ is called a stabilizer group if it is Abelian (all elements commute) and does not contain the element $-I$. The elements $S \in \mathcal{S}_{N_q}$ are referred to as stabilizer operators. A quantum state $|\psi\rangle$ is called a stabilizer state of $\mathcal{S}_{N_q}$ if it satisfies

$$S_i|\psi\rangle = +|\psi\rangle, \quad \text{for } \forall S_i \in \mathcal{S}_{N_q}.$$

Note that all stabilizer operators are Hermitian and the eigenvalues must be ± 1 since $-I \notin \mathcal{S}_{N_q}$. The framework for describing quantum states and subspaces using these groups is known as the stabilizer formalism. The theory of quantum error correction is mainly based on this formalism; logical qubits are embedded into the space spanned by stabilizer states called code space, and errors can be detected by measuring the eigenvalues of quantum states.

We also define the N_q-qubit Clifford group that plays an important role in describing stabilizer formalism. N_q-qubit Clifford group $\mathcal{C}_{N_q}$ is defined as the normalizer of the Pauli group,

$$\mathcal{C}_{N_q} := \{C \in \mathcal{U}(\mathbb{C}^2)^{\otimes N_q} | P \in \mathcal{P}_{N_q}, CPC^\dagger \in \mathcal{P}_{N_q}\}.$$

The Clifford group consists of Hadamard gates, phase gates, and CNOT gates. We now summarize some key properties of the stabilizer formalism to provide our proof.

Theorem A1 ([41]). Let the system size be N_q and the number of generators of a stabilizer group $\mathcal{S}_{N_q}$ be S. The dimension of the space spanned by the stabilizer states is given by

$$2^{N_q}/2^S = 2^{N_q - S}.$$

Proof. For any nontrivial stabilizer operator $S_i \in \mathcal{S}_{N_q} \setminus \{I\}$, its eigenvalues are ± 1. Since $\mathrm{Tr}[S_i] = 0$ for any Pauli string other than the identity, the multiplicities of the $+1$ and -1 eigenvalues must be equal. Consequently, the projection onto the $+1$ eigenspace of a single generator S_i bisects the Hilbert space, reducing its dimension by half. Given S independent

and commuting generators, each additional generator further constrains the space by a factor of 1/2. Therefore, the total dimension of the simultaneous +1 eigenspace is 2^{N_q-S}. ■

Next, we discuss the inner product of two stabilizer states $|\phi\rangle$ and $|\psi\rangle$ with different stabilizer groups.

Theorem A2 ([41]). Let $|\psi\rangle$ and $|\phi\rangle$ be stabilizer states in an N_q-qubit system, uniquely determined (up to a global phase) by their respective stabilizer groups $\mathcal{S}(|\psi\rangle) = \langle\{P_1, \ldots, P_{N_q}\}\rangle$ and $\mathcal{S}(|\phi\rangle) = \langle\{Q_1, \ldots, Q_{N_q}\}\rangle$. The magnitude of their inner product $|\langle\phi|\psi\rangle|$ is determined as follows:

(1) If there exists an operator $P \in \mathcal{P}_{N_q}$ such that $P \in \mathcal{S}(|\psi\rangle)$ and $-P \in \mathcal{S}(|\phi\rangle)$, the states are orthogonal: $|\langle\phi|\psi\rangle| = 0$.

(2) Let S be the minimum, over all sets of generators $\{P_1, \ldots, P_{N_q}\}$ for $\mathcal{S}(|\psi\rangle) = \langle\{P_1, \ldots, P_{N_q}\}\rangle$ and $\{Q_1, \ldots, Q_{N_q}\}$ for $\mathcal{S}(|\phi\rangle) = \langle\{Q_1, \ldots, Q_{N_q}\}\rangle$, of the number of different i values for which $P_i \neq Q_i$. Then, $|\langle\phi|\psi\rangle| = 2^{-S/2}$.

Note that the number of different generators cannot be uniquely determined and we must maximize the number of the same generators in step 2.

Proof. First, we prove the first step. If there exists P such that $P|\psi\rangle = |\psi\rangle$ and $-P|\phi\rangle = |\phi\rangle$, $|\psi\rangle$ and $|\phi\rangle$ are in different eigenspaces. Thus, we obtain $\langle\phi|\psi\rangle = 0$.

Next, we prove the second step. By Theorem A3, there always exists an element of the Clifford group W such that $|\psi\rangle$ transforms into canonical form $W|\psi\rangle = |0\rangle^{\otimes N_q}$, indicating that $W|\psi\rangle$ is stabilized by $\langle\{Z_1, \ldots, Z_{N_q}\}\rangle$. Now, choose generators of $\mathcal{S}(W|\phi\rangle)$ so that the number of generators shared with $\mathcal{S}(W|\psi\rangle)$ is maximized. Let R be the number of such common independent generators. By definition, the minimum number of different generators is therefore $S = N_q - R$. Because $|\psi\rangle$ and $|\phi\rangle$ are assumed to be nonorthogonal, $\mathcal{S}(W|\phi\rangle)$ cannot contain any operator of the form $-P$ with $P \in \mathcal{S}(W|\psi\rangle)$; otherwise, $W|\psi\rangle$ and $W|\phi\rangle$ would belong to different eigenspaces of P and hence be orthogonal. The remaining $N_q - R$ generators necessarily contain at least one X or Y literal. Each one relates pairs of computational-basis amplitudes up to a phase and hence contributes a factor $1/\sqrt{2}$ to the overlap. Therefore, $|\langle\phi|\psi\rangle| = |\langle W\phi|W\psi\rangle| = |\langle W\phi|0^{\otimes N_q}\rangle| = 2^{-S/2}$. ■

Next, we introduce the binary expression of the stabilizer formalism to explain the canonical form of stabilizer groups [41]. Given N_q-qubit systems, if we have S generators, stabilizers are represented by $S \times 2N_q$ matrices, where all elements are 0 or 1, that is, $\mathbb{F}_2^{S\times 2N_q}$. The rows correspond to each stabilizer and the columns to each qubit. The former part of the matrices represents X and the latter Z. Global phases get dropped. For example, the five-qubit code [26]

$$XZZXI, IXZZX, XIXZZ, ZXIXZ$$

can be equivalently expressed in the following matrix up to global phases:

$$\left(\begin{array}{ccccc|ccccc} 1 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \end{array}\right).$$

In this expression, $XZZXI$ corresponds to the first row (10010 | 01100). We denote the left (right) space as X (Z) space. In this expression, Clifford gates (Hadamard, phase, and CNOT) can be interpreted as elementary column operations. Since $\mathsf{H}X\mathsf{H} = Z, \mathsf{H}Y\mathsf{H} = -Y$, Hadamard gates play a role in swapping corresponding two columns. $\mathsf{P}Z\mathsf{P}^\dagger = Z, \mathsf{P}X\mathsf{P}^\dagger = Y$ means that phase gates P add the columns of X to the columns of Z modulo 2. $\mathrm{CNOT}_{i,j}$, which controls ith qubit and acts X gate to jth qubit, is equivalent to $x_j \to x_j \oplus x_i$ and $z_i \to z_i \oplus z_j$ since $\mathrm{CNOT}_{i,j} X_i \mathrm{CNOT}_{i,j}^\dagger = X_i X_j$ and $\mathrm{CNOT}_{i,j} Z_j \mathrm{CNOT}_{i,j}^\dagger = Z_i Z_j$, where $x_i(z_i)$ is the ith column vector and $\oplus$ represents addition modulo 2. Since generators are independent, we can make the X-space block have rank S by using H gates. Due to the nature of $\mathbb{F}_2$, CNOT gates play a role in adding and swapping column vectors in X space; therefore, we can perform Gaussian elimination in X space by using CNOT gates. In addition, the commutativity condition imposes a constraint on the matrix. If the space is represented by $(A|B)$, the commutativity condition requires AB^T to be symmetric.

Based on this expression, we prove the following theorem, which is the modified version of Theorem 8 in the literature [41].

Theorem A3. Given the binary representation of S independent commuting stabilizer generators in $\mathbb{F}_2^{S\times 2N_q}$, we can transform them into the canonical form by the application of Clifford gates:

$$\left(\begin{array}{cc|cc} 0 & 0 & I & 0 \end{array}\right),$$

where $I \in \mathbb{F}_2^{S\times S}$.

To prove the above theorem, we show a lemma below.

Lemma A1 ([41]). Let k be an integer $k \in \mathbb{N}$. For any symmetric matrix $A \in \mathbb{Z}_2^{k\times k}$, there exists a diagonal matrix Λ such that $A + \Lambda = MM^T$, with M some invertible binary matrix.

Proof. Let M be a lower-triangular matrix with 1s all along the diagonal:

$$M_{ii} = 1,$$
$$M_{ij} = 0 (i\ lt; j).$$

Such an M is always invertible. If we rewrite the claim in terms of elements,

$$A_{ij} + \Lambda_{ij} = \sum_k M_{ik} M_{jk}, \tag{A1}$$

for all pairs (i, j). In the case $i = j$, it always holds by taking Λ appropriately. Thus, considering the case $i > j$ is enough since Eq. (A1) is symmetric under i and j. Hence, we only consider the case $i > j$ below.

Since M is a lower-triangular matrix, $M_{ik}M_{jk} = 0$ unless $k \leqslant j$, following

$$A_{ij} = \sum_{k<j} M_{ik} M_{jk} + M_{ij}. \tag{A2}$$

Based on those observations, we prove that we can determine all elements of M by mathematical induction about j.

For $j = 1$, since there is no k such that $k < j = 1$, Eq. (A2) indicates that $A_{i1} = M_{i1}$ for $\forall\ i\ (> 1)$. As $M_{11} = 1$, we can determine the elements of M_{i1} for $\forall i$. Next, we assume that $M_{ij'}$ have already been determined for $\forall\ i$ and $j' < j$. Then,

we can obtain M_{ij} from Eq. (A2) because

$$M_{ij} = A_{ij} - \sum_{k<j} M_{ik} M_{jk}$$

and right-hand side is already determined from the hypothesis for $j' < j$. ■

Using the above lemma, let us derive the theorem.

Proof. Let $(A|B) \in \mathbb{F}_2^{S\times 2N_q}$ be such a binary representation, with $A, B \in \mathbb{F}_2^{S\times N_q}$. Consider the following steps.

(1) Use Hadamard gates to make A have rank S.

(2) Use CNOTS to perform Gaussian elimination on A, producing

$$\left(\begin{array}{cc|cc} I & 0 & C & D \end{array} \right),$$

where $C \in \mathbb{F}_2^{S\times S}$ and $D \in \mathbb{F}_2^{S\times(N_q-S)}$ are the former and latter parts of B after Gaussian elimination on X space.

(3) Use Hadamard gates to swap,

$$\left(\begin{array}{cc|cc} I & D & C & 0 \end{array} \right).$$

(4) Again, use CNOTS to perform Gaussian elimination on X space, producing

$$\left(\begin{array}{cc|cc} I & 0 & C & 0 \end{array} \right).$$

(5) Commutativity states that IC^T is symmetric, implying that C is symmetric. We can apply phase gates to add a diagonal matrix to C to convert C into the form $C = MM^T$ for some invertible M based on Lemma A1.

(6) Use CNOTS, producing

$$\left(\begin{array}{cc|cc} M & 0 & M & 0 \end{array} \right).$$

(7) Apply phase gates to S qubits to obtain

$$\left(\begin{array}{cc|cc} M & 0 & 0 & \end{array} \right).$$

(8) Again, use CNOTS to perform Gaussian elimination on M, producing

$$\left(\begin{array}{cc|cc} I & 0 & 0 & \end{array} \right).$$

(9) Use Hadamard gates for swapping X and Z space

$$\left(\begin{array}{cc|cc} 0 & I & 0 & \end{array} \right).$$

Let us evaluate the circuit complexity in this algorithm. The parallel application of Hadamard gates and phase gates implies that steps 1, 3, 5, 7, and 9 can be achieved with constant depth and $\mathcal{O}(N_q)$ gates. Steps 2, 4, 6, and 8 involving CNOTS can be constructed by $\mathcal{O}(N_q^2)$ gates with $\mathcal{O}(\log N_q)$ depth for each step based on Ref. [59]. In total, we can transform any stabilizer into canonical form using $\mathcal{O}(N_q^2)$ gates with $\mathcal{O}(\log N_q)$ depth.

2. Efficient construction of the MUPB

Here, we prove that measurement bases shown in Fig. 3 satisfy the MUPB condition (Definition 1) in arbitrary spatial dimensions and boundary conditions and they can be constructed efficiently, i.e., $\mathcal{O}(N_q^2)$ gates and $\mathcal{O}(\log N_q)$ depth for the system size N_q based on the stabilizer formalism.

For $\mathbb{Z}_2$ LGTs in arbitrary dimensions and with general boundary conditions, the gauge groups spanned by independent Gauss's law operators [60] $\mathcal{G}_{N_q} = \langle \{G_n\}_{n=1}^S \rangle$ can be efficiently decomposed into the trivial stabilizer group using Theorem A3 followed by the Corollary below.

Corollary A1. For any $\mathbb{Z}_2$ gauge groups $\mathcal{G}_{N_q}$ with S independent generators, there exists an element of Clifford group $W \in \mathcal{C}_{N_q}$ such that

$$\begin{aligned} W\mathcal{G}_{N_q}W^\dagger &= \{I, Z\}^{\otimes S} \otimes \{I\}^{\otimes N_q - S} \\ &= \langle \{Z_1, Z_2, \ldots, Z_S\} \rangle. \end{aligned} \tag{A3}$$

W can be efficiently computed classically and constructed using $\mathcal{O}(N_q^2)$ Hadamard gates, phase gates, and CNOT gates with $\mathcal{O}(\log N_q)$ depth.

Based on this corollary, we can derive the efficient quantum circuits to construct the MUPB.

Theorem A4. The circuits in Fig. 3 satisfy the conditions of the MUPB. W is the element of the Clifford group shown in Corollary A1. W can be constructed using $\mathcal{O}(N_q^2)$ Hadamard gates, phase gates and CNOT gates with $\mathcal{O}(\log N_q)$ depth.

Here, we denote the measurement bases represented in Fig. 3 by the physical Z basis and the physical X basis with their respective basis states denoted by $\{|e_i^{(Z_{\rm phys})}\rangle\}$ and $\{|e_i^{(X_{\rm phys})}\rangle\}$ again.

Proof. First, we show that the two measurement bases are eigenstates of Gauss's law operators. The measurement basis in Fig. 3(a) is given by $|e_i^{(Z_{\rm phys})}\rangle = W^\dagger |b_i\rangle$ in terms of bit string states $|b_i\rangle = |b_i(1)b_i(2)\cdots b_i(N_q)\rangle$, where $b_i(j) \in \{0, 1\}$ means the jth bit string of the state $|b_i\rangle$, $\{|e_i^{(Z_{\rm phys})}\rangle\}_i$. Therefore,

$$\begin{aligned} G_n \left|e_i^{(Z_{\rm phys})}\right\rangle &= W^\dagger (W G_n W^\dagger) |b_i\rangle \\ &= z W^\dagger |b_i\rangle \\ &= z \left|e_i^{(Z_{\rm phys})}\right\rangle, \end{aligned}$$

where z is the corresponding eigenvalue. Note that $WG_nW^\dagger \in \{I, Z\}^{\otimes S} \otimes \{I\}^{N_q - S}$ and $|b_i\rangle$ is one of the eigenvectors of $WG_nW^\dagger$. Similarly, we can show

$$G_n \left|e_i^{(X_{\rm phys})}\right\rangle = z \left|e_i^{(X_{\rm phys})}\right\rangle.$$

Finally, we show the mutually unbiasedness of two measurement bases. By Theorem A1, $d_{\rm phys} = 2^{N_q - S}$. A measurement basis state $|e_i^{(Z_{\rm phys})}\rangle$ is uniquely stabilized by

$$\left\langle \{G_n\}_{n=1}^S \cup \{W^\dagger (-1)^{b_i(n)} Z_n W\}_{n=S+1}^{N_q} \right\rangle.$$

On the other hand, a measurement basis $|e_i^{(X_{\rm phys})}\rangle$ is uniquely stabilized by

$$\left\langle \{G_n\}_{n=1}^S \cup \{W^\dagger (-1)^{b_i(n)} X_n W\}_{n=S+1}^{N_q} \right\rangle.$$

Therefore, the number of different stabilizers is $N_q - S$. Thus,

$$\left| \left\langle e_i^{(X_{\rm phys})} \middle| e_j^{(Z_{\rm phys})} \right\rangle \right| = \frac{1}{2^{(N_q - S)/2}} = \frac{1}{\sqrt{d_{\rm phys}}},$$

implying that two measurement bases satisfy the condition of the MUPB. ■

APPENDIX B: AUTOCORRELATION ANALYSIS

In this appendix, we first derive Eq. (12). Next, we show Eq. (13). Finally, we compare the single-shot strategy with direct independent measurements on the Gibbs state.

First, for $t \geqslant 1$, the shot noise is independent of the Markov chain, and therefore

$$\mathrm{Cov}(O_k, O_{k+t}) = \mathrm{Cov}(\mu_k, \mu_{k+t}). \tag{B1}$$

Using the definition of the integrated autocorrelation time and Eq. (B1), we obtain

$$\begin{aligned}\tau(N_{\mathrm{shot}}) &= 1 + 2\sum_t \rho_O(t) \\ &= 1 + 2\sum_t \frac{\mathrm{Cov}(O_k, O_{k+t})}{\sigma_\mu^2 + (1/N_{\mathrm{shot}})\sigma_{\mathrm{shot}}^2} \\ &= 1 + \frac{\sigma_\mu^2}{\sigma_\mu^2 + (1/N_{\mathrm{shot}})\sigma_{\mathrm{shot}}^2}\left(2\sum_t \rho_\mu(t)\right).\end{aligned}$$

Since $\tau_\mu = 1 + 2\sum_{t=1}^{\infty}\rho_\mu(t)$, we find

$$\tau(N_{\mathrm{shot}}) = 1 + \frac{N_{\mathrm{shot}}}{N_{\mathrm{shot}} + r}(\tau_\mu - 1), \tag{B2}$$

where we define $r := \sigma_{\mathrm{shot}}^2/\sigma_\mu^2$. The variance is then given by

$$\begin{aligned}&\mathrm{Var}[\bar{O}](N_{\mathrm{shot}}) \\ &= \frac{\tau(N_{\mathrm{shot}})}{N_{\mathrm{chain}}}\sigma_\mu^2 + \frac{\tau(N_{\mathrm{shot}})}{N_{\mathrm{shot}}N_{\mathrm{chain}}}\sigma_{\mathrm{shot}}^2 \\ &= \frac{\tau(N_{\mathrm{shot}})(N_{\mathrm{shot}}M + 1)}{N_{\mathrm{tot}}}\left(\sigma_\mu^2 + \frac{\sigma_{\mathrm{shot}}^2}{N_{\mathrm{shot}}}\right) \\ &= \frac{\sigma_\mu^2}{N_{\mathrm{tot}}}\frac{(N_{\mathrm{shot}}M + 1)(N_{\mathrm{shot}}\tau_\mu + r)}{N_{\mathrm{shot}}}.\end{aligned}$$

Treating N_{shot} as a continuous variable, the stationary point of this function is given by

$$N_{\mathrm{shot}}^* = \sqrt{\frac{r}{\tau_\mu M}} = \sqrt{\frac{\sigma_{\mathrm{shot}}^2}{\sigma_\mu^2 \tau_\mu M}}.$$

Because we require $N_{\mathrm{shot}} \geqslant 1$, the optimal shot number is $N_{\mathrm{shot}}^* \sim \max(1, \sqrt{\sigma_{\mathrm{shot}}^2/\sigma_\mu^2\tau_\mu M})$.

Next, we prove Eq. (13). When $r/\tau_\mu \leqslant M$, the constrained optimum is simply $N_{\mathrm{shot}}^* = 1$, and hence

$$\frac{\mathrm{Var}[\bar{O}](1)}{\mathrm{Var}[\bar{O}](N_{\mathrm{shot}}^*)} = 1. \tag{B3}$$

When $r/\tau_\mu > M$, the continuous optimum satisfies $N_{\mathrm{shot}} > 1$, and thus

$$\begin{aligned}\mathrm{Var}[\bar{O}](N_{\mathrm{shot}}^*) &= \frac{\sigma_\mu^2}{N_{\mathrm{tot}}}\frac{(M\sqrt{\frac{r}{\tau_\mu M}} + 1)(\tau_\mu\sqrt{\frac{r}{\tau_\mu M}} + r)}{\sqrt{\frac{r}{\tau_\mu M}}} \\ &= \frac{\sigma_\mu^2}{N_{\mathrm{tot}}}(\sqrt{rM} + \sqrt{\tau_\mu})^2.\end{aligned}$$

On the other hand,

$$\mathrm{Var}[\bar{O}](N_{\mathrm{shot}} = 1) = \frac{\sigma_\mu^2}{N_{\mathrm{tot}}}(M + 1)(\tau_\mu + r).$$

Therefore, the ratio becomes

$$R := \frac{\mathrm{Var}[\bar{O}](N_{\mathrm{shot}} = 1)}{\mathrm{Var}[\bar{O}](N_{\mathrm{shot}}^*)} = \frac{(M + 1)(\tau_\mu + r)}{(\sqrt{\tau_\mu} + \sqrt{rM})^2}.$$

If we set $x := \sqrt{r/\tau_\mu}$, then $R = [(M + 1)(1 + x^2)]/(1 + \sqrt{M}x)^2$ $(\sqrt{M} < x < \infty)$. Due to $dR(x)/dx > 0$, $R(x)$ is monotonically increasing function. Considering

$$\lim_{x\to\sqrt{M}} R(x) = 1$$

and

$$\lim_{x\to\infty} R(x) = \frac{M + 1}{M} = 1 + \frac{1}{M},$$

we conclude that

$$1 < R(x) < 1 + \frac{1}{M}. \tag{B4}$$

Combining the two cases [Eqs. (B3) and ((B4), we conclude that

$$1 \leqslant \frac{\mathrm{Var}[\bar{O}](1)}{\mathrm{Var}[\bar{O}](N_{\mathrm{shot}}^*)} < 1 + \frac{1}{M} \leqslant 2.$$

This proves Eq. (13).

Finally, we compare the single-shot strategy with direct independent measurements on the Gibbs state. As in the QMETTS setting considered above, we assume that the observable is decomposed as $\mathcal{O} = \sum_{m=1}^{M}\mathcal{O}_m$, and that each component $\mathcal{O}_m$ can be measured with one quantum circuit on the Gibbs state $\rho_{\beta,\mu} = e^{-\beta(H-\mu N)}/Z$. For a fixed total number of circuit executions N_{tot}, the number of shots allocated to each component $\mathcal{O}_m$ is $N_{\mathrm{shot}}^{\mathrm{Gibbs}} = N_{\mathrm{tot}}/M$. The corresponding estimator is

$$\bar{O}_{\mathrm{Gibbs}} = \frac{1}{N_{\mathrm{shot}}^{\mathrm{Gibbs}}}\sum_{l=1}^{N_{\mathrm{shot}}^{\mathrm{Gibbs}}}\sum_{m=1}^{M} Y_{m,l},$$

where $Y_{m,l}$ denotes the lth measurement outcome of $\mathcal{O}_m$ on $\rho_{\beta,\mu}$. Assuming that the measurements for different m and l are independent, its variance is

$$\begin{aligned}\mathrm{Var}_{\mathrm{Gibbs}}[\bar{O}_{\mathrm{Gibbs}}] &= \frac{1}{N_{\mathrm{shot}}^{\mathrm{Gibbs}}}\sum_{m=1}^{M}\mathrm{Var}_{\mathrm{Gibbs}}[Y_{m,l}] \\ &= \frac{1}{N_{\mathrm{shot}}^{\mathrm{Gibbs}}}\sigma_{\mathrm{Gibbs}}^2,\end{aligned}$$

where

$$\sigma_{\mathrm{Gibbs}}^2 := \sum_{m=1}^{M}\left[\mathrm{Tr}\left[\mathcal{O}_m^2\rho_{\beta,\mu}\right] - \left(\mathrm{Tr}\left[\mathcal{O}_m\rho_{\beta,\mu}\right]\right)^2\right].$$

We next relate this quantity to the variance components appearing in the QMETTS estimator. Let $\mu_{m,i} := \langle\phi_i|\mathcal{O}_m|\phi_i\rangle$, implying that $\mu_i := \sum_{m=1}^{M}\mu_{m,i}$. Using the METTS ensemble identity $\sum_i \mathrm{Prob}_i\langle\phi_i|A|\phi_i\rangle = \mathrm{Tr}[A\rho_{\beta,\mu}]$, for any observable A, we obtain

$$\begin{aligned}\sigma_{\mathrm{shot}}^2 + \sigma_\mu^2 = \sum_i \mathrm{Prob}_i \sum_{m=1}^{M}\left[\langle\phi_i|\mathcal{O}_m^2|\phi_i\rangle - \mu_{m,i}^2\right] \\ + \mathrm{Var}_{\mathrm{chain}}[\mu_i]\end{aligned}$$

$$= \sum_{m=1}^{M} \left[\mathrm{Tr}\left[\mathcal{O}_m^2 \rho_{\beta,\mu}\right] - \left(\mathrm{Tr}\left[\mathcal{O}_m \rho_{\beta,\mu}\right]\right)^2\right] + 2 \sum_{m<m'} \mathrm{Cov}_{\mathrm{chain}}(\mu_{m,i}, \mu_{m',i}) = \sigma^2_{\mathrm{Gibbs}} + 2 \sum_{m<m'} \mathrm{Cov}_{\mathrm{chain}}(\mu_{m,i}, \mu_{m',i}).$$

The covariance term

$$\mathrm{Cov}_{\mathrm{chain}}(\mu_{m,i}, \mu_{m',i}) = \sum_i \mathrm{Prob}_i \mu_{m,i} \mu_{m',i} - \left(\sum_i \mathrm{Prob}_i \mu_{m,i}\right) \times \left(\sum_i \mathrm{Prob}_i \mu_{m',i}\right)$$

represents the covariance between the components $\mathcal{O}_m$ and can be either positive or negative in general.

Substituting this relation into the single-shot QMETTS variance gives

$$\mathrm{Var}[\bar{O}](1) = \tau(1)\left(1 + \frac{1}{M}\right)\left[\mathrm{Var}_{\mathrm{Gibbs}}[\bar{O}_{\mathrm{Gibbs}}] + \frac{2\sum_{m<m'} \mathrm{Cov}(\mu_{m,i}, \mu_{m',i})}{N_{\mathrm{tot}}/M}\right].$$

Therefore, apart from the covariance correction among different components of the observable in the METTS ensemble, the single-shot QMETTS variance is related to the direct Gibbs measurement variance by the multiplicative factor $\tau(1)(1 + 1/M)$.

APPENDIX C: MUPB PROOF FOR THE CONCRETE SETUP

We now provide a rigorous proof of our main statement for the concrete setup considered in the numerical simulation. The improved measurement bases shown in Fig. 4 were obtained

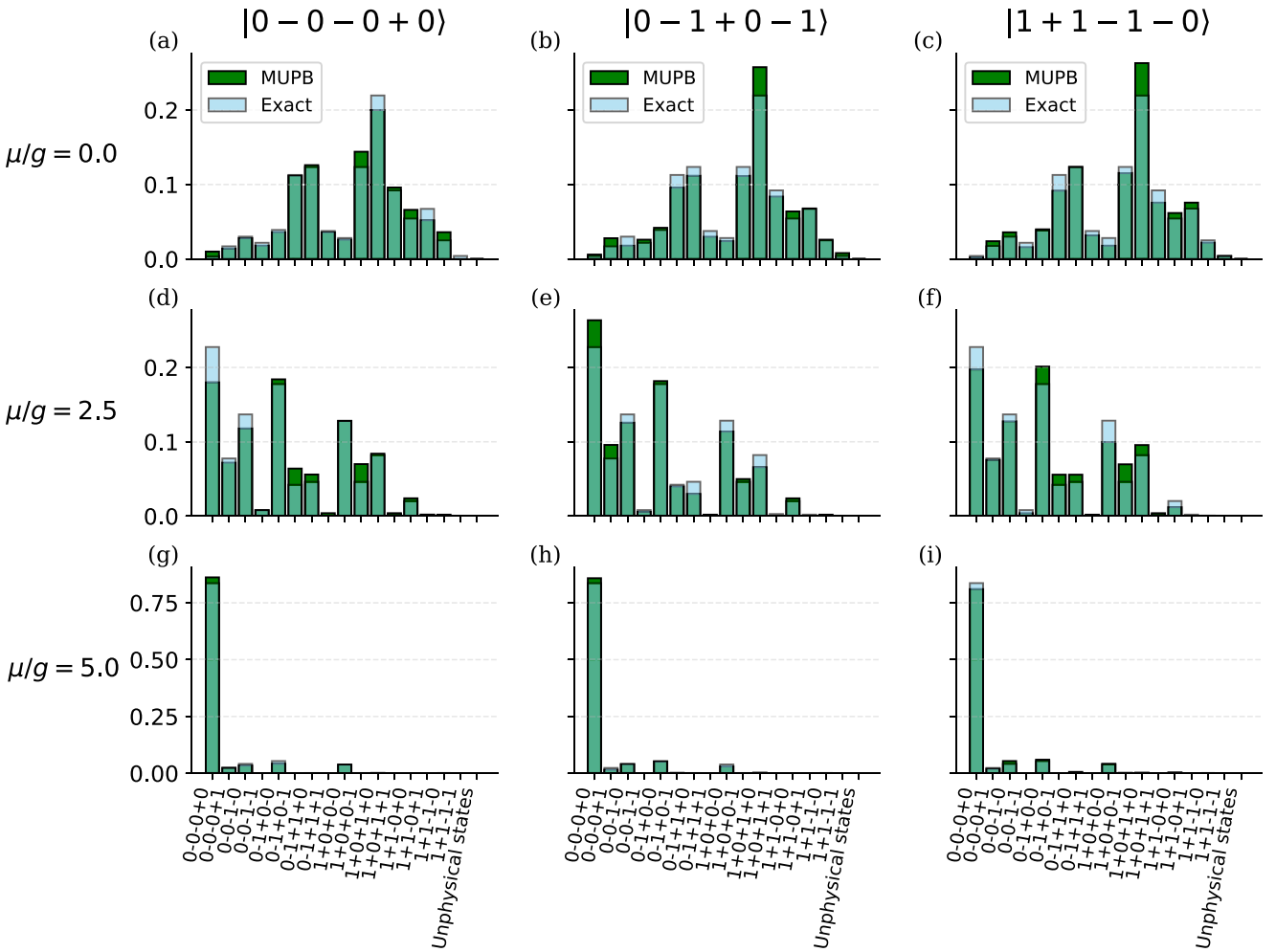

FIG. 12. Dependence of probability distributions on the initial state. Sampling distributions of collapse states obtained from projective measurements using the MUPBs during QMETTS sampling (green) and the exact probabilities (light blue) for $L_{\mathrm{KS}} = 4$. The ITE is implemented exactly using classical computation. The panels are arranged by chemical potential μ/g (rows) and by the initial state: $|0-0-0+0\rangle$ (left), $|0-1+0-1\rangle$ (middle), and $|1+1-1-0\rangle$ (right). The inverse temperature is fixed at $\beta g = 1.0$. Top panels (a)–(c) correspond to $\mu/g = 0.0$, middle panels (d)–(f) to $\mu/g = 2.5$, and bottom panels (g)–(i) to $\mu/g = 5.0$. The exact probabilities are calculated as $\mathrm{Prob}_i = \langle i^{(Z_{\mathrm{phys}})}|e^{-\beta(H-\mu N)}|i^{(Z_{\mathrm{phys}})}\rangle/Z$. All QMETTS results are obtained with $N_{\mathrm{chain}} = 1000$.

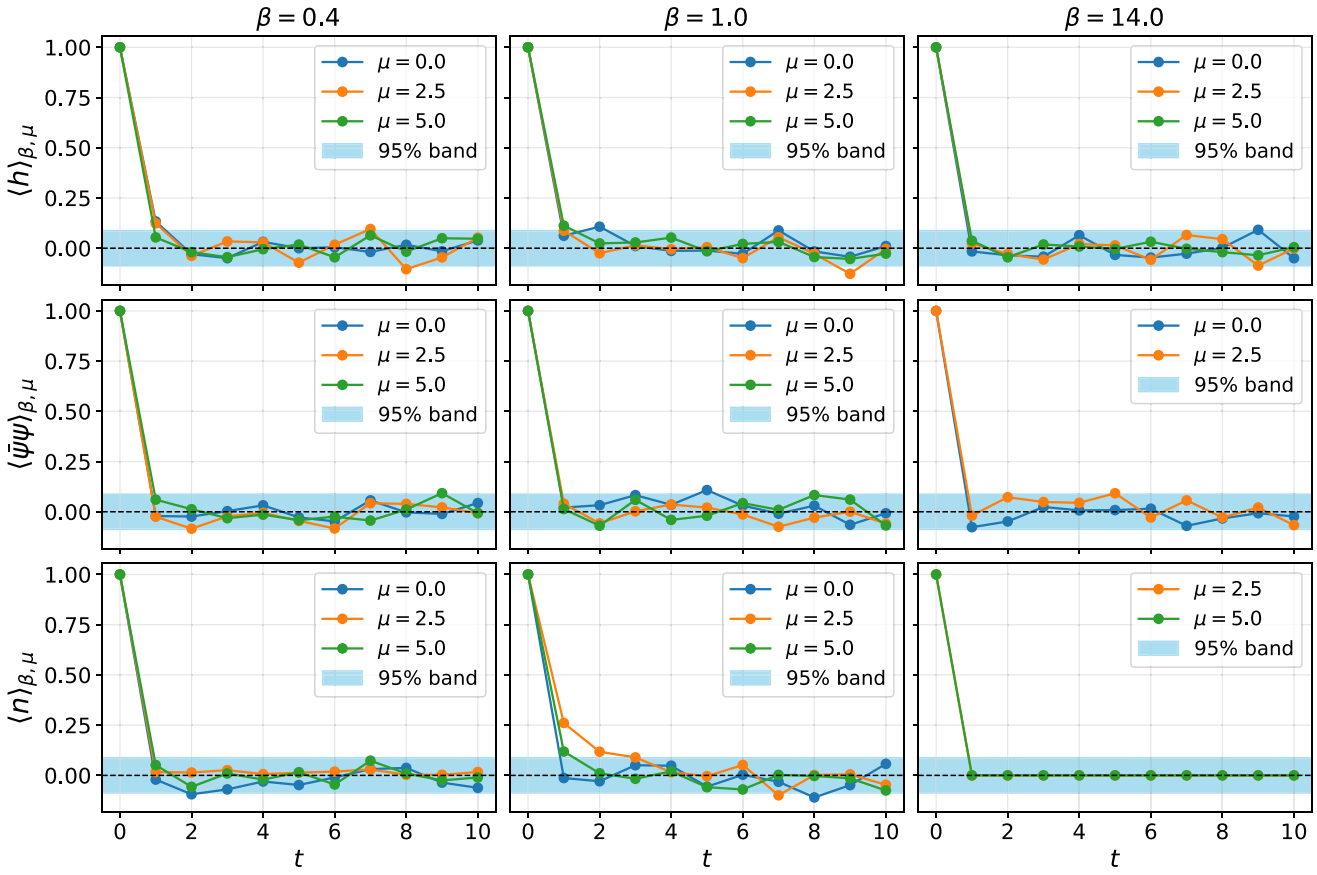

FIG. 13. Autocorrelation functions $\rho_{\mathcal{O}}(t, N_{\text{shot}} = 1)$ for the energy density (top row), chiral condensate (middle row), and fermion number density (bottom row). The inverse temperature is fixed at $\beta g = 0.4$ (left column), $\beta g = 1.0$ (middle column), and $\beta g = 14.0$ (right column). Blue, orange, and green markers correspond to chemical potentials $\mu/g = 0.0$, 2.5, and 5.0, respectively. The shaded "95% band" indicates the approximate 95% confidence interval around zero expected under the assumption of uncorrelated samples. For the cases $\mathcal{O} = \bar{\psi}\psi$, $\beta g = 14.0$, and $\mu/g = 5.0$ and $\mathcal{O} = n$, $\beta g = 14.0$, and $\mu/g = 0.0$, we obtain zero variance and cannot define the corresponding autocorrelation functions.

heuristically. In what follows, we show, using the stabilizer formalism again, that these bases satisfy the defining properties of MUPB. We begin by showing that the physical Z basis and X basis represented in Fig. 4 are eigenstates of the Gauss's law operators (Item 1 of Definition 1). The physical Z basis is given by

$$\left|e_i^{(Z_{\text{phys}})}\right\rangle = \left(\prod_{m=1}^{L_{\text{KS}}-1} \sigma^{\mathsf{H}}_{m,m+1}\right)^{\dagger} |b_i\rangle =: U_{\mathsf{H}}^{\dagger}|b_i\rangle.$$

Since $G_n|e_i^{(Z_{\text{phys}})}\rangle = G_n U_{\mathsf{H}}^{\dagger}|b_i\rangle = U_{\mathsf{H}}^{\dagger}\cdot(U_{\mathsf{H}} G_n U_{\mathsf{H}}^{\dagger})|b_i\rangle$ and $\sigma^{\mathsf{H}}_{m-1,m}\sigma^{X}_{m-1,m}\sigma^{\mathsf{H}}_{m-1,m} = \sigma^{Z}_{m-1,m}$, we obtain

$$U_{\mathsf{H}} G_n U_{\mathsf{H}}^{\dagger} = (-1)^n \sigma^{Z}_{n-1,n} Z_n \sigma^{Z}_{n,n+1}.$$

As $(-1)^n \sigma^{Z}_{n-1,n} Z_n \sigma^{Z}_{n,n+1}|b_i\rangle = z|b_i\rangle$, where $z \in \{\pm 1\}$ is the corresponding eigenvalue, we can show

$$U_{\mathsf{H}}^{\dagger}\cdot(U_{\mathsf{H}} G_n U_{\mathsf{H}}^{\dagger})|b_i\rangle = z U_{\mathsf{H}}^{\dagger}|b_i\rangle = z\left|e_i^{(Z_{\text{phys}})}\right\rangle.$$

As for $\{|e_i^{(X_{\text{phys}})}\rangle\}$ represented in Fig. 4, since it can be written as $|e_i^{(X_{\text{phys}})}\rangle = U^{\dagger}|e_i^{(Z_{\text{phys}})}\rangle$, where U is defined in Eq. (15),

$$\begin{aligned} G_n\left|e_i^{(X_{\text{phys}})}\right\rangle &= G_n U^{\dagger}\left|e_i^{(Z_{\text{phys}})}\right\rangle \\ &= U^{\dagger}(U G_n U^{\dagger})\left|e_i^{(Z_{\text{phys}})}\right\rangle. \end{aligned}$$

Then,

$$\begin{aligned} G_n &= (-1)^n \sigma^{X}_{n-1,n} Z_n \sigma^{X}_{n,n+1} \\ &\to V(-1)^n X_n \xrightarrow{\prod_m \sigma^{\mathsf{H}}_{m,m+1}} (-1)^n X_n \\ &\to V^{\dagger}(-1)^n \sigma^{X}_{n-1,n} Z_n \sigma^{X}_{n,n+1} = G_n, \end{aligned}$$

where $A \xrightarrow{X} B$ denotes the conjugation of A by X, $B = XAX^{\dagger}$, and V is defined in Eq. (16). Thus, the Gauss's law operators are invariant under the conjugation $U G_n U^{\dagger} = G_n$ and we can also prove that $\{|e_i^{(X_{\text{phys}})}\rangle\}$ are eigenstates of G_n.

Next, we prove the mutual unbiasedness in physical subspace: $|\langle e_i^{(Z_{\text{phys}})}|e_j^{(X_{\text{phys}})}\rangle| = 1/\sqrt{d_{\text{phys}}}$ (Item 2 in Definition 1). The stabilizer group that stabilizes $|e_i^{(Z_{\text{phys}})}\rangle$ is given by

$$\begin{aligned} \langle \{G_n\}_{n=1}^{L_{\text{KS}}-1} &\cup \{(-1)^{b_i(g_{n,n+1})}\sigma^{X}_{n,n+1}\}_{n=1}^{L_{\text{KS}}-1} \\ &\cup \{(-1)^{b_i(f_L)} Z_{L_{\text{KS}}}\}\rangle. \end{aligned}$$

From Theorem A1, we can see that the dimension of this stabilized space is $2^0 = 1$ and the state is uniquely determined up to a global phase. Then, the state $|e_i^{(X_{\text{phys}})}\rangle = U^{\dagger}|e_i^{(Z_{\text{phys}})}\rangle$ can be described by following stabilizer operators through the

conjugation under U:

$$\{(-1)^{b_i(g_{n,n+1})}\sigma^X_{n,n+1}\}_n \to U\{(-1)^{b_i(g_{n,n+1})}X_n\sigma^Z_{n,n+1}X_{n+1}\}_n,$$
$$\times\{Z_{L_{\rm KS}}\} \xrightarrow{U} \{X_{L_{\rm KS}}\},$$

and Gauss's law operators are invariant, $UG_nU^\dagger = G_n$. Since different operators are $\{(-1)^{b_i(g_{n,n+1})}\sigma^X_{n,n+1}\}_{n=1}^{L_{\rm KS}-1} \cup \{Z_{L_{\rm KS}}\}$ and $\{(-1)^{b_i(g_{n,n+1})}X_n\sigma^Z_{n,n+1}X_{n+1}\}_{n=1}^{L_{\rm KS}-1} \cup \{X_{L_{\rm KS}}\}$, the number of different generators is $L_{\rm KS}$. Then, by Theorem A2, we can compute the inner product as

$$\left|\langle e_i^{(Z_{\rm phys})}|e_j^{(X_{\rm phys})}\rangle\right| = \frac{1}{2^{L_{\rm KS}/2}} = \frac{1}{\sqrt{d_{\rm phys}}}.$$

Therefore, the physical Z basis and the physical X basis satisfy the condition of the MUPB.

APPENDIX D: IMPLEMENTATION OF QITE

In this appendix, we briefly provide the details of the implementation of QITE by Motta *et al.* [12]. To implement the QITE via Trotterization, we decompose the Hamiltonian with the chemical potential term into three parts, $H - \mu N = H_e + H_o + H_D$, where

$$H_{e/o} = -\frac{1}{4a}\sum_{n\in\text{even/odd}}^{L_{\rm KS}-1}\left[X_n\sigma^Z_{n,n+1}X_{n+1} + Y_n\sigma^Z_{n,n+1}Y_{n+1}\right],$$
$$H_D = ag^2\sum_{n=1}^{L_{\rm KS}-1}\left(1-\sigma^X_{n,n+1}\right) + \frac{m}{2}\sum_{n=1}^{L_{\rm KS}}(-1)^{n+1}Z_n - \frac{\mu}{2}\sum_{n=1}^{L_{\rm KS}}Z_n.$$

As discussed below, this decomposition suppresses Trotter errors. We employ a second-order Trotter decomposition,

$$e^{-\beta(H-\mu N)/2} = (e^{-\Delta\beta(H-\mu N)/2})^{N_s} \approx \left(\overleftarrow{\prod_{\gamma\in\{e,o,D\}}}\ \overrightarrow{\prod_{\gamma\in\{e,o,D\}}} e^{-\Delta\beta H_\gamma/4}\right)^{N_s},$$

where $\Delta\beta = \beta/N_s$ and N_s is the number of Trotter steps. The left- and right-ordered products are defined as

$$\overleftarrow{\prod_{\gamma\in\{e,o,D\}}} e^{-\Delta\beta H_\gamma/4} := e^{-\Delta\beta H_e/4}e^{-\Delta\beta H_o/4}e^{-\Delta\beta H_D/4},$$
$$\overrightarrow{\prod_{\gamma\in\{e,o,D\}}} e^{-\Delta\beta H_\gamma/4} := e^{-\Delta\beta H_D/4}e^{-\Delta\beta H_o/4}e^{-\Delta\beta H_e/4}.$$

Following Ref. [61], the Trotter error can be bounded in the spectral norm as

$$\left\|\overleftarrow{\prod_\gamma}\overrightarrow{\prod_\gamma}e^{-\Delta\beta H_\gamma/4} - e^{-\Delta\beta(H-\mu N)/2}\right\|_\infty = \mathcal{O}\left(\tilde{\alpha}_{\rm comm}(\Delta\beta)^3 e^{4\Delta\beta\sum_\gamma\|H_\gamma\|_\infty}\right),$$

where $\tilde{\alpha}_{\rm comm} = \sum_{\gamma_1,\gamma_2,\gamma_3}\|[H_{\gamma_1},[H_{\gamma_2},H_{\gamma_3}]]\|_\infty$. Thus, $\Delta\beta \to 0$ recovers the exact imaginary-time evolution operators. In the present work, we take $\Delta\beta = 0.01$.

APPENDIX E: DEPENDENCE OF THE PROBABILITY DISTRIBUTIONS ON THE INITIAL STATE IN THE NUMERICAL BENCHMARK

In this appendix, we show the sampling distributions obtained from several different initial states using the MUPBs for the numerical benchmark discussed in Sec. VI. We use $N_{\rm chain} = 1000$ for each initial state, matching the setup in Sec. VI. As seen in Fig. 12, the sampled collapse-state distributions obtained with the MUPBs are all consistent with the exact stationary distribution within finite-sampling fluctuations, indicating that no clear dependence on the initial state is observed in the present setup.

APPENDIX F: AUTOCORRELATION FUNCTIONS

In this appendix, we examine the behavior of the autocorrelation function $\rho_{\mathcal{O}}(t, N_{\rm shot} = 1)$ defined in Eq. (11) for the three observables $\mathcal{O} = h$, $\bar{\psi}\psi$, and n. Figure 13 shows the corresponding autocorrelation functions. The shaded "95% band" indicates the approximate 95% bounds around zero expected under the assumption of uncorrelated samples. For all cases, the correlations decay rapidly with increasing lag

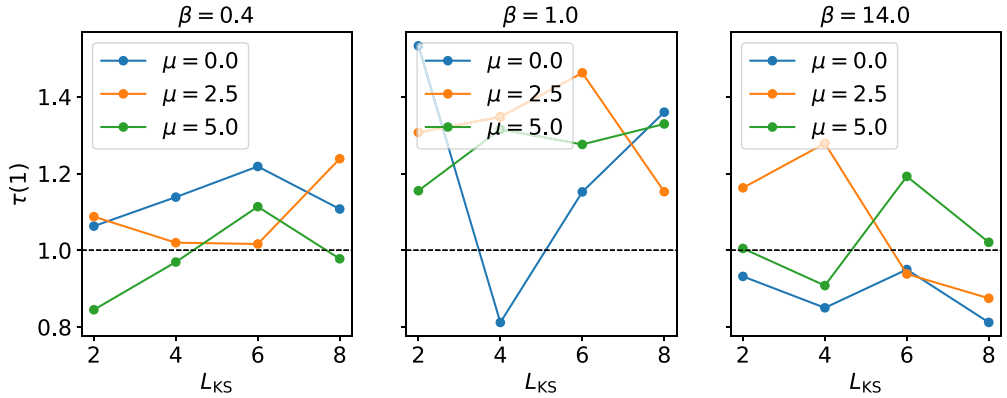

FIG. 14. System-size dependence of the integrated autocorrelation time for the energy density. The left, middle, and right panels correspond to $\beta g = 0.4$, 1.0, and 14.0, respectively. In each panel, blue, orange, and green markers correspond to $\mu/g = 0.0$, 2.5, and 5.0, respectively.

time t and enter the shaded regions within a few ($\sim$3) steps. Based on this behavior, we choose the window size $w = 3$ for estimating the integrated autocorrelation time.

APPENDIX G: SYSTEM-SIZE DEPENDENCE OF THE INTEGRATED AUTOCORRELATION TIME

In this appendix, we investigate the system-size dependence of the integrated autocorrelation time by varying $L_{\mathrm{KS}} = 2, 4, 6$, and 8. We use the window parameter w introduced in Sec. VI and set $w = 3$ based on the analysis of autocorrelation functions in Appendix F. Figure 14 shows the resulting system-size dependence of the integrated autocorrelation times $\tau(1)$. Within the range of system sizes considered, we observe no clear systematic dependence of the integrated autocorrelation time on the system size. In all cases, the estimated autocorrelation times remain of order unity.

[1] M. A. Stephanov, QCD phase diagram: An overview, PoS LAT **2006**, 024 (2006).

[2] K. Fukushima and T. Hatsuda, The phase diagram of dense QCD, Rep. Prog. Phys. **74**, 014001 (2011).

[3] E. V. Shuryak, Quantum chromodynamics and the theory of superdense matter, Phys. Rep. **61**, 71 (1980).

[4] K. G. Wilson, Confinement of quarks, Phys. Rev. D **10**, 2445 (1974).

[5] J. B. Kogut and L. Susskind, Hamiltonian formulation of Wilson's lattice gauge theories, Phys. Rev. D **11**, 395 (1975).

[6] S. Durr *et al.*, *Ab initio* determination of light hadron masses, Science **322**, 1224 (2008).

[7] K. Nagata, Finite-density lattice QCD and sign problem: Current status and open problems, Prog. Part. Nucl. Phys. **127**, 103991 (2022).

[8] C.-F. Chen, M. J. Kastoryano, F. G. S. L. Brandão, and A. Gilyén, Efficient quantum thermal simulation, Nature (London) **646**, 561 (2025).

[9] C.-F. Chen, M. J. Kastoryano, and A. Gilyén, An efficient and exact noncommutative quantum Gibbs sampler, arXiv:2311.09207.

[10] M. Consiglio, J. Settino, A. Giordano, C. Mastroianni, F. Plastina, S. Lorenzo, S. Maniscalco, J. Goold, and T. J. G. Apollaro, Variational Gibbs state preparation on noisy intermediate-scale quantum devices, Phys. Rev. A **110**, 012445 (2024).

[11] T. J. Sewell, C. D. White, and B. Swingle, Thermal multi-scale entanglement renormalization ansatz for variational Gibbs state preparation, arXiv:2210.16419.

[12] M. Motta, C. Sun, A. T. K. Tan, M. J. O. Rourke, E. Ye, A. J. Minnich, F. G. S. L. Brandão, and G. K.- L. Chan, Determining eigenstates and thermal states on a quantum computer using quantum imaginary time evolution, Nat. Phys. **16**, 205 (2020).

[13] I.-C. Chen, J. C. Getelina, K. Pollock, A. Khindanov, S. Sen, Y.-X. Yao, and T. Iadecola, Classical and quantum simulations of $1+1$-dimensional $\mathbb{Z}_2$ gauge theory at finite temperature and density, Commun. Phys. **8**, 375 (2025).

[14] A. Tomiya, Schwinger model at finite temperature and density with beta VQE, arXiv:2205.08860.

[15] F. Turro, A. Ciavarella, and X. Yao, Classical and quantum computing of shear viscosity for $(2+1)D$ SU(2) gauge theory, Phys. Rev. D **109**, 114511 (2024).

[16] F. Turro and X. Yao, Emergent hydrodynamic mode on SU(2) plaquette chains and quantum simulation, Phys. Rev. D **111**, 094502 (2025).

[17] A. T. Than *et al.*, The phase diagram of quantum chromodynamics in one dimension on a quantum computer, Nat. Commun. **16**, 10288 (2025).

[18] T. Schmale and H. Weimer, Stabilizing quantum simulations of lattice gauge theories by dissipation, Phys. Rev. Res. **6**, 033306 (2024).

[19] E. Ballini, J. Mildenberger, M. M. Wauters, and P. Hauke, Symmetry verification for noisy quantum simulations of non-Abelian lattice gauge theories, Quantum **9**, 1802 (2025).

[20] A. Rajput, A. Roggero, and N. Wiebe, Quantum error correction with gauge symmetries, npj Quantum Inf. **9**, 41 (2023).

[21] M. Carena, H. Lamm, Y.-Y. Li, and W. Liu, Quantum error thresholds for gauge-redundant digitizations of lattice field theories, Phys. Rev. D **110**, 054516 (2024).

[22] L. Spagnoli, A. Roggero, and N. Wiebe, Fault-tolerant simulation of lattice gauge theories with gauge covariant codes, Quantum **10**, 1968 (2026).

[23] Z. Davoudi, N. Mueller, and C. Powers, Towards quantum computing phase diagrams of gauge theories with thermal pure quantum states, Phys. Rev. Lett. **131**, 081901 (2023).

[24] M. Fromm, O. Philipsen, M. Spannowsky, and C. Winterowd, Simulating Z_2 lattice gauge theory with the variational quantum thermalizer, EPJ Quantum Technol. **11**, 20 (2024).

[25] E. Ballini, G. Clemente, M. D'Elia, L. Maio, and K. Zambello, Quantum computation of thermal averages for a non-Abelian D_4 lattice gauge theory via quantum Metropolis sampling, Phys. Rev. D **109**, 034510 (2024).

[26] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2010).

[27] D. Gottesman, Stabilizer codes and quantum error correction, Caltech, Ph.D. thesis, 1997.

[28] M. Binder and T. Barthel, Symmetric minimally entangled typical thermal states for canonical and grand-canonical ensembles, Phys. Rev. B **95**, 195148 (2017).

[29] S. Goto and I. Danshita, Quasiexact Kondo dynamics of fermionic alkaline-earth-like atoms at finite temperatures, Phys. Rev. Lett. **123**, 143002 (2019).

[30] D. Horn, M. Weinstein, and S. Yankielowicz, Hamiltonian approach to $Z(N)$ lattice gauge theories, Phys. Rev. D **19**, 3715 (1979).

[31] J. Mildenberger, W. Mruczkiewicz, J. C. Halimeh, Z. Jiang, and P. Hauke, Confinement in a $\mathbb{Z}_2$ lattice gauge theory on a quantum computer, Nat. Phys. **21**, 312 (2025).

[32] T. Hayata, K. Seki, and A. Yamamoto, Floquet prethermalization of Z_2 lattice gauge theory on superconducting qubits, Phys. Rev. D **110**, 114503 (2024).

[33] P. Jordan and E. P. Wigner, About the Pauli exclusion principle, Z. Phys. **47**, 631 (1928).

[34] S. R. White and E. M. Stoudenmire, Minimally entangled typical thermal state algorithms, New J. Phys. **12**, 023026 (2010).

[35] S. R. White, Minimally entangled typical quantum states at finite temperature, Phys. Rev. Lett. **102**, 190601 (2009).
[36] N. Madras and A. D. Sokal, The pivot algorithm: A highly efficient Monte Carlo method for the self-avoiding walk, J. Stat. Phys. **50**, 109 (1988).
[37] A. Sokal, Monte Carlo methods in statistical mechanics: Foundations and new algorithms, in *Functional Integration: Basics and Applications* (Springer, Berlin, 1997), pp. 131–192.
[38] U. Wolff (ALPHA Collaboration), Monte Carlo errors with less errors, Comput. Phys. Commun. **156**, 143 (2004).
[39] B. Bruognolo, J. von Delft, and A. Weichselbaum, Symmetric minimally entangled typical thermal states, Phys. Rev. B **92**, 115105 (2015).
[40] W. K. Wootters and B. D. Fields, Optimal state-determination by mutually unbiased measurements, Ann. Phys. **191**, 363 (1989).
[41] S. Aaronson and D. Gottesman, Improved simulation of stabilizer circuits, Phys. Rev. A **70**, 052328 (2004).
[42] In Figs. 7 and 8, no initial samples are discarded as burn-in, yet the QMETTS estimates for $L_{\mathrm{KS}} = 8$ remain consistent with the exact results within statistical fluctuations. This provides additional evidence that burn-in effects are not practically significant for the chain lengths and parameter ranges considered here.
[43] The integrated autocorrelation time can become slightly negative due to statistical fluctuations for the short-chain case with $N_{\mathrm{chain}} = 40$. To avoid underestimating the statistical uncertainty, we conservatively replace estimates below unity by $\max(\tau(N_{\mathrm{shot}}), 1)$.
[44] M. Kliesch, C. Gogolin, M. J. Kastoryano, A. Riera, and J. Eisert, Locality of temperature, Phys. Rev. X **4**, 031019 (2014).
[45] D. Gottesman, Fault tolerant quantum computation with higher dimensional systems, Chaos, Solitons Fractals **10**, 1749 (1999).
[46] V. Gheorghiu, Standard form of qudit stabilizer groups, Phys. Lett. A **378**, 505 (2014).
[47] L. Spagnoli, A. Roggero, and N. Wiebe, Qudit stabiliser codes for $\mathbb{Z}_N$ lattice gauge theories with matter, arXiv:2602.20661.
[48] M. Turco, L. Spagnoli, and A. Roggero, Binary Gauss stabilizers for Abelian lattice gauge theories, arXiv:2607.14861.
[49] J. R. Stryker, Oracles for Gauss's law on digital quantum computers, Phys. Rev. A **99**, 042301 (2019).
[50] J. P. Lacambra, A. Chatwin-Davies, M. Honda, and P. A. Hoehn, Gauss law codes and vacuum codes from lattice gauge theories, arXiv:2604.06087.
[51] X. Yao, Quantum error correction codes for truncated SU(2) lattice gauge theories, Phys. Rev. D **113**, 114512 (2026).
[52] M. Sekiyama and L. Nagano, Ground state preparation in two-dimensional pure $\mathbb{Z}_2$ lattice gauge theory via deterministic quantum imaginary time evolution, arXiv:2604.17874 [Phys. Rev. D (to be published)].
[53] T. Kosugi, Y. Nishiya, H. Nishi, and Y.-I. Matsushita, Imaginary-time evolution using forward and backward real-time evolution with a single ancilla: First-quantized eigensolver algorithm for quantum chemistry, Phys. Rev. Res. **4**, 033121 (2022).
[54] F. Turro, A. Roggero, V. Amitrano, P. Luchi, K. A. Wendt, J. L. DuBois, S. Quaglioni, and F. Pederiva, Imaginary-time propagation on a quantum chip, Phys. Rev. A **105**, 022440 (2022).
[55] A. Gilyén, Y. Su, G. H. Low, and N. Wiebe, Quantum singular value transformation and beyond: Exponential improvements for quantum matrix arithmetics, in 51st annual, ACM SIGACT Symp. Theory Comput. **6**, 193 (2018).
[56] K. Hejazi, M. Motta, and G. K.-L. Chan, Adiabatic quantum imaginary time evolution, Phys. Rev. Res. **6**, 033084 (2024).
[57] M. Gluza, J. Son, B. H. Tiang, R. Zander, R. Seidel, Y. Suzuki, Z. Holmes, and N. H. Y. Ng, Double-bracket quantum algorithms for quantum imaginary-time evolution, Phys. Rev. Lett. **136**, 020601 (2026).
[58] N. Matsumoto, S. Tsutsui, Y. O. Nakagawa, Y. Hidaka, S. Kanasugi, K. Maruyama, H. Oshima, and S. Sato, Quantum many-body simulation of finite-temperature systems with sampling a series expansion of a quantum imaginary-time evolution, Phys. Rev. Res. **7**, 013254 (2025).
[59] C. Moore and M. Nilsson, Parallel quantum computation and quantum codes, SIAM J. Comput. **31**, 799 (2012).
[60] As follows from Theorem A3, the stabilizer generators corresponding to Gauss's law constraints are assumed to be independent. If the original set of Gauss's law operators contains redundant constraints, we first replace it with an independent generating subset.
[61] A. M. Childs, Y. Su, M. C. Tran, N. Wiebe, and S. Zhu, Theory of Trotter error with commutator scaling, Phys. Rev. X **11**, 011020 (2021).