Multiple-basis representation of quantum states

Adrián Pérez-Salinas,[1,2] Patrick Emonts,[1,2] Jordi Tura,[1,2] and Vedran Dunjko[1,3]

[1]$\langle aQa^L \rangle$ Applied Quantum Algorithms, Universiteit Leiden
[2]Instituut-Lorentz, Universiteit Leiden, Niels Bohrweg 2, 2333 CA Leiden, Netherlands
[3]LIACS, Universiteit Leiden, Niels Bohrweg 1, 2333 CA Leiden, Netherlands

Classical simulation of quantum physics is a central approach to investigating physical phenomena. Quantum computers enhance computational capabilities beyond those of classical resources, but it remains unclear to what extent existing limited quantum computers can contribute to this enhancement. In this work, we explore a new hybrid, efficient quantum-classical representation of quantum states, the multiple-basis representation. This representation consists of a linear combination of states that are sparse in some given and different bases, specified by quantum circuits. Such representation is particularly appealing when considering depth-limited quantum circuits within reach of current hardware. We analyze the expressivity of multiple-basis representation states depending on the classical simulability of their quantum circuits. In particular, we show that multiple-basis representation states include, but are not restricted to, both matrix-product states and stabilizer states. Furthermore, we find cases in which this representation can be used, namely approximation of ground states, simulation of deeper computations by specifying bases with shallow circuits, and a tomographical protocol to describe states as multiple-basis representations. We envision this work to open the path of simultaneous use of several hardware-friendly bases, a natural description of hybrid computational methods accessible for near-term hardware.

I. INTRODUCTION

Classical simulation of quantum processes is broadly utilized approach to investigate physics, such as material discovery [1], electronic structure problems [2] or chemistry [3]. Through the years, a large variety of methods has been developed to efficiently simulate certain classes of quantum states, such as tensor networks [4, 5], the stabilizer formalism [6–8], or decision diagrams [9] among others [10–12]. ulation methods pushes the borders of current knowledge to represent quantum states using classical resources, but there still exist quantum states of interest that elude an efficient characterization by existing classical methods.

Quantum computers offer new computational capabilities as compared to classical machines. As a motivation, quantum computing allows, in principle, the simulation quantum processes beyond the possibilities of classical computers [13]. In addition, there exist certain problems, outside the domain of physics,where it is expected that the best classical algorithms will always require super-polynomial runtime in the size of the problem, and that have been proven to become efficiently solvable with quantum resources [14–18]. Nowadays existing prototypes of quantum devices are pushing the frontier of addressable quantum computations, in the quest for quantum advantage [19–21]. However, these experimental devices still suffer from limitations in qubit numbers, depth or fidelity of operations. Recent research has explored the combination of limited quantum resources with classical algorithms to harness the full potential of existing quantum devices [22–26].

In this work, we explore a new representation of quantum states built on the idea of leveraging limited quantum resources for reaching certain classically inaccessible computations. We call our construction multiple-basis representation (MBR). The ingredients for the MBR are a $\mathcal{O}(\text{poly}(n))$-sized set ($n$ being the number of qubits) of classical parameters, and a set of $\mathcal{O}(\text{poly}(n))$-sized quantum circuits. Each such quantum circuit determines a basis. The MBR is built as a superposition of states from each of these bases, each state admitting sparse descriptions, i.e. with $\mathcal{O}(\text{poly}(n))$ many coefficients, in the corresponding bases. Given these coefficients and the quantum circuits specifying the bases, access to a quantum computer is required to compute relevant quantities of the state of interest, for instance energies, using the MBR. The MBR then relies on a classical combination of quantum resources. The MBR enables a systematic recipe to obtain approximations of states of interest, such as ground states. Additionally, it is possible to relax the conditions on the quantum resources of MBR until making the computation accessible to classical computer, at the expense of losing generality. Therefore, as a byproduct of our construction, we obtain a novel fully classical method to represent quantum states.

The goal of MBR is to exhibit strictly more representation capabilities than single-basis approaches. These representation capabilities will be influenced by the choice of bases in the MBR. We explore the conditions these bases must fulfill to deliver MBRs with the following properties. (i) As a first requirement, the bases should be chosen such that each of them delivers some information of the global state that is inaccessible to all other considered bases (maintaining the sparsity constraint). (ii) Additionally, the MBR should be made inaccessible to classical simulation methods. The MBR carries a natural notion of optimality with respect to condition (ii), which we show it is met when the chosen basis are MUBs.

We show how MBR can be used in several tasks of interest. We explore first the approximation of ground states of physically motivated Hamiltonians. We as-

sume that the Hamiltonian admits the form $H = \sum_i h_i$ such that each term h_i can be easily diagonalized, and $[h_i, h_j] \neq 0$ (possibly with different bases). The results from these individual terms h_i can be combined through the MBR formalism, yielding approximations to the ground state. Our numerical study suggests the MBR method could be used to approximate ground states of many-body physics problems on arbitrary graphs with an advantage over tensor networks approaches. The second application is the simulation of deeper circuits with shallow computations, as a follow up of Ref. [26]. The MBR formalism allows for a further reduction in depth of the circuits of interest. Finally, we propose a tomography protocol of quantum states using the MBR form. This tomography protocol is related to a relaxed notion of peaked circuits. Peaked circuits are interesting, since they admit efficient verification of quantum advantages [27]. These three applications enhance the computational capabilities of hybrid quantum classical computations, following the recently open research path on classical combination of limited quantum computations [22, 23, 25].

This paper is organized as follows. Section II contains the necessary background for the rest of the manuscript. In Section III, we present the MBR and define several subclasses of MBR, exploring their relationships among them and with other classical representations of quantum states, and optimal choices of bases. Three distinct applications, ground state approximation, optimization of quantum algorithms, and tomography, are described in Section IV. Finally, we conclude in Section V.

II. PRELIMINARIES

A. Classical representations of quantum states

We begin our discussion by looking at numerous classical representations of pure quantum states from the perspective of various *ranks* that will become clear from context. The rank of a quantum state or unitary, with respect to certain decomposition criteria, is the minimal number of components needed to exactly represent the state of interest in that particular decomposition. Classical representability of quantum states crucially depends on the concept of rank. By convention, for a quantum state $|\psi\rangle$ to admit an efficient classical representation, there necessarily exist a decomposition in which a) $|\psi\rangle$ has a $\mathcal{O}\left(\text{poly}\left(n\right)\right)$ rank, and b) each component in which $|\psi\rangle$ is decomposed admits an efficient classical representation.

As an introduction to the arguments in the main paper we describe in this section three classical representations of quantum states: sparse states, matrix product states (MPS) and stabilizer representations of quantum states. These methods are well-known simulation methods for quantum computing, but the list is far from exhaustive. Approaches such as $\mathfrak{g}$-sim [10], LOWESA [28], specifi-

cally designed for variational algorithms, or combinations between stabilizer and tensor networks methods [29] are also available in literature.

1. Sparse representation

The first rank we discuss is the computational-basis (CB) rank of a quantum state, measuring how sparse a state is with respect to the computational basis.

Definition 1 (CB_ε-rank [26]). *Let $|\psi\rangle$ be an n-qubit quantum state, and let $\varepsilon > 0$. The ε-approximate CB rank of a state $|\psi\rangle$ is*

$$CB_\varepsilon(|\psi\rangle) := \min_{K \in \mathbb{N}} \left(K \text{ s.t. } \max_{|\psi_K\rangle \in \Phi_K} |\langle \psi | \phi_K \rangle|^2 \geq 1 - \varepsilon \right),$$
(1)

with

$$\Phi_K = \left\{ |\phi_K\rangle \ \middle| \ |\phi_K\rangle = \sum_{\substack{i \in S \\ |S| = K}} \alpha_i |i\rangle \right\},$$
(2)

and $S \subset \{0, 1, 2, \ldots, 2^n - 1\}$.

It can be further shown that the optimal $|\phi_K\rangle$ in Definition 1 is obtained by retaining the K elements in $|\psi\rangle$ with largest amplitudes and discarding everything else, with proper normalization [26]. Note that it is possible to extend the definition of the CB_ε-rank to the rank in any other basis. Consider the basis $\{U |i\rangle\}_i$, with U unitary and $|i\rangle$ being the elements of the computational basis. Then, the CB_ε-rank in the basis specified by U is just $CB_\varepsilon(U^\dagger |\psi\rangle)$. In the remainder of the paper $CB(|\psi\rangle) \equiv CB_{\varepsilon=0}(|\psi\rangle)$. This extends to other ranks. States of the form $|\phi_K\rangle$ are referred to as K-sparse states.

Small CB_ε-ranks enable classical simulations of quantum computations by storing only sparse states. Consider a quantum circuit $U = U_L U_{L-1} \ldots U_2 U_1$, where U_l are arbitrary $\mathcal{O}(1)$-local gates, such that for all $l \leq L$, $CB_\varepsilon(U_l \ldots U_2 U_1 |0\rangle) \in \mathcal{O}\left(\text{poly}\left(n\right)\right)$. Under this assumption, the simulation can be conducted by updating the sparse representation of the state at each layer. We refer to this concept as sparse simulation.

Definition 2 (SPARSE simulation). *Let $U = U_L U_{L-1} \ldots U_2 U_1$ be a quantum circuit of n qubits, such that U_l are arbitrary operations acting on $\mathcal{O}(1)$ qubits. U is SPARSE -simulatable with precision ε and with respect to an observable $\hat{O}$ if*

$$CB_\varepsilon(U_l \ldots U_2 U_1 |0\rangle) \in \mathcal{O}\left(\text{poly}\left(n\right)\right), \quad \forall l \leq L, \quad (3)$$

and the SPARSE representation of $U |0\rangle$ permits to efficiently compute expectation values of this state with respect to the observable $\hat{O}$.

2. Matrix product state representation

The Schmidt rank is our second rank to review, closely related to matrix product states. The Schmidt rank provides a natural decomposition for studying correlations between quantum subsystems, as well as a robust framework for understanding bipartite entanglement. To define the Schmidt rank, we need to first partition the available Hilbert space into two subspaces, as $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$, and then perform the Schmidt decomposition.

Definition 3 (Schmidt rank). *Let $|\psi\rangle$ be an n-qubit quantum state, and let $\varepsilon > 0$. Let the Hilbert space be decomposed into two partitions $(A, \bar{A})$, as $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_{\bar{A}}$. The ε-approximate Schmidt rank of a state $|\psi\rangle$ is*

$$\chi_\varepsilon^{(A)}(|\psi\rangle) = \min_{K \in \mathbb{N}} \left(K \text{ s.t. } \max_{|\phi_K\rangle \in \Phi_K} |\langle\psi|\phi_K\rangle|^2 \geq 1 - \varepsilon \right), \tag{4}$$

with

$$\Phi_K = \left\{ |\phi_K\rangle \,\middle|\, |\phi_K\rangle = \sum_{k=1}^{K} \lambda_k |u_k\rangle_A \otimes |v_k\rangle_{\bar{A}} \right\}, \tag{5}$$

and $\lambda_k \geq 0$, $\langle v_k|v_l\rangle = \delta_{k,l}$, $\langle u_k|u_l\rangle = \delta_{k,l}$.

Notice that the Schmidt rank may vary depending on the partition A. It is upper-bounded by $\chi_{(A)} = 2^{\min(|A|,|\bar{A}|)}$.

The Schmidt decomposition has a direct application in simulation of quantum states and circuits via tensor networks [5], and in particular matrix product states (MPS) [4]. An MPS is a representation of quantum states in which the qubits are ordered in a one dimensional chain. The partitions A, B are defined by establishing a border between *left* and *right*. An MPS is constructed by chaining Schmidt decompositions in this order, as

$$|\psi_{\text{MPS}}\rangle := \sum_{i_1 \ldots i_n} \sum_{j_1 \ldots j_n}$$
$$A_{j_1}^{(1),i_1} A_{j_1,j_2}^{(2),i_2} \ldots A_{j_{n-1},j_n}^{(n-1),i_{n-1}} A_{j_n}^{(n),i_n} |i_1 i_2 \ldots i_{n-1} i_n\rangle, \tag{6}$$

with $A_{j_{k-1}j_k}^{(k),i_k}$ being a tensor of size $d \times D \times D$, with d being the local dimension and D being the so-called bond dimension of the MPS. The indices $i \in \{1, 2, \ldots d\}$ are related to output states, usually in the computational basis. The indices j are associated to the correlations among partitions. For the MPS to be an efficient classical representation of a quantum states $D \in \mathcal{O}(\text{poly}(n))$. For many interesting cases, this constraint still provides an accurate approximation for MPS to constitute an useful technique.

Definition 4 (MPS simulation). *Let $U = U_L U_{L-1} \ldots U_2 U_1$ be a quantum circuit on n qubits, where U_l are arbitrary gates acting on $\mathcal{O}(1)$ gates. U is MPS simulatable with precision ε and with respect to an observable $\hat{O}$ if for any bipartition $(A, \bar{A})$*

$$\chi_\varepsilon^{(A)}(U_l \ldots U_2 U_1 |0\rangle) \in \mathcal{O}(\text{poly}(n)), \quad \forall l \leq L, \tag{7}$$

and the MPS *representation of $U |0\rangle$ permits to efficiently compute expectation values of this state with respect to $\hat{O}$.*

Computing expectation values through MPS is possible as long as the observable of interest admits a Schmidt decomposition with $\chi_\varepsilon^{(A)} \in \mathcal{O}(\text{poly}(n))$, what is usually called a matrix product operator, for instance for Pauli strings. The MPS method allows for classical simulations of quantum circuits with nearest-neighbor connectivity and $\mathcal{O}(\log n)$ depth.

3. Stabilizer representation

Finally, we address the stabilizer rank, naturally connecting quantum states with the stabilizer formalism [6]. The stabilizer rank captures the optimal decomposition of state $|\psi\rangle$ into a superposition of states that are outputs of quantum circuits composed only of Clifford gates, that is the group of unitary operations that stabilizes the Pauli group.

Definition 5 (Stabilizer rank). *Let $|\psi\rangle$ be an n-qubit quantum state, and let $\varepsilon > 0$. The ε-approximate stabilizer rank of a state $|\psi\rangle$ is*

$$\sigma_\varepsilon(|\psi\rangle) := \min_{K \in \mathbb{N}} \left(K \text{ s.t. } \max_{|\phi_K\rangle} |\langle\psi|\phi_K\rangle|^2 \geq 1 - \varepsilon \right), \tag{8}$$

with

$$|\phi_K\rangle = \sum_{k=1}^{K} \beta_k U_k |0\rangle, \tag{9}$$

and U_k being circuits composed only by Clifford gates.

The stabilizer rank is hard to compute [8], but it can be easily bounded with the number of non-Clifford gates needed to prepare $|\psi\rangle$. The reason is that Clifford circuits allow one to perform quantum computations by keeping track of the updates in Pauli string stabilizing the initial state [6]. These operations are classically accessible. When a non-Clifford gate appears, e.g. a T-gate, one can extend the stabilizer formalism to decompose the gate as a superposition of Clifford operations, thus effectively increasing $\sigma_\varepsilon(|\psi\rangle)$ by a constant factor. Hence, circuits with $\mathcal{O}(\log n)$ non-Clifford gates admit classical simulations [6].

Definition 6 (STAB simulation). *Let $U = U_L U_{L-1} \ldots U_2 U_1$ be a quantum circuit of n qubits, where U_l act on at most $\mathcal{O}(1)$ qubits. U is STAB simulatable with precision ε and with respect to an observable $\hat{O}$ if U has at most $\mathcal{O}(\log T)$ non-Clifford gates, and $\hat{O}$ is decomposable on at most $\mathcal{O}(\text{poly}(n))$ Pauli strings. Under that assumption,*

$$\sigma_\varepsilon(U_l \ldots U_2 U_1 |0\rangle) \in \mathcal{O}(\text{poly}(n)), \quad \forall l \leq L. \tag{10}$$

B. Mutually unbiased bases

Throughout this paper the concept of mutually unbiased bases (MUB) will become useful to specify the a subser of MBR with optimal properties in some sense, to be detailed. MUBs are collections of bases such that any element of any basis can only be written as an equal superposition of all elements of other basis. The formal definition is as follows.

Definition 7 (Mutually unbiased bases [30]). *Let* $\{U_b\}$ *be a set of* $N \times N$ *unitary matrices, each one defining the basis* $\{U_b |i\rangle\}$, *for* $|i\rangle$ *in the computational basis. The set* $\{U_b |i\rangle\}$ *is a set of mutually unbiased basis if*

$$\forall (i,j); \forall (a,b), a \neq b, \qquad |\langle j| U_a^\dagger U_b |i\rangle|^2 = \frac{1}{N}. \quad (11)$$

Sets of MUBs in a N-dimensional system have cardinality bounded by $2 \leq |\{U_b\}| \leq N + 1$ [31]. If the upper bound can be saturated, then the MUBs form a complete set. Existence of complete sets of MUBs are guaranteed if N is a prime power, but remains an open problem for general N [32]. There exist numerous ways to find MUBs [31, 33–36]. It is in fact possible to construct complete sets of MUBs using only Clifford circuits [34], thus thus further motivating the study of the STAB representation of quantum states. Let us remark that there exist non-Clifford constructions for complete sets of MUBs, for instance through the bases defined as VU_b, where $\{U_b\}$ are MUBs and V is an arbitrary non-Clifford unitary.

III. MULTIPLE-BASES REPRESENTATION OF QUANTUM STATES

The multiple-basis representation (MBR) of a quantum state is the central element of this work. The MBR aims to approximately describe an arbitrary quantum state with a few elements combining classical resources and quantum operations.

Inspired by a combination of the CB_ε- and STAB -ranks, we define the MBR as follows.

Definition 8 (Multiple-basis representation (MBR)). *Let* $\{U_b\}_{b=1}^B$ *be a set of quantum circuits,* $\{\alpha_b\}$ *be a set of real positive values,* $\{c_{i_b}^{(b)}\}$ *a set of complex coefficients and* $\{S_b\}$ *a set of supports of cardinality* K, *independent for each basis. The values* $K, B \in \mathcal{O}(\text{poly}(n))$. *A MBR state is given by*

$$|\psi_{K,B}\rangle = \sum_{b=1}^B \alpha_b U_b |\psi_b\rangle \quad (12)$$

$$\text{with} \quad |\psi_b\rangle = \sum_{\substack{i_b \in S_b \\ |S_b|=K}} c_{i_b}^{(b)} |i_b\rangle \quad (13)$$

$$\text{and} \quad \langle \psi_b | \psi_b \rangle = 1, \quad (14)$$

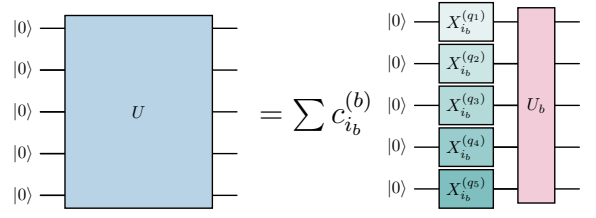*with* $|i_b\rangle$ *being elements of the computational basis.*

FIG. 1: A motivation for the MBR is to describe states using superpositions of quantum circuits that are more easily implementable. See Definition 8 for further details.

The interpretation of a MBR states can be done as follows. We start with the states $|\psi_b\rangle$. Those will be states with low CB_ε-ranks, and thus have limited power to represent all quantum states. The role of U_b is to perform a unitary operation that changes the basis in which the state is sparse. Each new term in Equation (12), noted with the index b, will add a new state which is sparse in a completely new basis U_b, but not sparse in any of the previously considered bases. The superposition of sparse states in different bases is intended to create a new state which is not sparse in any of the previous bases, and hence does not admit an efficient representation. On the other hand, there must exist bases such that $|\psi_{K,B}\rangle$ is sparse in these bases, even with unit CB_ε-rank. However, we expect that constructing these bases will require much deeper circuits than each of the individual unitaries U_b. A graphical interpretation of MBR is available in Figure 1.

We now detail the requirements to fully specify a MBR. The unitary gates $\{U_b\}$ must be given in an efficient form, for instance through a $\mathcal{O}(\text{poly}(n))$-sized circuit. For each basis we need $\mathcal{O}(K)$ memory to store (up to digital precision) α_b, $\{c_{i_b}^{(b)}\}$, and the set S_b. By convention we will order S_b so that $|c_{i_b}^{(b)}|^2 \geq |c_{i_b+1}^{(b)}|^2$. Notice that these elements are not sufficient to fully specify the state due to the linear dependence of its different components. An additional necessary component will be the overlap matrix

$$F_{i_b, j_a} = \langle i_b | U_b^\dagger U_a | j_a \rangle. \quad (15)$$

Whether or not this Gram matrix can be computed classically determines whether the MBR can be treated as a fully classical or hybrid classical-quantum representation, although its storage is always possible, by construction. This matrix has a clear structure of $B \times B$ blocks, each of size $K \times K$[1]. The blocks in the diagonal are just the identity matrix relating elements within the same basis.

The motivation to create the MBR can be easily outlined in terms of the CB_ε-rank. Take an arbitrary state

---

[1] The sizes of these blocks may change in the case the support of each basis has a different number of elements.

$|\psi\rangle$, and a bases U. By definition, there are certain bases in which $CB_\varepsilon(U|\psi\rangle)$ is a large quantity, and some others in which it is small, or even minimal, see Ref. [26] for an in-depth discussion. There exists a trade-off between allowed CB_ε-rank and number of gates needed to implement a proper basis. In the MBR representation, we aim to reduce this requirement. Each basis is now asked to capture a fraction of the quantum state within a sparse approximation, that is small CB_ε-rank for moderately large ε. The combined action of all bases, each capturing a different fraction of the state, gives representation power to the MBR, allowing us to efficiently describe (certain) states that require exponentially many coefficients using only one basis. This change cannot significantly change the volume of ε-approximately representable states, as in the MPS or STAB cases. However, it can make new states reachable within modest quantum resources. We illustrate this idea in Figure 1, a).

A. Example of a MBR

We give an example for a state that offers an efficient MBR representation in this section. Consider the quantum state given by

$$|\psi_{+0}\rangle = \frac{1}{\sqrt{2C}}\left(|0\rangle^{\otimes n} + |+\rangle^{\otimes n}\right), \qquad (16)$$

where $C = 1 + 2^{-n/2}$ is the normalization constant. The CB_ε-rank is given by $CB_\varepsilon(|\psi_{+0}\rangle) = 1 + \max(\varepsilon - C/2, 0)(2^n - 1)$. Consequently, an approximate representation of $|\psi_{+0}\rangle$ is achievable even with CB_ε-rank 1 only with error above approximately $1/2$. More accurate representations demand an exponential CB_ε-rank in n, since adding any extra component $|i\rangle, i \neq 0$ can only increase the fidelity by capturing factor $\langle i|\psi_{+0}\rangle = 2^{-n/2}$, existing $2^n - 1$ possible values for i. Explicitly,

$$CB_\varepsilon(|\psi_{+0}\rangle) =$$
$$1 + \left\lfloor \left(1 - \varepsilon - \frac{1 + 2^{-n/2}}{2}\right)(2^n - 1)\right\rfloor. \quad (17)$$

The Hadamard basis provides no alternative, since $|\psi_{+0}\rangle$ stabilizes $H^{\otimes n}$. Nevertheless, simultaneous use of both the computational and Hadamard basis allows for a sparse representation, as intuited in the explicit writing of Equation (16).

We can explore alternative classical representations for $|\psi_{+0}\rangle$. For instance, it is possible to consider a basis with elements $|0\rangle^{\otimes n}$ and $|+\rangle^{\otimes n} - 2^{-n/2}|0\rangle^{\otimes n}$. Preparation of the second state can be done in depth $\mathcal{O}(n)$ [37] to the required precision $\varepsilon \in \mathcal{O}(2^{-n})$. Alternatively, one can use matrix product states to exactly represent $|\psi_{+0}\rangle$. Such matrix product state has bond dimension 2, and thus it is required to store $\mathcal{O}(n)$ real numbers. Alternatively, one can use a stabilizer representation, with stabilizer rank 2 [7, 8]. The memory requirement for this construction scales also as $\mathcal{O}(n)$, achieved by specifying the Clifford circuits of linear size needed to prepare $|\psi_{+0}\rangle$. If we address the depth required to prepare this state from the $|0\rangle^{\otimes n}$ state, we find $\Omega(\log n)$ through optimal recipes for matrix product states [38]. The MBR in this case is a special case of the stabilizer formalism. The minimal depth required to construct a state with small stabilizer rank is, to the best of our knowledge, an open problem.

B. Comparison to other representations

In this section, we will connect the MBR to other simulation methods of quantum computing, reviewed in Section II A. To do so, we will define several subclasses for the MBR, distinguishing between different properties of the circuits, and will compare them to existing methods. The first element of discussion will be the classical tractability of a MBR. In all this discussion, we restrict ourselves to observables that are Pauli strings or sums of moderately many Pauli strings.

Definition 9 (Classical tractability). *Consider a quantum state $|\psi\rangle$, and an observable $\hat{O}$. Then, $|\psi\rangle$ is classically tractable with respect to $\hat{O}$ if computing the expectation $\langle\psi|\hat{O}|\psi\rangle$ is efficient with classical resources.*

In the context of MBR we can give a more restricted definition.

Definition 10 (Classical tractability of MBR). *Let $\{U_b, S_b, \boldsymbol{\alpha}_b\}$ be respectively the unitaries, integers sets and coefficients defining a MBR, as in Definition 8. The MBR is classically tractable, for an observable $\hat{O}$, if it is possible to classically and efficiently compute*

$$F^{\hat{O}}_{i_b,j_a} = \langle i_b| U_b^\dagger \hat{O} U_a |j_a\rangle. \qquad (18)$$

Note that classical tractability of MBR immediately implies that computing expectations values $\langle\psi_{K,B}|\hat{O}|\psi_{K,B}\rangle$ becomes classically feasible since it is just a combination of $\mathcal{O}(\text{poly}(n))$-many elements of $F^{\hat{O}}_{i_b,j_a}$.

Lemma 1. *Classical tractability of the MBR implies classical tractability.*

The proof is available in Appendix 1

The MBR definition allows for several subclasses of MBR, depending on the properties of the circuits U_b. We define four different classes.

Definition 11 (MBR - Classical (MBR_C)). *We define MBR_C to be the set MBR_C consists of all classically tractable MBRs.*

Definition 12 (MBR - Non classical ($\text{MBR}_{\not C}$)). *We define $\text{MBR}_{\not C}$ to be the set of all classically intractable MBRs. For $\text{MBR}_{\not C}$, $F^{\hat{O}}_{i_b,j_a}$ is efficiently computable with quantum resources.*
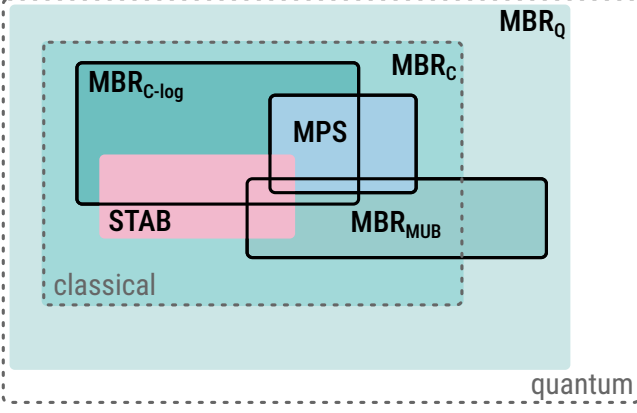
FIG. 2: Known inclusion relationships of MBR and its subclasses with respect to other classical representations of quantum states. This diagram is a graphical description of Theorem 1.

Definition 13 (MBR - classical logarithmic (MBR$_{C-\log}$))**.** *We define* MBR$_{C-\log}$ *to be a subset of* MBR$_C$ *with the constraint that the depth of* $\{U_b\}$ *is* $\mathcal{O}(\log n)$.

Definition 14 (MBR - MUBs (MBR$_{\text{MUB}}$))**.** *We define* MBR$_{\text{MUB}}$ *to be the subset of* MBR$_{\mathscr{C}}$ *with the constraint that the set of unitaries* $\{U_b\}$ *form a set of MUBs.*

The definitions here given, together with the representations given in Section II A, allow us to detail the relationships depicted in Figure 2, as a pictorial representation of Theorem 1.

Theorem 1. *The relationship among the MBR subclasses and other representations are the following:*

$$\text{STAB} \subsetneq \text{MBR}_C \tag{19}$$

$$\text{MPS} \subsetneq \text{MBR}_C \tag{20}$$

$$\text{MBR}_{\text{MUB}} \cap \text{STAB} \neq \emptyset \tag{21}$$

$$\text{MBR}_{\text{MUB}} \cap (\text{MBR}_{\mathscr{C}} \cap \text{MBR}_C) = \emptyset \tag{22}$$

The proof can be found in Appendix A.

C. Optimality of MBR$_{\text{MUB}}$

In this section we explore certain properties that are desirable when choosing the bases that will be used for the MBR, and argue that MBR$_{\text{MUB}}$ is optimal with respect to these properties. Intuitively, we would like each basis to capture a different fraction of the quantum state of interest, that is,

$$\langle \psi_a | U_a^\dagger U_b | \psi_b \rangle \approx 0. \tag{23}$$

The reason is that if two bases address the same components of the state, then we encounter redundancy. For a second condition, we require the multiple bases to be

necessary, i.e., it is not possible to mimic the effect of a new basis with a modest increment of K. We focus on the bases, therefore we give freedom to choose S_b and coefficients α_b, $\boldsymbol{c}^{(b)}$ specifying $|\psi_b\rangle$. The concept of overlaps among bases hinted in Equation (23) is then formalized as follows.

Definition 15 ((U, K)-fidelity)**.** *Let U be a unitary operation specifying a change of basis. The states $|\psi_1\rangle$ and $|\psi_2\rangle$ are any two with CB-rank upper bounded by K. The (U, K)-fidelity is given by*

$$\mathcal{F}_{K,U} = \max_{\substack{|\psi_1\rangle, |\psi_2\rangle \\ CB(|\psi_1\rangle) \leq K \\ CB(|\psi_2\rangle) \leq K}} |\langle \psi_1 | U | \psi_2 \rangle|. \tag{24}$$

We discuss now the conditions under which $\mathcal{F}_{(U,K)} = 0$. Consider first two arbitrary states with supports S_1, S_2 in the computational bases. The changes of bases U that satisfies this condition are those that fulfill $U|i\rangle = |j\rangle$, for $i \in S_1, j \notin S_2$. That is, the supports are moved to orthogonal subspaces. Certain permutations meet this requirement. However, MBRs where U_b are permutations are trivially reduced to a single-basis representation with larger K. Another allowed kind of basis changes are block diagonalizable operations U. For illustrative purposes, we consider two extremal cases, (a) U is diagonalizable in $\mathcal{O}(\exp(n))$ blocks of $\mathcal{O}(\text{poly}(n))$ size, (b) U is diagonalizable in $\mathcal{O}(\text{poly}(n))$ blocks of $\mathcal{O}(\exp(n))$ size. In the case (a), the MBR can be substituted by a single-basis representation, with $\mathcal{O}(\text{poly}(n))$ increment in K. The exponential scaling is never seen, since there are *only* polynomially many elements. On the contrary, the case (b) leads to MBRs that cannot be reduced to singular bases. However, this implies that only a small number of supports will be capable of satisfying Equation (23). This reasoning provides the intuition that: if $\langle \psi_1 | U | \psi_2 \rangle = 0$, then it is likely that the MBR can be substituted by single-basis representation with a moderate increase in the CB_ε-rank.

We formulate the following statement inspired by the previous reasoning, hinting for a tension between reducing $\mathcal{F}_{(U,K)}$ and maintaining the condition in Equation (23).

Lemma 2 (Optimal choice of bases)**.** *Let $\mathcal{F}_{K,U}$ be the basis-dependent sparse-fidelity. Let $\mathcal{F}_K$ be*

$$\mathcal{F}_K = \min_U \mathcal{F}_{K,U}, \tag{25}$$

for U unitary. $\mathcal{F}_K$ is minimized if U satisfies the condition

$$|\langle i | U | j \rangle|^2 = \frac{1}{2^n} \; \forall (i, j), \tag{26}$$

with

$$\mathcal{F}_K \leq \frac{K}{\sqrt{2^n}}. \tag{27}$$

The proof can be found in Appendix B. The condition given by Lemma 2 is known as mutually unbiased bases (MUB), see Section II B.

Properties of MBR with MUBs

The conditions from Lemma 2 allows us to distill some properties of the $\mathrm{MBR_{MUB}}$ of a quantum states. All subsequent properties can be proven using the following following observation. When $\{U_b\}$ form a set of MUBs, then the low-CB_ε-rank state corresponding to each bases has an exponentially vanishing overlap with all other bases. This means that new basis in the MBR adds an exponentially small — but non zero — redundancy in the previous steps.

Lemma 3 (Accuracy of a MBR). *Let $|\psi_{B,K}\rangle$ be the MBR representation of a state $|\psi\rangle$, with bases $\{U_b\}$, and let $|\psi_b\rangle$ be the optimal approximation of $U_b^\dagger |\psi\rangle$ with CB-rank K. The bases $\{U_b\}_b$ form a set of MUBs. The overlap achievable by the MBR is bounded by*

$$\max_{\{\alpha_b\}} |\langle \psi_{B,K}|\psi\rangle| \leq \sqrt{\sum_b |\langle \psi_{b,K}|\psi\rangle|^2 \left(1 + \frac{KB}{2\sqrt{2^n}}\right)} \quad (28)$$

$$\max_{\{\alpha_b\}} |\langle \psi_{B,K}|\psi\rangle| \geq \sqrt{\sum_b |\langle \psi_{b,K}|\psi\rangle|^2 \left(1 - \frac{KB}{2\sqrt{2^n}}\right)}, \quad (29)$$

and it is obtained by choosing $\alpha_b \propto |\langle \psi_b|\psi\rangle|$, with proper normalization.

The proof can be found in Appendix C.

We can also study the dimensionality of the space spanned by the MBR.

Lemma 4 (Dimensionality). *Let $\{U_b\}$ be a set of unitary operations that transforms the computational basis into other bases, such that those bases are pairwise MUBs, and let $\{S_b\}$ be a set of sets of states in the computational basis, with $|S_b| = K$. It holds that the linear space spanned by*

$$\{U_b |i_b\rangle\}_{\substack{i_b \in S_b \\ b \in \{1,\dots B\}}}$$

has dimension KB as long as

$$B \leq \frac{\sqrt{2^n}}{K} + 1. \quad (30)$$

The proof can be found in Appendix D.

Finally, we address first the difference in the volume of states that can be approximated in the single- and multiple-basis representations.

As an auxiliary result, we begin by computing the relative volume of states that can be approximated by a K sparse state with fixed support.

Lemma 5. *Let $|\psi\rangle$ be an n-qubit random quantum state drawn from the Haar distribution, and let $|\phi_S\rangle$ be any quantum state of the form*

$$|\phi_S\rangle = \sum_{i \in S} c_i |i\rangle, \quad (31)$$

where S is a fixed set of size $K \in \mathcal{O}\left(\mathrm{poly}\,(n)\right)$. Then

$$\mathrm{Prob}\left(\max_{|\phi_S\rangle} \left(|\langle\psi|\phi_S\rangle|^2\right) \geq 1 - \varepsilon\right) = $$
$$I\left(\varepsilon; 2^n - K, K\right), \quad (32)$$

with $I(x; a, b)$ being the incomplete regularized beta function.

The proof can be found in Appendix E, and a detailed description of $I(\cdot; \cdot, \cdot)$ is available at Appendix F. The idea behind this proof is as follows. A quantum state can be mapped to a sphere of dimension 2^{n+1}, and a quantum state with CB_ε-rank K to another sphere of dimension $2K$. The calculation consists in computing the volumen of a ε-ball of dimension $2^{n+1} - K$ around any state with CB_ε-rank K, and normalizing appropriately. For the interest of the paper, we refer the approximation

$$I\left(\varepsilon; 2^n - K - \frac{1}{2}, K - \frac{1}{2}\right) \approx $$
$$\Phi\left(\varepsilon; \frac{2^n - K}{2^n}, \frac{(2^n - K)(K)}{(2^n + 1)2^{2n}}\right), \quad (33)$$

with $\Phi(\cdot; \mu, \sigma^2)$ being the cumulative distribution function of a Gaussian distribution with mean μ and variance σ^2 [39].

For completeness, we provide the intuitively more comprehensible bound, valid for $K \leq (1 - \varepsilon)2^n$,

$$\mathrm{Prob}\left(\max_{|\phi_S\rangle} \left(|\langle\psi|\phi_S\rangle|^2\right) \geq 1 - \varepsilon\right) \in $$
$$\mathcal{O}\left(\exp\left(-2^n \left(\varepsilon - \frac{2^n - K}{2^n}\right)^2\right)\right). \quad (34)$$

This bound comes from connecting the function $I(x; a, b)$ to the binomial distribution and applying Hoeffding's inequality to such binomial distribution. We leverage the result to compute the fraction of states that can be represented within ε accuracy using the MBR.

Corollary 1. *Let $|\psi\rangle$ be an n-qubit random quantum state drawn from the Haar distribution, and let $|\psi_{K,B}\rangle$ be its optimal MBR. The B bases are chosen from a set of M different MUBs. Then, for $K \leq (1 - \varepsilon)2^n$,*

$$\mathrm{Prob}\left(|\langle\psi|\psi_{K,B}\rangle|^2 \geq 1 - \varepsilon\right) \in $$
$$\mathcal{O}\left(\exp\left(-2^n \left(\varepsilon - \frac{2^n - K}{2^n}\right)^2\right) C_M(K, B)\right). \quad (35)$$

with

$$C_M(K, B) = \binom{M}{B}\binom{2^n}{K}^B \quad (36)$$

The proof follows immediately from Lemma 5 together with two considerations. First, the MBR spans in a linear subspace with dimension KB. Second, there are M

different MUBs, from which we choose B of them. However, the M bases do not count any ordering, and therefore there are $\binom{2^n}{K}$ different choices for K-sparse states for each basis. The overlaps among volumes are absorbed in the upper bound.

It is interesting to compare the bounds in the probabilities given by Corollary 1 for extremal cases, in which $K = 1$ or $B = 1$ while keeping the same number of degrees of freedom. We obtain the quotients

$$\frac{C_M(K,B)}{C_M(KB,1)} \leq M^{B-1} \frac{(KB)^{KB}}{B!(K!)^B} \tag{37}$$

$$\frac{C_M(K,B)}{C_M(KB,1)} \geq M^{B-1} \frac{(KB)!}{(BK^K)^B} \tag{38}$$

$$\frac{C_M(K,B)}{C_M(1,KB)} \leq \frac{M^B}{M^{K-1}} \frac{(KB)^{KB}}{B!(K!)^B} \tag{39}$$

$$\frac{C_M(K,B)}{C_M(1,KB)} \geq \frac{M^B}{M^{K-1}} \frac{(KB)!}{(BK^K)^B} \tag{40}$$

These numbers indicate that the volume of representable states increases if the number of considered bases B increases.

We give an interpretation of the results obtained in this section. The bounds given by Lemma 5 and subsequent approximations state that the volume of captured states rapidly increases at the transition point $\varepsilon \approx 2^{-n}(2^n - K)$, yielding exponentially vanishing relative volumes for reasonable $K \in \mathcal{O}(\text{poly}(n)), 1 - \varepsilon \in \Omega(\text{poly}^{-1}(n))$. This observation aligns with the common observation that no classical representation of states can reach all (or a significant fraction) of quantum states. This, however, does not contradict the results on classical (in)tractability detailed in this manuscript.

IV. APPLICATIONS OF MBR

Succinct representations of quantum states acquire importance depending on their utility to conduct certain tasks. Tensor networks, and in particular MPS, have been broadly utilized to investigate ground states of many-body physics. The stabilizer formalism found great applicability in quantum error correction [40]. In order to highlight the potential of MBR, we outline in this section three different applications to it.

A. Ground-state approximation via MBR

The MBR can be used as a tool to obtain classical approximations of ground states of Hamiltonians, by leveraging quantum subspace expansions [41]. Notice that finding the ground states even for physically motivated Hamiltonians is QMA-complete [42], hence proposals for heuristic approaches becomes practically relevant. In the case we deal with MBR_C then this approximation is also classically efficient to compute.

We assume a decomposition of the Hamiltonian as

$$\hat{H} = \sum_c \hat{h}_c, \tag{41}$$

such that each term $\hat{h}_c$ can be diagonalized in a basis defined by U_c, that is

$$\hat{h}_c U_c \left|i\right\rangle = \lambda_i U_c \left|i\right\rangle, \tag{42}$$

for $i \in \{0, \ldots 2^n - 1\}$. Starting from this point, we want to obtain the best possible MBR approximation to the ground state. For that, we consider $\{U_c\}$ as our set of unitary operations, and develop a constructive method to find the coefficients $\boldsymbol{c}^{(b)}, \alpha_b$, as well as the sets S_b.

We assume first that S_b is fixed. Then, the Hamiltonian as evaluated in the considered subspace is given by

$$\hat{H}_{i_b,j_a} = \sum_c \left\langle i_b\right| U_b^\dagger \hat{h}_c U_a \left|j_a\right\rangle, \tag{43}$$

where the indices $\{a, b, c\}$ run over the same values, and $\{i_b, j_a\}$ are given by the subsets S_b from the MBR. Since this truncated Hamiltonian has $\mathcal{O}(\text{poly}(n))$ size, the next step is just to diagonalize. However, the basis in which this Hamiltonian is expressed is not orthonormal due to the presence of the unitary matrices U_b, and the relationship between basis elements is encoded in the matrix F from Equation (15). The exact diagonalization can be performed through the generalized eigenvalue problem, where the solution state lives in the truncated space explored by the MBR, that is solving the equation [2]

$$\det\left(\hat{H}_{i_b,j_a} - E F_{i_b,j_a}\right) = 0. \tag{44}$$

Assuming a polynomial MBR_C, solving the generalized eigenvalue problem is classically feasible if H decomposes in polynomially many terms $\hat{h}_c$. On the contrary, assuming $\text{MBR}_{\mathscr{Q}}$, both matrices $\hat{H}_{i_b,j_a}$ and F_{i_b,j_a} must be obtained with the assistance of a quantum computer via Hadamard tests or similar techniques for overlap estimation, while the diagonalization is a classical task. This diagonalization provides exactly the ground state if $\det F = 0$, that is the MBR, for given K and unitaries, can represent any quantum state. This approach yields an upper bound on the energy of the approximation of the ground state of interest captured by our ansatz. This scenario is only possible for few qubits, and MBR is not needed to approximate the ground state.

We address now the question of finding the subspaces in which the approximation to the ground state is optimal, that is the choice of S_b. Searching among all possibilities is a hard problem if no further information is

---

[2] Notice the small abuse of notation by writing in the previous equations the matrices with the same symbol we used to define their components.

provided. There are $\binom{2^n}{K}$ different sets to choose from, and any brute-force search becomes unfeasible. Only informed searches provide tractable scenarios. As a first candidate, one can use heuristic optimization methods to look for sufficiently good $\{S_b\}$, e. g. via genetic algorithms [43]. One can also choose a physics-informed alternative by selecting the lowest-energy states in each of the Hamiltonians H_b, provided this information is efficiently computable with classical resources.

Transverse-Field Ising model

We illustrate the method here described with the example of approximating the ground state of a transverse-field Ising model (TFIM), defined as

$$H_G(J,h) = J \sum_{(i,j)\in E(G)} X_i X_j + h \sum_{i\in N(G)} Z_i \qquad (45)$$
$$\equiv JH_{xx} + hH_z,$$

where G is a graph defining the model and $N(G), E(G)$ are respectively nodes and edges of the graph.

We first choose the bases to apply the MBR. In this case, the bases can be chosen as those diagonalizing H_{zz} and H_z respectively. H_z is diagonalized in the computational basis, thus $U_z = I$, and H_{xx} is diagonalized in the Hadamard basis, thus $U_{xx} = H^{\otimes n}$.

For the subsets $\{S_b\}$, we follow the physical intuition in the considered problem. Since we look for a ground state, we apply the heuristic of keeping those states with the lowest energies in each term $\hat{h}_c$. For the Hamiltonians considered, the energies of a basis element defined by a bitstring can be easily computed. For H_z, one just needs to obtain the Hamming weight. For H_{xx}, the energy is related to the number of cuts of the graph G induced by the bitstring, in the Hadamard basis. Depending on the value of J, the heuristics are different. For the ferromagnetic phase $J < 0$, the ground state of H_{zz} is equivalent to solving the MinCut problem, which has a polynomial cost [44]. The antiferromagnetic phase $J > 0$ needs the heuristic to solve a MaxCut problem, which is NP-hard for arbitrary graphs Heuristic methods exist to find approximations to the solutions in [45, 46], and each approximation found by any method can be used as one element in H_{zz}. In some physics-motivated cases, such as lattice graphs, the MaxCut problem is trivial to solve.

We numerically approximate the ground state of the TFIM [cf. Equation (45)] on a 2D square lattice of size $L \times L$. The considered space includes the ground states of H_{xx} and H_z in the computational and Hadamard basis respectively, and all states up to a fixed Hamming distance D_H to these ground states. The dimensionality of the expanded subspace grows with the system size as $\mathcal{O}\left(\left(L^2/H_W\right)^{D_H}\right)$. The regularity of the lattice allows for the efficient computation of these state. The accuracy is assessed through the comparison between the energies
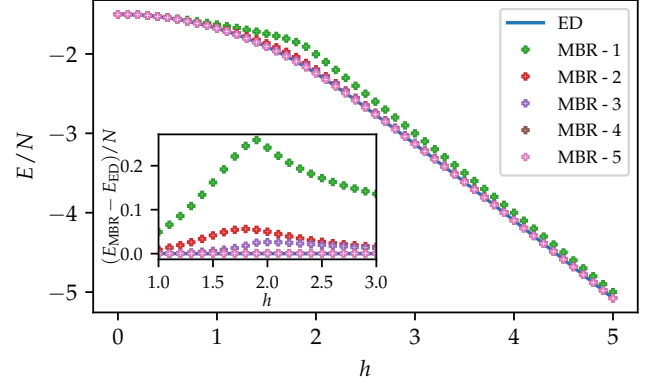

FIG. 3: Different MBR approximations of the ground state of an anti-ferromagnetic Ising model. The system is defined on a square lattice of size 4×4 with open boundary conditions, and $J = 1$. The inset shows the deviation from exact diagonalization data with increasing degree of the MBR approximation in the middle of the parameter range, i.e. the hardest regime for the MBR approximation.
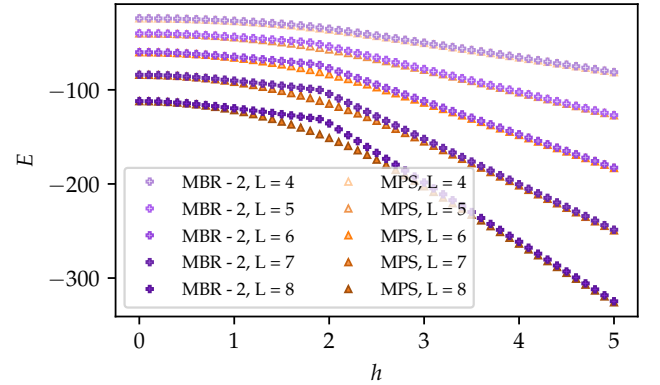

FIG. 4: Finite size effects in the approximation for the TFIM on an $L \times L$ lattice. As the system grows, the deviations around the phase transition between MPS data and MBR data become more pronounced. Larger systems can need a larger MBR basis.

of found states through MBR, and that obtained with exact diagonalization or MPS approximation, depending on the size of the lattice.

First, we benchmark the MBR state against exact results (cf. Figure 3) on a small lattice in the anti-ferromagnetic case ($J = 1$) with varying transverse magnetic field h. The degree of the MBR approximation is given by the dimension of the considered subspace, here depicted as $\mathrm{MBR} - D_H$. The energies of the MBR approximations to the ground state energy match the values obtained by exact diagonalization increasingly well as D_H increases (cf. inset of Figure 3). The MBR approximation is capable of accurately representing the regimes $h \approx 0$ and $h \gg J$ even with of $D_H = 1$. We attribute this

to the choice of the subspace expansion which includes the ground state of each term in the Hamiltonian. In both regimes, high and low transverse field, the ground state becomes more similar to a state in our ansatz.

For the regime around the phase transition $3J \approx h$ [47, 48], the MBR approximation yields less accurate results as compared to the reference value. The accuracy improves as the degree of the MBR is increased, fulfilling the interpretation that the approximation will become a exact representation with enough computational power. In the MBR formulation, the deviation from the MPS simulation happens slightly earlier than the expected formulation. This could be related to a breakdown of the heuristic in this parameter regime, even independent of the actual phase transition.

In Figure 4, we compare the performance of a fixed MBR approximation for increasing system size. While a small MBR degree is sufficient to obtain good estimates for the ground state energy on small lattice sizes, a larger degree becomes necessary when the lattice increases.

It is interesting to analyze the differences between the methods using MBR and MPS to obtain ground states (cf. Figure 4) MBR is entirely based on heuristics over the subspace to consider in the search for the ground state. Candidates to ground states are accumulated, without adding correlations to the selection criteria. On the other hand, MPS are highly dependent on correlations of the ground state to be found, and are successful only in the case of low entanglement. This hinders the applicability of MPS for systems allowing long-distance correlations, such as 2D lattices, or lattices defined by arbitrary and irregular graphs. Therefore, the candidates to ground states given by MBR might be more accurate in those cases, where MPS demand large overheads. This phenomenon can extend to quantum circuits with arbitrary connectivity which do not admit classical simulations. Investigations on these topics is left for future research.

B. MBR and reduce&chop

We inspect now the interplay between MBR and the reduce&chop algorithm [26] to show how using several bases might be useful to further decrease the depth requirement.

1. Background

The reduce&chop method leverages states with small CB_ε-ranks to conduct simulations of deep quantum computations making use of shallower circuits [26]. Consider the problem of computing $p(x) = |\langle x| \mathcal{U} |0\rangle|^2$, where x is a given bitstring and U is a given (deep) quantum circuit.

We can approximate $p(x)$ by

$$p(x) = |\langle x| \mathcal{U} |0\rangle|^2 = \left| \sum_{i=0}^{2^n-1} \langle x| \mathcal{U}_2 |i\rangle \langle i| \mathcal{U}_1 |0\rangle \right|^2$$

$$\approx \left| \sum_{\substack{i \in S \\ |S|=K}} \langle x| \mathcal{U}_2 R^\dagger |i\rangle \langle i| R\mathcal{U}_1 |0\rangle \right|^2 , \quad (46)$$

where S K is the support over bitstrings i, $\mathcal{U} = \mathcal{U}_2\mathcal{U}_1$ and R is a unitary operation called the reducer. For the approximation to be accurate, it is sufficient that the state $R\mathcal{U}_1 |0\rangle$ fulfills $CB_\varepsilon(|\psi\rangle) \in \mathcal{O}(\text{poly}(n))$, with n the number of qubits. Under this assumption, the reduce&chop method provides a recipe to approximate $p(x)$ with classical overhead in $\mathcal{O}(\text{poly}(n))$. The depth requirements to compute $p(x)$ decrease from $d(U)$ to $d(U)/2 + d(R)$ through reduce&chop (assuming $\mathcal{U}$ is splitted in halves), with $d(\cdot)$ being the depth of a quantum circuit.

A fundamental step for reduce&chop to work is to find the reducer R. The function of R is to identify a change of basis under which the state $\mathcal{U}_1 |0\rangle$ has small CB_ε-rank. The hardness of finding the appropriate R hinges on the internal structure of state $\mathcal{U}_1 |0\rangle$. For example, there exist straightforward circuits with constant depth with the capability to transform a state with maximal CB rank into another state with minimal rank. The search for R can be approached variationally [26], and a successful search for R dramatically affects the performance of reduce&chop . Recent advances in peaked states [27], which can be understood as states with CB_ε-rank 1, showed that it is possible to peak a random quantum circuit $d(V)$ by adding a reducer satisfying

$$d(R) \in \Omega \left(\left(\frac{d(U)}{n} \right)^{0.19} \right). \quad (47)$$

in the same domain, it was shown that $O(1)$-depth peaked circuits can be classically and optimally simulated in quasi-polynomial runtime $n^{\mathcal{O}(\log(n))}$ [49], and therefore reduce&chop can bring quantum advantages starting from $d(R\mathcal{U}_1) \in \Omega(1)$.

2. MBR for depth reduction

The reduce&chop method relies on identifying an easily implementable change of basis $R(\theta)$ to achieve a small CB_ε-rank for a given state. The MBR offers a more nuanced representation of a quantum state by employing multiple bases simultaneously, each representing a distinct component of the state of interest. This representation further explores the depth reduction motivating reduce&chop . The MBR is interesting as long as the bases required for it are easy to implement in a quantum computer. Ideally, these bases are designed to be even

simpler than the original $R(\theta)$ from reduce&chop , striking a balance between CB_ε-ranks, number of bases, and the associated cost of basis change.

To clarify this, consider a state $|\psi\rangle$ and a poly-sized set of MUBs $\{U_b\}$ such that $(CB_\varepsilon(U_b\psi)),(1-\varepsilon)^{-1} \in \mathcal{O}(\text{poly}(n))$. By virtue of Lemma 3, the accuracy of the MBR can reach constant values. The action of R needed to obtain sparse representation is now split among several more simple unitaries. The complexity of running each circuit is therefore simpler.

Estimating the depth savings attained from combining the MBR with the reduce&chop method is more complicated. As motivated by the example in Section III A, it is possible in some cases to find bases represented by shallow circuits for which an accurate MBR of the state is obtained, while the basis required to lower the CB_ε-rank of the entire state requires larger depth. In the reduce&chop method, the output of the first half of the circuit must admit an accurate MBR, for $K, B \in \mathcal{O}(\text{poly}(n))$. The results in this section bound the depth needed to attain such states, thus allowing us the to substitute this circuit with classical combinations of quantum circuits of depth at most $d(U)$. If this state serves as input for the reduce&chop algorithm, then the required depth in a quantum circuit to simulate the values of $p(x) = |\langle x| U |\psi_{K,B}\rangle|^2$ is upper bounded by $d(U) + \max_b d(U_b)$, yielded from concatenating the change of basis and the circuit U.

The first result we give is related to lower bounds to create arbitrary n-qubit states with CB-rank K.

Lemma 6 (Lower-bound in creating sparse states). *Arbitrary n-qubit states with CB-rank K require at least m gates in depth d to be prepared in a quantum computer from the $|0\rangle$ state, with*

$$m \in \Omega \left(\frac{K \log(\varepsilon^{-1})}{\log \log K} + n \right) \tag{48}$$

$$d \in \Omega \left(\frac{K \log(\varepsilon^{-1})}{\log K \log \log K} + n \right) \tag{49}$$

The proof immediately follows from Section 4.5.5 in [50] to construct states of dimension K in a register of $\log K$ qubits. The additional term n accounts for arbitrary correlations among qubits.

For the next step, we recall that any state in MBR_{MUB} has dimension at most KB, assuming no redundancy in the bases. Therefore, we can further leverage Lemma 6 to obtain the following result.

Corollary 2 (Lower-bound in MBR). *Let $|\psi\rangle$ be an n-qubit such that it can be ε-approximated with an MBR_{MUB} as in Equation (12), with sparsity K and B different bases. This state requires at least m gates and*

depth d to be created, with

$$m \in \Omega \left(\frac{KB \log(\varepsilon^{-1})}{\log \log(KB)} + n \right), \tag{50}$$

$$d \in \Omega \left(\frac{KB \log(\varepsilon^{-1})}{\log(KB) \log \log(KB)} + n \right). \tag{51}$$

The proof uses a modification of the proof in Section 4.5.5 in [50], which does not rely on any choice of basis, but only on counting dimensions. This result holds but is less tight for arbitrary MBR.

We give now upper bounds by providing a constructive algorithm to prepare MBR states into a quantum computer of the form of Equation (12). This algorithm assumes availability of preparation of states with CB-rank K and at circuit specifications of the bases U_b. Note that black-box query to U_b is enough to prepare $|\psi_{K,B}\rangle$, but descriptions are needed to classically compute relative overlaps among states. The algorithm we propose utilizes the linear combination of unitaries as a subroutine [51]. The first step is to create a state with a non-zero overlap with the desired state $|\psi\rangle$ in the MBR.

Lemma 7. *Let $|\psi\rangle$ be an MBR of a state with sparsity K and number of basis B, where the weight for each basis is given by $\boldsymbol{\alpha}$, as in Equation (12). Let $|\Psi\rangle$ be the output of a quantum state created as a linear combination of unitaries. The success probability is bounded by*

$$|\langle\Psi|\psi_{K,B}\rangle|^2 \le \frac{\|\boldsymbol{\alpha}\|_2^2}{\|\boldsymbol{\alpha}\|_1^2} \left(1 - \max_{a \ne b} \|F_{a,b}\|_\infty \right) + \max_{a \ne b} \|F_{a,b}\|_\infty, \tag{52}$$

where $F_{a,b}$ is the block of the overlap matrix in Equation (15) corresponding to the bases (a,b).

The proof can be found in Appendix G.

We can leverage the previous result to obtain upper and lower bounds to create the state $|\psi\rangle$ with arbitrary precision, using the recipe from the previous paragraph.

Lemma 8 (Upper-bounds to prepare MBR states on a quantum computer). *Let $|\psi\rangle$ be the MBR of a state with sparsity K and number of basis B, with weights α, as in Equation (12). Let $d(K)$ be the required depth to prepare a K-sparse state in the computational basis, and let $d(U) = \max_b d(U_b)$ be the maximum depth required for the basis change. The depth D required to load $|\psi\rangle$ with success probability at least $1 - \delta$ is given by*

$$D \le d(W) + 2 \left(d(U) + d(K) \right) \sqrt{B} \log(2/\delta) \frac{\|\boldsymbol{\alpha}\|_1^2}{\|\boldsymbol{\alpha}\|_2^2}$$

$$\left(1 - \frac{\max_{a \ne b} \|F_{a,b}\|_\infty}{2} \left(\frac{\|\boldsymbol{\alpha}\|_1^2}{\|\boldsymbol{\alpha}\|_2^2} + 1 \right) \right). \tag{53}$$

The proof can be found in Appendix H.

The previous bounds purposely did not specify $d(K)$ and $d(U)$. First, we do not impose yet any constraint on the depth needed to apply the changes of basis. These
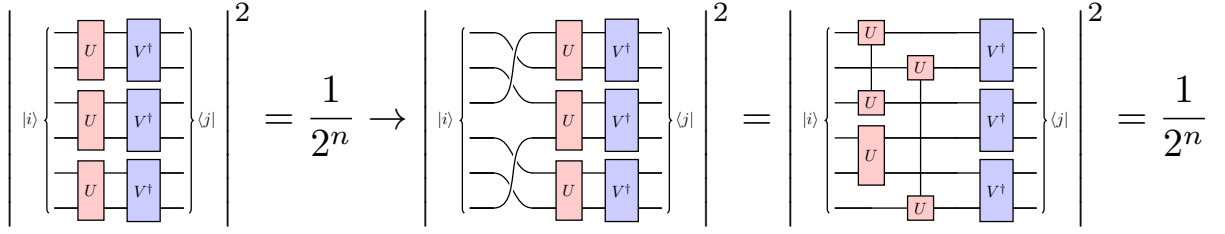
FIG. 5: Let U, V be two k-qubit MUBs, such that its tensor products $U^{\otimes n/k}, V^{\otimes n/k}$ are MUBs for n qubits. Since MUBs do not have ordering, it is possible to construct new sets of MUBs by applying arbitrary permutations to the bases $U^{\otimes n/k} \to U^{\otimes n/k}\Pi$. This allows for finding many different sets of MUBs.

will depend on the bases themselves, and choosing these bases is a challenge on its own. Second, the number $d(K)$ strongly depends on the number of ancillary qubits available [52–54]. These bounds become tight and imply an approximate lower-bound under the assumption that the only available tools are the preparation of sparse states in the CB and the basis unitaries U_b separately. This scenario acquires relevance in the context that all states for each basis are peaked [49]. Finally, we use ancillary qubits to prepare the control register $|b\rangle$. The use of ancillary qubits is necessary to create this superposition in the aforementioned scenario [55].

Additionally, we can adapt the results from Lemma 7 and Lemma 8 to obtain new upper bounds in MBR-state preparation in the case the bases form sets of MUBs.

Corollary 3. *Given the conditions of Lemma 7, for* $\mathrm{MBR}_{\mathrm{MUB}}$, *then*

$$P_{\text{success}} \le B \frac{\|\boldsymbol{\alpha}\|_\infty^2}{\|\boldsymbol{\alpha}\|_1^2} \left(1 - \frac{K\,B}{\sqrt{2^n}}\right) + \frac{K\,B}{\sqrt{2^n}}. \quad (54)$$

The proof is analogous to that of Lemma 7, usingalso Lemma 2.

Combinatorial considerations

Finally, we restrict the depth of the unitary operations. In order to maintain the MUB condition, we focus in the case in which the MUBs are local in subsystems of l qubits, with $l \in \mathcal{O}(1)$. This example illustrates that interesting MBRs are accessible with minimal resources.

Consider now the example of complete sets of MUBs in subsystems of size l, with $M = 2^l + 1$. We can compute the asymptotic scaling in Equation (36) as

$$C_M(K, B) \in \exp\left(\tilde{\mathcal{O}}\left(\text{poly}(n)\right)\right), \quad (55)$$

assuming $K, B \in \mathcal{O}\left(\text{poly}\,(n)\right)$. If we restrict the permutations to qubits, there are several choices to be done on how to select the (n/l) partitions of size l. In addition, one can choose several different sets of MUBs. This leads to large numbers of possible combinations

$$\sharp_{\text{MUB sets}}(n, l) = \left((2^l + 1)!\right)^{n/l} \quad (56)$$

$$\sharp_{\text{Partitions}}(n, l) = \frac{n!}{(l!)^{n/l}(n/l)!} \quad (57)$$

$$\sharp_{\text{Partitions}}(n, l)\sharp_{\text{MUB sets}}(n, l) \in e^{\tilde{\mathcal{O}}\left(n2^l\right)}. \quad (58)$$

We dive now into the numerous possibilities opened by the k-qubit MUBs. We first note that the MUB condition $|\langle i| U_2^\dagger U_1 |j\rangle| = 2^{-l/2}$ holds under permutations Π over the elements $|i\rangle, |j\rangle$. That is, if U_1, U_2 are MUB, then all bases $U_1\Pi_1, U_2\Pi_2$ are MUBs, for any pair of permutations Π_1, Π_2. The permutations can be understood as reorderings of the elements in the computational basis. Thus, if we choose two l-qubit MUBs to be simultaneously applied in a n-qubit system, then all permutations of basis elements allow for the MUB condition as well. See Figure 5 for a graphical description of qubit permutation. This argument increases the number of sets of MUBs, but not the size of each set. If we consider l-qubit subsystems in a n-qubit state, without permutations, the number of different sets grows as

$$\sharp_{\text{MUB sets}}(n, l) = \left((2^l + 1)!\right)^{n/l} \in \tilde{\Theta}\left(e^{n2^l}\right), \quad (59)$$

where the scaling is obtained following Stirling's approximation. The number of different size-l partitions of a set of size n grows as

$$\sharp_{\text{Partitions}}(n, l) = \frac{n!}{l!^{n/l}(n/l)!}$$

$$\in \tilde{\Theta}\left(\left(\frac{n}{l}\right)^{n\frac{l-1}{l}} e^{n/l}\right), \quad (60)$$

yielding a total number of combinations

$$\sharp_{\text{Partitions}}(n, l)\sharp_{\text{MUB sets}}(n, l) \in \tilde{\mathcal{O}}\left(\left(ne^{2^l}\right)^n\right). \quad (61)$$

Imposing the condition $l \in \mathcal{O}(\log n)$ in this reasoning, the number of different available sets of MUBs scales as

$$\sharp_{\text{Partitions}}(n, l)\sharp_{\text{MUB sets}}(n, l) \in \tilde{\mathcal{O}}\left(n^n\right). \quad (62)$$

Notice that it is also possible to use the l-qubit MUBs to construct bases that are no longer MUBs in the full

Input: - Several copies of $|\psi\rangle$

   - Set of bases $\{U_b\}$, error ε

Output: - Sets of coefficients $\left(\{\alpha_b\}, \{c_{i_b}^{(b)}\}\right)$, see Equation (12)

   - Sets of supports $\{S_b\}$, see Equation (12)

Steps:

   a. For each basis, apply the operation $U_b^\dagger$ to the quantum state $|\psi\rangle$.

   b. For each basis, Measure the quantum state in the computational basis to identify the most relevant elements $\{S_b\}$.

   c. Merge data from all bases and compute the fidelity matrix from Equation (15).

   d. Measure the amplitudes associated to $\{S_b\}$ for each basis.

   e. Revert the effect of non-orthogonality to obtain the coefficients $\left(\{\alpha_b\}, \{c_{i_b}^{(b)}\}\right)$.

TABLE I: Overview of the algorithm needed to obtain the MBR representation of a quantum state. A more detailed description of each step is available in Section IV C, including the costs associated to each of them.

n-qubit system. Such condition allows us to adapt the orthogonality relationships derived in Section II B to the use of only some l-qubit bases. MBR as a general framework is still usable, since the basis can still be constructed as Clifford circuits, and Equation (15) is classically accessible.

C. Finding multiple-basis representations of unknown states

In this section, we outline a protocol to find a MBR description of pure states. The input inputs are the quantum state $|\psi\rangle$ and the bases set $\{U_b\}$, and output are the coefficients and supports of the MBR. Note that finding a set of bases that effectively capture a given state of interest is in general a hard problem, which will be left as future research. However, once the sets of bases is given, obtaining the MBR is easy, since the supports can be obtained by sampling, and the coefficients by conducting experiments with small overheads on the corresponding supports.

We are interested in finding an approximation to a quantum state $|\psi\rangle$ in its MBR, as in Equation (12). The recipe to find the MBR of an unknown quantum state is summarized in Table I, and detailed in the following paragraphs.

a. Applying $U_b^\dagger$ The first step of the algorithm is to apply $U_b^\dagger$ to the input state. This unitary operations has the role of expressing the quantum state in a basis in which its CB_ε-rank is small, for a moderately large ε.

b. Sampling The obtained samples of a quantum state, in the limit of infinite samples, gives an approximation of the (absolute values of the) amplitudes in $|\psi\rangle$, in the sampling basis. In the context of MBR, sampling is not strictly necessary to obtain these amplitudes, this step is addressed later, but it is required to obtain the supports S_b that define the MBR.

Assume sampling from an unknonwn state $|\psi\rangle$, with sampling budget $M \in \mathcal{O}(\mathrm{poly}(n))$. As an extreme example, assume that $|\psi\rangle$ induces a uniform distribution. With high probability, M measurements output M different bitstrings, and therefore it is statistically impossible to distinguish uniform distributions over all bitstrings and only the observed ones. On the other hand, assume a peaked quantum circuit U satisfying [27, 49]

$$\max_{x\in\{0,1\}^n} |\langle x| U |0\rangle| \in \Omega(\mathrm{poly}^{-1}(n)). \qquad (63)$$

It is efficient to classically verify if U is peaked just by circuit. Intuitively speaking, this tomographical protocol requires a stronger peaked-condition, meaning at least K outcomes can be sampled with probability $\Omega(\mathrm{poly}^{-1}(n))$. We proceed now to describe the concentration properties needed for accurate sampling.

First, we define the rank of a probability distribution, in a similar way to the CB_ε-rank Definition 1.

Definition 16 (ε-rank of a probability distribution). *Let i be a random variable that can take values $i = \{0, 1, \ldots, N-1\}$, each with probability p_i, and $\sum_i p_i = 1$. The outcomes are ordered such that $p_i \geq p_{i+1}$ without loss of generality. Let*

$$S_K(p) = \sum_{i=1}^{K} p_i. \qquad (64)$$

The ε-rank of p is defined as

$$R_\varepsilon(p) = \min_K \{K : S_K \geq 1 - \varepsilon\} \qquad (65)$$

This definition allows us to formulate the following statement to give the conditions needed to identify the most relevant components of the quantum state by sampling.

Theorem 2 (Concentration properties). *Let i be a random variable that can take values $i = \{0, 1, \ldots, N-1\}$, each with probability p_i, and $\sum_i p_i = 1$. The outcomes are ordered such that $p_i \geq p_{i+1}$ without loss of generality. The probability distribution is sampled M times, with m_i samples corresponding to outcome i. For every ε, the outcomes m_i allow to estimate the ε-rank $R_\varepsilon(p)$ as K. The estimation K is the only possible value compatible*

with the observed outcomes with probability at least $1 - \delta$ if

$$m_{K+1} \geq \sqrt{M} \log(2/\delta). \qquad (66)$$

The true value S_K is estimated within precision

$$\Delta S_K \in \mathcal{O}\left(\frac{1}{\sqrt{M}}\right). \qquad (67)$$

The formal proof can be found in Appendix I, but a sketch is enough to comprehend the gist of this result. For a quantum state $|\psi\rangle$ we sample its outcomes in the computational basis according to some underlying probability specified by the state. We are interested in computing K, through the quantity $S_K(p)$, which can be estimated with accuracy $\mathcal{O}(M^{-1/2})$ via standard estimators and Monte Carlo arguments. We repeat this process for $S_{K+1}(p)$. Notice that only one sampling round is needed. Since we assume $p_i \geq p_{i+1}$, then $S_{K+1}(p) - S_K(p) \geq S_{K+2}(p) - S_{K+1}(p)$. Due to the sampling uncertainty, we can accurately compute K as long as $S_{K+1}(p) - S_K(p) \in \Omega(M^{-1/2})$, which forces the constraint in Equation (66) and constitutes the main result in Theorem 2.

In the light of the algorithm to obtain MBR, the supports $\{S_b\}$ for each change of basis are obtained by just sampling the state $U_b^\dagger |\psi\rangle$ a total amount of M times, and saving all outcomes signaling more than $\sqrt{M}$ times.

c. *Computing amplitudes* After identification of supports $\{S_b\}$ is complete, we can proceed to estimate the quantities $\alpha_b c_{i_b}^{(b)}$. However, the only quantities that are accessible through quantum measurements are

$$a_{i_b}^{(b)} = \langle i_b | U_b^\dagger | \psi \rangle. \qquad (68)$$

Both sets of quantities are related by

$$\begin{aligned}
a_{i_b}^{(b)} &= \alpha^{(b)} c_{i_b}^{(b)} + \sum_{\substack{a=1 \\ a \neq b}}^{B} \sum_{j \in S_a} \left(\alpha^{(a)} c_{j_a}^{(a)} \langle i_b | U_b^\dagger U_a | j_a \rangle \right) \\
&= \alpha^{(b)} c_{i_b}^{(b)} + \sum_{\substack{a=1 \\ a \neq b}}^{B} \sum_{j \in S_a} F_{i_b, j_a} \alpha^{(a)} c_{j_a}^{(a)}.
\end{aligned} \qquad (69)$$

For simplicity we group the elements $a_{i_b}^{(b)}$ in a vector $\boldsymbol{a}$. The comparison with Equation (12) is that $a_{i_b}^{(b)} = \alpha_b c_{i_b}^{(b)}$, and the value α_b is obtained by imposing the normalization constraint from Equation (14). In a more compact fashion, we observe

$$\boldsymbol{a} = F\boldsymbol{\alpha}. \qquad (70)$$

We first focus on obtaining the quantities given by Equation (68). Such estimation requires both the real and imaginary parts of $a_{i_b}^{(b)}$. For a fixed basis U_b, we perform the i-th Hadamard test with $M_H m_i / M$ measurements, where m_i is the times of appearances of the $|i\rangle$ in the sampling step. With this budget, we obtain the following error bounds.

Lemma 9 (Measurement errors). *Let $|\psi\rangle$ be a quantum state, and let $\boldsymbol{a}$ be coefficients defined in Equation (68). The estimation of one element is conducted via Hadamard tests of M_H samples, and both real and imaginary parts are obtained. For the error between the estimate $\hat{\boldsymbol{a}}$ and the true value $\boldsymbol{a}$ it holds that*

$$\|\boldsymbol{a} - \hat{\boldsymbol{a}}\|_2^2 \leq 2\sqrt{2} K B \sqrt{\frac{\log 2/\delta}{M_H}}. \qquad (71)$$

The proof can be found in Appendix J.

d. *Computing overlaps* The next step is to estimate the values F_{i_b, j_a} to obtain the coefficients in the MBR. Notice that only the relevant overlaps are needed, thus incurring in a computational cost of $\mathcal{O}((K\,B) \times C)$, where C is the cost associated to computing one overlap. Assuming we have access to a quantum computer capable of implementing the changes of basis given by $\{U_b\}$, computing the real and imaginary parts of F_{i_b, j_a} is efficient via Hadamard tests in the case $F_{i_a, j_b} \in \Omega(1/\mathcal{O}(\text{poly}\,(n)))$, since the measurement budget scales as $|F_{i_b, j_a}|^2$ for constant multiplicative error. In fact, analogous to Lemma 9, the error for the overlap matrix will scale as

$$\|F - \hat{F}\|_2^2 \leq 2\sqrt{2} K^2 B^2 \sqrt{\frac{\log 2/\delta}{M_H}} \qquad (72)$$

In certain cases, it is sometimes possible to compute the overlaps via classical algorithms, even when these overlaps decay exponentially with the number of qubits. For more information, we refer to Section III.

e. *Overlap inversion* The only calculation left is computing and inverting F to obtain $\boldsymbol{\alpha}$. Notice that we can easily extract the values of α_b by imposing the normalization constraint in Equation (14). The inversion of F is in principle only possible under the assumption that no vector $U_b |i_b\rangle$ is linearly dependent with the rest of the considered states. In the case this constraint is not fulfilled, it means that some measurements are redundant, and can therefore be neglected, and no fundamental caveat arises. We can easily circumvent this problem through the pseudo-inverse of F [56]. This step outputs the coefficients

$$\boldsymbol{\alpha} = F^{-1}\boldsymbol{a}, \qquad (73)$$

which translates into the sets of coefficients $\left(\{\alpha_b\}, \{c_{i_b}^{(b)}\}\right)$. Standard error propagation bounds the errors in $\boldsymbol{\alpha}$ to first order as

$$\begin{aligned}
\|\boldsymbol{\alpha} - \hat{\boldsymbol{\alpha}}\|_2 \leq \\
\|F^{-1}\|_2 \|\boldsymbol{a} - \hat{\boldsymbol{a}}\|_2 + \|F^{-1}\|_2^2 \|F - \hat{F}\|_2 \|\hat{\boldsymbol{a}}\|_2.
\end{aligned} \qquad (74)$$

These errors strongly depend on the condition of the overlap matrix F. Notice that F is a Gramian matrix, therefore it will become ill-conditioned if the considered bases are linearly dependent with each other. In particular, if $\mathcal{F}_{K, U_b^\dagger U_a} \sim 0$ for all pairs $(a, b), a \neq b$, the error

scale as the measurement error. Finally, the values $\boldsymbol{\alpha}$ are fully specified if the normalization condition from Equation (14) is imposed.

V. CONCLUSIONS

This work introduces the multiple-basis representation (MBR) of quantum states. The novelty of MBR consists in using several bases simultaneously, chosen in such a way that the state of interest has a significant overlap with some state with small CB_ε-rank, when expressed in the corresponding basis. The choice of bases determines whether the MBR can be used to compute overlaps using only classical resources, or quantum computers are required. We further provide a classification of subclasses of MBR depending on the properties of these unitaries and distill conditions under which quantum computers are needed to perform classical computations of expectation values. We also discuss the optimality of the choice of bases to apply MBR, and prove that mutually unbiased bases (MUB) minimize the relative overlaps among the components associated to each basis. We also characterize several properties of MBR under the MUB assumption.

This manuscript also gives several possible applications for the MBR. The first application is using MBR to find approximations of ground states of physical systems. The results here depicted show comparable performance to tensor-network methods. We foresee that our method can capture long-range correlations more accurately than standard algorithms. The second application is the merge of MBR with the reduce&chop method to simulate deep circuits through classical combination of shallow circuits. We provide rigorous upper and lower bounds to the depth savings obtained by merging these two techniques together. Finally, we outline a tomographical protocol based on the MBR representation for pure quantum states, when many copies of this state can be processed through a quantum computer, providing rigurous bounds on the attained error.

We envision this work to open the path to systematically use several bases simultaneously to represent quantum states, and exploit their advantages. In particular, we envision these bases to be implementable with limited-depth quantum circuits, thus being executable in near-term quantum computers. Even with constant-depth quantum circuits, the combinatorial magnitudes arising when considering all possible MBRs, the question becomes how to find useful bases, rather than searching for optimal ones. Hence, any approach to this problem is inherently heuristic. As shown in the ground state approximation problem addressed in this work, physically motivated choices lead to classically tractable ways to select bases, and opens the path towards efficiently solving problems using heuristic methods.

The ultimate goal of MBR is to help understanding, and ideally pushing, the boundaries of what can be efficiently done with a classical computer, or with quantum-assisted classical approaches. To this end, this work provides conditions under which computing expectation values of MBR-states is classically simulable. The MBR efficiently represents quantum states with certain properties, along the lines of tensor networks for states with limited entanglement or stabilizer states for circuits with Clifford gates. Since the MBR can be utilized with and without assistance of a quantum computer, it allows for studying hybrid representations of quantum states. This is left as an open research to be addressed with access to experimental resources.

DATA AVAILABILITY

The code to simulate the numerical data in this manuscript is available online [57], as well as the data shown in plots [58].

ACKNOWLEDGMENTS

[1] H. Ma, M. Govoni, and G. Galli, Quantum simulations of materials on near-term quantum computers, npj Computational Materials **6**, 1 (2020).

[2] J. R. McClean, K. J. Sung, I. D. Kivlichan, Y. Cao, C. Dai, E. S. Fried, C. Gidney, B. Gimby, P. Gokhale, T. Häner, T. Hardikar, V. Havlíček, O. Higgott, C. Huang, J. Izaac, Z. Jiang, X. Liu, S. McArdle, M. Neeley, T. O'Brien, B. O'Gorman, I. Ozfidan, M. D. Radin, J. Romero, N. Rubin, N. P. D. Sawaya, K. Setia, S. Sim, D. S. Steiger, M. Steudtner, Q. Sun, W. Sun, D. Wang, F. Zhang, and R. Babbush, OpenFermion: The Electronic Structure Package for Quantum Computers (2019), arXiv:1710.07629 [physics, physics:quant-ph].

[3] S. N. Genin, I. G. Ryabinkin, and A. F. Izmaylov, Quantum chemistry on quantum annealers, arXiv:1901.04715 [physics, physics:quant-ph] (2019), arXiv:1901.04715 [physics, physics:quant-ph].

[4] J. I. Cirac, D. Pérez-García, N. Schuch, and F. Verstraete, Matrix product states and projected entangled pair states: Concepts, symmetries, theorems, Reviews of Modern Physics **93**, 045003 (2021).

[5] U. Schollwöck, The density-matrix renormalization group in the age of matrix product states, Annals of Physics January 2011 Special Issue, **326**, 96 (2011).

[6] D. Gottesman, The Heisenberg Representation of Quantum Computers (1998), arXiv:quant-ph/9807006.

[7] S. Aaronson and D. Gottesman, Improved simulation of stabilizer circuits, Physical Review A **70**, 052328 (2004).

[8] S. Bravyi, D. Browne, P. Calpin, E. Campbell, D. Gosset, and M. Howard, Simulation of quantum circuits by low-rank stabilizer decompositions, Quantum **3**, 181 (2019).

[9] L. Vinkhuijzen, T. Coopmans, D. Elkouss, V. Dunjko, and A. Laarman, LIMDD: A Decision Diagram for Simulation of Quantum Computing Including Stabilizer States, Quantum **7**, 1108 (2023).

[10] M. L. Goh, M. Larocca, L. Cincio, M. Cerezo, and F. Sauvage, Lie-algebraic classical simulations for variational quantum computing (2023), arXiv:2308.01432 [quant-ph].

[11] M. Cerezo, M. Larocca, D. García-Martín, N. L. Diaz, P. Braccia, E. Fontana, M. S. Rudolph, P. Bermejo, A. Ijaz, S. Thanasilp, E. R. Anschuetz, and Z. Holmes, Does provable absence of barren plateaus imply classical simulability? Or, why we need to rethink variational quantum computing (2023), arXiv:2312.09121 [quant-ph, stat].

[12] M. S. Rudolph, E. Fontana, Z. Holmes, and L. Cincio, Classical surrogate simulation of quantum systems with LOWESA (2023), arXiv:2308.09109 [cond-mat, physics:quant-ph].

[13] R. P. Feynman, Simulating physics with computers, International Journal of Theoretical Physics **21**, 467 (1982).

[14] P. W. Shor, Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer, SIAM Journal on Computing **26**, 1484 (1997).

[15] Y. Liu, S. Arunachalam, and K. Temme, A rigorous and robust quantum speed-up in supervised machine learning, Nature Physics **17**, 1013 (2021).

[16] R. Molteni, C. Gyurik, and V. Dunjko, Exponential quantum advantages in learning quantum observables from classical data (2024), arXiv:2405.02027 [quant-ph].

[17] A. W. Harrow, A. Hassidim, and S. Lloyd, Quantum algorithm for solving linear systems of equations, Physical Review Letters **103**, 150502 (2009), arXiv:0811.3171 [quant-ph].

[18] H.-Y. Huang, M. Broughton, J. Cotler, S. Chen, J. Li, M. Mohseni, H. Neven, R. Babbush, R. Kueng, J. Preskill, and J. R. McClean, Quantum advantage in learning from experiments, arXiv:2112.00778 [quant-ph] (2021), arXiv:2112.00778 [quant-ph].

[19] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, R. Biswas, S. Boixo, F. G. S. L. Brandao, D. A. Buell, B. Burkett, Y. Chen, Z. Chen, B. Chiaro, R. Collins, W. Courtney, A. Dunsworth, E. Farhi, B. Foxen, A. Fowler, C. Gidney, M. Giustina, R. Graff, K. Guerin, S. Habegger, M. P. Harrigan, M. J. Hartmann, A. Ho, M. Hoffmann, T. Huang, T. S. Humble, S. V. Isakov, E. Jeffrey, Z. Jiang, D. Kafri, K. Kechedzhi, J. Kelly, P. V. Klimov, S. Knysh, A. Korotkov, F. Kostritsa, D. Landhuis, M. Lindmark, E. Lucero, D. Lyakh, S. Mandrà, J. R. McClean, M. McEwen, A. Megrant, X. Mi, K. Michielsen, M. Mohseni, J. Mutus, O. Naaman, M. Neeley, C. Neill, M. Y. Niu, E. Ostby, A. Petukhov, J. C. Platt, C. Quintana, E. G. Rieffel, P. Roushan, N. C. Rubin, D. Sank, K. J. Satzinger, V. Smelyanskiy, K. J. Sung, M. D. Trevithick, A. Vainsencher, B. Villalonga, T. White, Z. J. Yao, P. Yeh, A. Zalcman, H. Neven, and J. M. Martinis, Quantum supremacy using a programmable superconducting processor, Nature **574**, 505 (2019).

[20] L. S. Madsen, F. Laudenbach, M. F. Askarani, F. Rortais, T. Vincent, J. F. F. Bulmer, F. M. Miatto, L. Neuhaus, L. G. Helt, M. J. Collins, A. E. Lita, T. Gerrits, S. W. Nam, V. D. Vaidya, M. Menotti, I. Dhand, Z. Vernon, N. Quesada, and J. Lavoie, Quantum computational advantage with a programmable photonic processor, Nature **606**, 75 (2022).

[21] H.-S. Zhong, H. Wang, Y.-H. Deng, M.-C. Chen, L.-C. Peng, Y.-H. Luo, J. Qin, D. Wu, X. Ding, Y. Hu, P. Hu, X.-Y. Yang, W.-J. Zhang, H. Li, Y. Li, X. Jiang, L. Gan, G. Yang, L. You, Z. Wang, L. Li, N.-L. Liu, C.-Y. Lu, and J.-W. Pan, Quantum computational advantage using photons, Science **370**, 1460 (2020).

[22] C. Piveteau and D. Sutter, Circuit knitting with classical communication, IEEE Transactions on Information Theory , 1 (2023), arXiv:2205.00016 [quant-ph].

[23] P. Huembeli, G. Carleo, and A. Mezzacapo, Entanglement Forging with generative neural network models (2022), arXiv:2205.00933 [quant-ph].

[24] T. Peng, A. W. Harrow, M. Ozols, and X. Wu, Simulating Large Quantum Circuits on a Small Quantum Computer, Physical Review Letters **125**, 150504 (2020).

[25] S. C. Marshall, C. Gyurik, and V. Dunjko, High Dimensional Quantum Machine Learning With Small Quantum Computers (2022), arXiv:2203.13739 [quant-ph].

[26] A. Pérez-Salinas, R. Draškić, J. Tura, and V. Dunjko, Shallow quantum circuits for deeper problems, Physical Review A **108**, 062423 (2023).

[27] S. Aaronson and Y. Zhang, On verifiable quantum advantage with peaked circuit sampling (2024), arXiv:2404.14493 [quant-ph].

[28] M. S. Rudolph, J. Chen, J. Miller, A. Acharya, and A. Perdomo-Ortiz, Decomposition of Matrix Product States into Shallow Quantum Circuits (2022), arXiv:2209.00595 [quant-ph].

[29] S. Masot-Llima and A. Garcia-Saez, Stabilizer Tensor Networks: Universal quantum simulator on a basis of stabilizer states (2024), arXiv:2403.08724 [quant-ph].

[30] T. Durt, B.-G. Englert, I. Bengtsson, and K. Życzkowski, On mutually unbiased bases, International Journal of Quantum Information **08**, 535 (2010), arXiv:1004.3348 [quant-ph].

[31] A. Klappenecker and M. Roetteler, Constructions of Mutually Unbiased Bases (2003), arXiv:quant-ph/0309120.

[32] M. Grassl, On SIC-POVMs and MUBs in Dimension 6 (2009), arXiv:quant-ph/0406175.

[33] S. Bandyopadhyay, P. O. Boykin, V. Roychowdhury, and F. Vatan, A new proof for the existence of mutually unbiased bases (2001), arXiv:quant-ph/0103162.

[34] O. Kern, K. S. Ranade, and U. Seyfarth, Complete sets of cyclic mutually unbiased bases in even prime power dimensions, Journal of Physics A: Mathematical and Theoretical **43**, 275305 (2010), arXiv:0912.4661 [quant-ph].

[35] U. Seyfarth and K. S. Ranade, Construction of mutually unbiased bases with cyclic symmetry for qubit systems, Physical Review A **84**, 042327 (2011), arXiv:1106.3515 [quant-ph].

[36] U. Seyfarth, L. L. Sánchez-Soto, and G. Leuchs, Practical implementation of mutually unbiased bases using quantum circuits, Physical Review A **91**, 032102 (2015).

[37] T. J. Yoder, G. H. Low, and I. L. Chuang, Fixed-point quantum search with an optimal number of queries, Physical Review Letters **113**, 210501 (2014), arXiv:1409.3305 [quant-ph].

[38] D. Malz, G. Styliaris, Z.-Y. Wei, and J. I. Cirac, Preparation of matrix product states with log-depth quantum circuits (2023), arXiv:2307.01696 [cond-mat, physics:quant-ph].

[39] A. Pérez-Salinas, H. Wang, and X. Bonet-Monroig, Analyzing variational quantum landscapes with information content, npj Quantum Information **10**, 27 (2024), arXiv:2303.16893 [quant-ph].

[40] B. M. Terhal, Quantum error correction for quantum memories, Reviews of Modern Physics **87**, 307 (2015).

[41] J. R. McClean, M. E. Kimchi-Schwartz, J. Carter, and W. A. de Jong, Hybrid quantum-classical hierarchy for mitigation of decoherence and determination of excited states, Physical Review A **95**, 042308 (2017).

[42] N. Schuch and F. Verstraete, Computational complexity of interacting electrons and fundamental limitations of density functional theory, Nature Physics **5**, 732 (2009).

[43] M. Mitchell, *An Introduction to Genetic Algorithms* (The MIT Press, 1996).

[44] O. Goldschmidt and D. S. Hochbaum, A Polynomial Algorithm for the k-cut Problem for Fixed k, Mathematics of Operations Research **19**, 24 (1994).

[45] M. X. Goemans and D. P. Williamson, Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming, Journal of the ACM **42**, 1115 (1995).

[46] I. Dunning, S. Gupta, and J. Silberholz, What Works Best When? A Systematic Evaluation of Heuristics for Max-Cut and QUBO, INFORMS Journal on Computing **30**, 608 (2018).

[47] H. W. J. Blöte and Y. Deng, Cluster Monte Carlo simulation of the transverse Ising model, Physical Review E **66**, 066110 (2002).

[48] M. S. L. du Croo de Jongh and J. M. J. van Leeuwen, Critical behavior of the two-dimensional Ising model in a transverse field: A density-matrix renormalization calculation, Physical Review B **57**, 8494 (1998).

[49] S. Bravyi, D. Gosset, and Y. Liu, Classical simulation of peaked shallow quantum circuits (2023), arXiv:2309.08405 [quant-ph].

[50] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition* (Cambridge University Press, Cambridge, 2010).

[51] R. de Wolf, Quantum Computing: Lecture Notes (2023), arXiv:1907.09415 [quant-ph].

[52] X.-M. Zhang, M.-H. Yung, and X. Yuan, Low-depth Quantum State Preparation (2021), arXiv:2102.07533 [quant-ph].

[53] X. Sun, G. Tian, S. Yang, P. Yuan, and S. Zhang, Asymptotically Optimal Circuit Depth for Quantum State Preparation and General Unitary Synthesis (2023), arXiv:2108.06150 [quant-ph].

[54] D. Ramacciotti, A.-I. Lefterovici, and A. F. Rotundo, A simple quantum algorithm to efficiently prepare sparse states (2023), arXiv:2310.19309.

[55] M. Oszmaniec, A. Grudka, M. Horodecki, and A. Wójcik, Creation of superposition of unknown quantum states, Physical Review Letters **116**, 110403 (2016), arXiv:1505.04955 [math-ph, physics:quant-ph].

[56] R. Penrose, A generalized inverse for matrices, Mathematical Proceedings of the Cambridge Philosophical Society **51**, 406 (1955).

[57] P. Emonts and A. Pérez-Salinas, Github repository: `github/patrickemonts/multiple-basis-representation` (2024).

[58] P. Emonts and A. Pérez-Salinas, Github repository: `github/patrickemonts/mbr-data` (2024).

[59] Y. Zhou, E. M. Stoudenmire, and X. Waintal, What Limits the Simulation of Quantum Computers?, Physical Review X **10**, 041038 (2020).

[60] E. W. Cheney and D. R. Kincaid, *Linear Algebra: Theory and Applications* (Jones & Bartlett Learning, 2009).

[61] R. W. Bailey, Distributional Identities of Beta and Chi-Squared Variates: A Geometrical Interpretation, The American Statistician **46**, 117 (1992), 2684178.

[62] E. B. Wilson, Probable Inference, the Law of Succession, and Statistical Inference, Journal of the American Statistical Association **22**, 209 (1927).

[63] P. S. de Laplace, *Théorie Analytique Des Probabilités*, Vol. 7 (Courcier, 1820).

1. Proof of Lemma 1

Proof. The expectaction of a MBR state with respect to an observable is simply

$$\langle \psi_{K,B} | \hat{O} | \psi_{K,B} \rangle = \sum_{b=1}^{B} \sum_{a=1}^{B} \sum_{\substack{i_b \in S_b \\ |S_b|=K}} \sum_{\substack{j_a \in S_a \\ |S_a|=K}} \alpha_a \alpha_b \left(c_{i_b}^{(b)} \right)^* c_{j_a}^{(a)} \langle i_b | U_b^\dagger \hat{O} U_a | j_a \rangle. \tag{75}$$

The cost of this calculation scales as $\mathcal{O}\left(\text{poly}\left(K,B\right)\right) \times C = \mathcal{O}\left(\text{poly}\left(n\right)\right) \times C$, where C is the cost of computing one element, that is

$$\hat{O}_{i_b, j_a} = \langle i_b | U_b^\dagger \hat{O} U_a | j_a \rangle. \tag{76}$$

If this operation can be handled with classical resources, that is $C \in \mathcal{O}\left(\text{poly}\left(n\right)\right)$, then computing expectation values is doable with classical resources. $\qquad\square$

Appendix A: Proof of Theorem 1

We provide an extended version of the theorem here.

Theorem 3. *The relationship among the MBR subclasses and other representations are the following:*

$$\text{STAB} \subsetneq \text{MBR}_C \tag{A1}$$
$$\text{MPS} \subsetneq \text{MBR}_C \tag{A2}$$
$$\text{STAB} \not\subset \text{MBR}_{C-\log} \tag{A3}$$
$$\text{STAB} \cap \text{MBR}_{C-\log} \neq \emptyset \tag{A4}$$
$$\text{MPS} \cap \text{MBR}_{C-\log} \neq \emptyset \tag{A5}$$
$$\text{MPS} \not\subset \text{STAB} \tag{A6}$$
$$\text{MPS} \cap \text{STAB} \neq \emptyset \tag{A7}$$
$$\text{MBR}_{\text{MUB}} \cap \text{STAB} \neq \emptyset \tag{A8}$$
$$\text{MBR}_{\text{MUB}} \cap \text{MPS} \neq \emptyset \tag{A9}$$
$$\text{MBR}_{\text{MUB}} \cap \left(\text{MBR}_C \cap \text{STAB}\right) \neq \emptyset \tag{A10}$$
$$\text{MBR}_{\text{MUB}} \cap \left(\text{MBR}_{\not\subset} \cap \text{MBR}_C\right) = \emptyset \tag{A11}$$

Proof. We proof each statement separately.

STAB $\subsetneq$ MBR$_C$: The STAB formalism deals with Clifford circuits, possibly doped with at most $\mathcal{O}(\log n)$ non-Clifford gates. Computing $\langle i | U^\dagger \hat{O} V | j \rangle$ is possible (assuming U, V satisfy the Clifford condition) since $\hat{O}$ is a combination of $\mathcal{O}\left(\text{poly}\left(n\right)\right)$-many Pauli strings. To do so, we just need to realize that $U^\dagger \hat{O} V$ is a combination of $\mathcal{O}\left(\text{poly}\left(n\right)\right)$-many Clifford circuits, and propagate the stabilizers of the $|0\rangle$ state [6], thus showing STAB $\subset$ MBR$_C$.

To show STAB $\neq$ MBR$_C$ we just need to find an example of circuits for which computing $\langle i | U^\dagger \hat{O} V | j \rangle$ is doable with classical methods outside the stabilizer formalism. Choose U, V to be a single-layer of arbitrary unitary rotations, thus $\mathcal{O}(n)$ non-Clifford gates and outside the reach of the stabilizer formalism. However, for $\hat{O}$ Pauli strings, computing the quantity of interest is possible. It is only required to perform $\mathcal{O}(n)$ single-qubit computations, with linear cost.

a. MPS $\subsetneq$ MBR$_C$ The MPS formalism deals with circuits with low entanglement. Computing $\langle i | U^\dagger \hat{O} V | j \rangle$ is possible (assuming U, V satisfy the MPS condition) since $\hat{O}$ is a combination of $\mathcal{O}\left(\text{poly}\left(n\right)\right)$-many Pauli strings. To do so, we just need to realize that $U^\dagger \hat{O} V$ is matrix product operator with bond dimension in $\mathcal{O}\left(\text{poly}\left(n\right)\right)$, thus showing MPS $\subset$ MBR$_C$.

To show MPS $\neq$ MBR$_C$ we just need to find an example of circuits for which computing $\langle i | U^\dagger \hat{O} V | j \rangle$ is doable with classical methods outside the MPS formalism. Assume $\hat{O}$ to be a Pauli string. Choose U, V to be a Clifford circuit with long-term connectivity. The bond dimension to compute observables using the MPS formalism increases exponentially with the range. However, the stabilizer formalism trivially allows to compute observables for these circuits. .

STAB $\not\subset$ MBR$_{C-\log}$: Following the reasoning of the previous paragraph, the STAB formalism allows us to compute $\langle i | U^\dagger \hat{O} V | j \rangle$ for arbitrary Clifford circuits. This includes circuits with $\mathcal{O}\left(\text{poly}\left(n\right)\right)$-depth, not only $\mathcal{O}(\log n)$. Existing results show that it is possible to efficiently reduce any Clifford circuit to another one with at most $\mathcal{O}\left(\text{poly}\left(n\right)\right)$-depth, thus no possible reduction can be applied to circumvent this result.

STAB $\cap\, \mathrm{MBR}_{C-\log} \neq \emptyset$: To show this it is enough to find an example of a STAB circuit with logarithmic depth. It is enough with restricting the depth of the Clifford circuits.

MPS $\cap\, \mathrm{MBR}_{C-\log} \neq \emptyset$: MPS representation can compute $\langle i|\, U^\dagger \hat{O} V\, |j\rangle$ for certain circuits with $\mathcal{O}(\log n)$-depth. For circuits with nearest-neighbours connectivity, each two-qubit gate can be decomposed as a superposition $\mathcal{O}(1)$ tensor-product operations. Recurrent application of this principle allows to find an MPS representation with $\chi \in \mathcal{O}\left(\mathrm{poly}\left(n\right)\right)$, thus efficient. This proof has been widely explored in the literature [59].

It is possible to find examples of $\mathrm{MBR}_{C-\log}$ that are not tractable through MPS. As a preliminary result, a state with CB_ε rank K is MPS izable with bond dimension at most K. The reason is that if we perform the Schmidt decomposition between any two partitions (A, B), such that $2^{\min(A,B)} > K$, we separate the digits in the binary representation of the integers. This provides a Schmidt rank at most K, and holds for all partitions. Now, if $\{U_b\}$ admits a representation as matrix product operator with bond dimension χ, then $|\psi_b\rangle$ admits an MPS representation of bond dimension $\chi(|\psi_b\rangle)$, since $\chi(U\,|\psi\rangle) \leq \chi(U)\chi(|\psi\rangle)$. In addition, the property $\chi(|\psi\rangle + |\phi\rangle) \leq \chi(|\phi\rangle) + \chi(|\phi\rangle)$, and therefore the maximal bond dimension is upper bounded, $\chi \in \mathcal{O}\left(\mathrm{poly}\left(n\right)\right)$.

Additionally, one can find circuits with depth $\omega(\log n)$ that are classically tractable with MPS. It suffices with padding a MPS-representable states with gates that merge into identity. Those fall into the category of MBR_C, but not $\mathrm{MBR}_{C-\log}$. Finally MPS $\in \mathrm{MBR}_C$ by definition.

MPS $\not\subset$ STAB: A Clifford circuit with long-range correlations, as $\mathcal{O}(n)$, even with $\mathcal{O}(1)$-depth, demands an exponential overhead in the bond dimension χ of the MPS representation, thus making it inefficient.

MPS $\cap$ STAB $\neq \emptyset$: In alignment with the previous paragraph, any $\mathcal{O}(\log n)$-circuit with nearest neighbour connectivity admits MPS representation, in particular a circuit composed only by Clifford gates.

$\mathrm{MBR}_{\mathrm{MUB}} \cap$ STAB $\neq \emptyset$: It is possible to construct a complete set of MUB using only Clifford circuits [34].

$\mathrm{MBR}_{\mathrm{MUB}} \cap$ MPS $\neq \emptyset$: Construct a pair of bases that are locally-MUBs, with locality $\mathcal{O}(1)$. These are by construction MUBs, and quantum circuits of depth $\mathcal{O}(1)$, which are representable by MPS.

$\mathrm{MBR}_{\mathrm{MUB}} \cap (\mathrm{MBR}_C \cap$ STAB $) \neq \emptyset$: It is not needed to have Clifford circuits to obtain MUBs. One can consider, for instance, two unitaries given by single-qubit arbitrary rotations with the constraint that they form —MUBs. Since they are single-qubit unitaries, computing $\langle i|\, R^\dagger U^\dagger \hat{O} V R\, |j\rangle$ becomes trivial, with $\mathcal{O}(n)$ non-Clifford gates.

$\mathrm{MBR}_{\mathrm{MUB}} \cap (\mathrm{MBR}_{\not{C}} \cap \mathrm{MBR}_C) = \emptyset$: Assume that there are sets of MUBs that are not computable classically. Then, one needs to resort to quantum computers to compute $|i\rangle\, U^\dagger \hat{O} V\, |j\rangle$, in particular for $\hat{O} = I$. In that case, this quantity is by definition (in amplitude) $N^{-1} = 2^{-n}$. Quantum methods would require an exponential measurement budget to resolve the relative phases.

$\square$

Appendix B: Proof of Lemma 2

The quantity of interest can be computed as

$$
\mathcal{F}_K = \min_{U \in \mathcal{SU}(2^n)} \max_{\substack{|\psi_1\rangle, |\psi_2\rangle \\ CB_\varepsilon\left(\left|\psi_{\{1,2\}}\right\rangle\right)=K}} |\langle\psi_1|\, U\, |\psi_2\rangle|. \tag{B1}
$$

More explicitly, we can rewrite it as

$$
\mathcal{F}_K = \min_{U \in \mathcal{SU}(2^n)} \max_{\substack{\{c_i^{(b)}, b \in \{1,2\}\} \\ S_1, S_2; |S_b|=K}} \left|d_j^* c_i \langle j|\, U\, |i\rangle\right|. \tag{B2}
$$

Since we can optimize over S_b, given a fixed U and set $\{c_i, d_i\}$ it is always possible to choose the sparse sets such that the largest coefficients correspond to the largest overlaps. Thus, the operation U must fulfill a condition that $\mathcal{F}_{U,K}$ is independent of the sets S_b. Such requirement is met when U defines a change of bases between MUBs. Finally, for any pair of sets S_1, S_2, it is always possible to choose the coefficients $\{c_i, d_i\}$ such that they perfectly align with the overlap between bases given by $\langle j|\, U\, |i\rangle$. Notice that the phases can be complex. Thus, the maximization is solved by choosing equally distributed coefficients $\{c_i, d_i\}$ aligned with the maximum values of $\langle j|\, U\, |i\rangle$, and the minimization problem is solved by making all overlaps equal

$$
|\langle i|\, U\, |j\rangle| = |\langle i'|\, U\, |j'\rangle|. \tag{B3}
$$

Normalization conditions imposed by unitarity of U suffice to obtain

$$
|\langle i|\, U\, |j\rangle| = \frac{1}{\sqrt{2^n}} \tag{B4}
$$

The relative overlap, with respect to the change of basis U, is given by

$$|\langle\psi_2|U|\psi_1\rangle| = \left|\sum_{j\in S_2}\sum_{i\in S_1} d_j^* c_i \langle j|U|i\rangle\right|. \tag{B5}$$

We recall $|S_b| = K$. Using the triangular inequality we can bound

$$|\langle\psi_2|U|\psi_1\rangle| \le \sum_{j\in S_2}\sum_{i\in S_1} |d_j||c_i||\langle j|U|i\rangle|. \tag{B6}$$

Each piece can be bounded separatedly. First, by assumption U defines a change between MUB, and thus the overlap between all elements is given by $2^{-n/2}$. Secondly, the coefficients $\{c_i\}$ are subject to the condition of $\sum_i |c_i|^2 = 1$, and equivalently for d. Therefore, the maximum possible overlap is attained when the quantum state is equally distributed among all the considered elements. Thus,

$$|\langle\psi_2|U|\psi_1\rangle| \le \sum_{j\in S_2}\sum_{i\in S_1} \frac{1}{K\sqrt{2^n}} = \frac{K}{\sqrt{2^n}}, \tag{B7}$$

leading to the desired result. $\qquad\square$

Appendix C: Proof of Lemma 3

We start by writing the MBR representation. The index K is ommited for convenience.

$$|\psi_B\rangle = \sum_b \alpha_b |\psi_b\rangle. \tag{C1}$$

The overlap of $|\psi_B\rangle$ with the target state $|\psi\rangle$ is given by

$$|\langle\psi_B|\psi\rangle| = \left|\sum_b \alpha_b \langle\psi|\psi_b\rangle\right|. \tag{C2}$$

This overlap is maximized by choosing α_b in such a way that it aligns with the partial overlaps $\langle\psi|\psi_b\rangle$. Since we are allowed to choose α in any way, under this condition

$$|\langle\psi_B|\psi\rangle| = \frac{1}{N}\sum_b |\langle\psi|\psi_b\rangle|^2, \tag{C3}$$

with N being a normalization constant. We can fix its value by imposing the condition

$$1 = |\langle\psi_B|\psi_B\rangle| = \frac{1}{N^2}\sum_{a,b} \langle\psi_b|\psi\rangle \langle\psi|\psi_a\rangle \langle\psi_a|\psi_b\rangle. \tag{C4}$$

Separating terms with $a = b$, we obtain

$$|\langle\psi_B|\psi_B\rangle| = \frac{1}{N^2}\left(\sum_b |\langle\psi_b|\psi\rangle|^2 + \sum_{a\neq b} \langle\psi_b|\psi\rangle \langle\psi|\psi_a\rangle \langle\psi_a|\psi_b\rangle\right). \tag{C5}$$

Using Lemma 2, we can bound the second term as

$$\left|\sum_{a\neq b} \langle\psi_b|\psi\rangle \langle\psi|\psi_a\rangle \langle\psi_a|\psi_b\rangle\right| \le \frac{K}{\sqrt{2^n}}\left|\sum_{a\neq b} \langle\psi_b|\psi\rangle \langle\psi|\psi_a\rangle\right|. \tag{C6}$$

The elements inside the sum can be rewritten as

$$\left|\sum_{a\neq b} \langle\psi_b|\psi\rangle \langle\psi|\psi_a\rangle\right| = \left|\sum_{a,b} \langle\psi_b|\psi\rangle \langle\psi|\psi_a\rangle - \sum_b \langle\psi_b|\psi\rangle \langle\psi|\psi_b\rangle\right|, \tag{C7}$$

and each of these terms is respectively the squared 1-norm and 2-norm of the vector specified by the overlaps. Since the $\| \cdot \|_1 \leq \sqrt{B} \| \cdot \|_2$, we can further simplify the previous equations to

$$\left| \sum_{a \neq b} \langle \psi_b | \psi \rangle \langle \psi | \psi_a \rangle \langle \psi_a | \psi_b \rangle \right| \leq \frac{KB}{\sqrt{2^n}} \sum_b |\langle \psi_b | \psi \rangle|^2 . \tag{C8}$$

Plugging this result into Equation (C4) we obtain

$$N^2 = \sqrt{\sum_b |\langle \psi_b | \psi \rangle|^2} \left(1 + z \frac{KB}{\sqrt{2^n}} \right), \tag{C9}$$

for some $z \in [-1, 1]$. Using this result in Equation (C2) and using the Taylor approximation in the $KB\sqrt{2^{-n}}$ term, we obtain the desired result. $\qquad\square$

Appendix D: Proof of Lemma 4

Consider B different K-sparse states $|\psi_b\rangle$, and B changes of bases $\{U_b\}_{b \leq B}$ each defining a basis used to to express the corresponding state. We are interested in finding the dimension of the space opened by $U_b |\psi_b\rangle$.

We can argue by Gram-Schmidt orthogonalization process [60]. The input of the usual formulation of the Gram-Schmidt process is a set of non-orthogonal vectors, and the output is another set of orthonormal vectors spanning the same linear space. The process is deterministic, provided the input vectors are ordered. The key step of Gram-Schmidt is subtracting from one vector its own projection over the previously considered one, so that each vector is only kept if it opens a new dimension.

In our case, the input set is substituted by the spaces opened by sparse states with different bases. Let $\mathcal{U}_b$ be the linear spaced spanned by $\{U_b |i\rangle\}_{i \in S_b}$, where $|S_b| = K$. The corresponding projector is $\Pi_b = U \left(\sum_{i \in S_b} |i\rangle \langle i| \right) U^\dagger$. We apply Gram-Schmidt orthogonalization to the set $\{\mathcal{U}_b\}_b$ to obtain a set of orthogonal spaces $\{\mathcal{V}_b\}_b$. We do not consider normalization in this step. Each new orthogonal space is given by

$$\mathcal{V}_b = \mathcal{U}_b - \sum_{a=1}^{b-1} \mathrm{Tr}\left(\mathcal{U}_b \Pi_a \right) \mathcal{U}_a. \tag{D1}$$

By recalling Lemma 2, we know that the overlap between any two K-sparse states each one in a MUB is bounded by $K2^{-n/2}$. Therefore, it is guaranteed that $\mathcal{V}_b$ has a non-zero dimension as long as

$$\frac{K(B-1)}{\sqrt{2^n}} \leq 1, \tag{D2}$$

leading to the desired result. $\qquad\square$

Appendix E: Proof of Lemma 5

We are interested in computing the fraction of states that is ε-close to a K sparse state. To do so, we need to integrate the surface of the sphere for quantum states within the desired distance. To do so we just need to recall that a the amplitudes of a n-qubit state can be understood as points on a sphere of dimension 2^{n+1}. Correspondingly, all states with CB-rank K can be understood as the unit sphere of dimensions $2K$. The fraction of states within ε distance to any state of CB-rank K is the surface of the hypersphere around the $(2K-1)$-dimensional equator. The assumed Haar randomness of Haar allows us to treat the corresponding subspaces with the beta distribution [61]. Let $(x_1, x_2, \ldots, x_N)$ be the coordinates of points on a unit sphere, thus $\sum_i x_i^2 = 1$. We sum a subset of these variables as $X_K = \sum_{i=1}^{2K} x_i^2$. Then X_K follows a random distribution

$$X_K \sim \mathrm{Beta}\left(K, N - K \right). \tag{E1}$$

The corresponding surface is obtained by computing the probabilities of $X_K \geq 1 - \varepsilon$, thus

$$S_K(\varepsilon) = \int_0^\varepsilon dt \; t^{2^n - K - 1} (1 - t)^{K-1}. \tag{E2}$$

The normalization is computed by dividing over the total surface, for $\varepsilon = 1$. This normalization is known as the regularized incomplete beta function, detailed in the next section. $\qquad\square$

Appendix F: Regularized incomplete beta function

The beta function is defined as

$$B(\alpha, \beta) = \int_0^1 t^{1-\alpha}(1-t)^{1-\beta}dt = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}, \tag{F1}$$

where $\Gamma(\cdot)$ is the Euler gamma function, functioning as an extension to the body of complex numbers of the factorial, that is

$$\Gamma(z) = \int_0^\infty t^{z-1}e^{-t}dt \tag{F2}$$

$$\Gamma(n) = (n-1)!; \quad n \in \mathbb{N}. \tag{F3}$$

The incomplete beta function is defined as

$$B(x; \alpha, \beta) = \int_0^x t^{1-\alpha}(1-t)^{1-\beta}dt. \tag{F4}$$

It is commonly regularized to adapt this function to the beta distribution [61] as

$$I(x; \alpha, \beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \int_0^x t^{1-\alpha}(1-t)^{1-\beta}dt. \tag{F5}$$

Appendix G: Proof of Lemma 7

We begin by detailing the linear combination of unitaries.

We first apply an auxiliary operation W on an ancilla register of size $\log B$ with the effect

$$W |0\rangle = \frac{1}{\sqrt{\|\alpha\|_1}} \sum_b \sqrt{\alpha_b} |b\rangle. \tag{G1}$$

This register controls the application of unitary operations generating each state $U_b |\psi_b\rangle$. By reverting the W operator we end up with a state

$$|\Psi\rangle = \frac{1}{\|\alpha\|_1} |0\rangle |\psi_{K,B}\rangle + \sqrt{1 - \frac{\| |\psi_{K,B}\rangle \|^2}{\|\alpha\|_1^2}} |1\rangle |\eta\rangle, \tag{G2}$$

with $|\psi_{K,B}\rangle = \sum_b \alpha_b U_b |\psi_b\rangle$ as in Equation (12). By creating $|\Psi\rangle$ and post-selecting on $|0\rangle$ in the first register, we obtain $|\psi\rangle$ in the second register with a certain success probability P_{success}.

We are only interested in the first term of the equation above, marked with $|0\rangle$, since it has the state we are interested in creating. The overlap of $|\Psi\rangle$ with $|0\rangle |\psi\rangle$ is given by

$$|\langle\Psi|\psi_{K,B}\rangle|^2 = \frac{\| |\psi_{K,B}\rangle \|^2}{\|\alpha\|_1^2}. \tag{G3}$$

Such success probability can be readily upper bounded as

$$\| |\psi\rangle \|^2 = \sum_b \alpha_b^2 \langle\psi_b|\psi_b\rangle + \sum_{a \neq b} \alpha_a \alpha_b \langle\psi_a| U_a^\dagger U_b |\psi_b\rangle. \tag{G4}$$

We consider now the block $F_{a,b}$ of the overlap matrix. The corresponding element can be upper bounded as

$$\langle\psi_a| U_a^\dagger U_b |\psi_b\rangle \leq \|F_{a,b}\|_\infty. \tag{G5}$$

On the other hand, we recall

$$\sum_{a \neq b} \alpha_a \alpha_b = \sum_{a,b} \alpha_a \alpha_b - \sum_b \alpha_b^2 = \|\alpha\|_1^2 - \|\alpha\|_2^2. \tag{G6}$$

Therefore, we can further simplify

$$|\langle\Psi|\psi_{K,B}\rangle|^2 \leq \frac{\|\boldsymbol{\alpha}\|_2^2}{\|\boldsymbol{\alpha}\|_1^2} \left(1 - \max_{a \neq b} \|F_{a,b}\|_\infty\right) + \max_{a \neq b} \|F_{a,b}\|_\infty. \tag{G7}$$

$$\square$$

Appendix H: Proof of Lemma 8

The starting point is the fixed-point amplification procedure from Ref. [37]. It is possible to load a state δ-close to $|\psi\rangle$ with a depth lower-bounded by

$$L \geq \frac{\log\left(2\delta^{-1}\right)}{\sqrt{P_{\text{success}}}}. \tag{H1}$$

Using the results from Lemma 7 it is straightforward to see that

$$L \geq \log\left(2\delta^{-1}\right) \frac{\|\boldsymbol{\alpha}\|_1}{\sqrt{B}\|\boldsymbol{\alpha}\|_\infty} \left(1 + \max_{a\neq b} \|F_{a,b}\|_\infty \frac{\|\boldsymbol{\alpha}\|_1^2}{\|\boldsymbol{\alpha}\|_2^2}\right)^{-1/2}. \tag{H2}$$

Making use of Taylor's expansion and the property

$$\frac{1}{\sqrt{1-x}} > 1 - \frac{x}{2} \quad \forall x > 0, \tag{H3}$$

we can bound

$$L \geq \log\left(2\delta^{-1}\right) \frac{\|\boldsymbol{\alpha}\|_1}{\sqrt{B}\|\boldsymbol{\alpha}\|_\infty} \left(1 - \frac{\max_{a\neq b} \|F_{a,b}\|_\infty}{2} \frac{\|\boldsymbol{\alpha}\|_1^2}{\|\boldsymbol{\alpha}\|_2^2}\right). \tag{H4}$$

Fixed-point amplification is protected against overcooking. Thus, we can apply a circuit with depth twice the previous lower-bound and maintain the guarantees.

From this point on, we just need to count. Each sparse state can be created in depth at most $d(K)$, and each basis change can be applied in at most $d(U)$. This is repeated for each basis, leading to a depth $(d(U) + d(K))B$. The fixed point amplification is given by L repetitions of the same circuits. Additionally, we need to count the depth of the operation W, $d(W)$. In total, we obtain a depth upper bounded by

$$D \leq d(W) + 2\left(d(U) + d(K)\right)\sqrt{B}\log(2/\delta)\frac{\|\boldsymbol{\alpha}\|_1^2}{\|\boldsymbol{\alpha}\|_2^2}\left(1 - \frac{F}{2}\left(\frac{\|\boldsymbol{\alpha}\|_1^2}{\|\boldsymbol{\alpha}\|_2^2} + 1\right)\right). \tag{H5}$$

$\square$

Appendix I: Proof of Theorem 2

We start by considering a probability distribution with N possible different outcomes. We sample this probability distribution M times. Let p_i be the probability of sampling outcome i, and let $\hat{p}_i = m_i/M$ be an estimator given by sampling. The number m_i is the number of appearances of the outcomes i. Without loss of generality, we assume $\hat{p}_i \geq \hat{p}_{i+1}$. Notice we cannot in principle assume any ordering in the true values p_i since those are not exactly accessible. For an integer K we define

$$\hat{S}_K = \sum_{i=1}^{K} \hat{p}_i, \tag{I1}$$

as the estimator of the joint probability of the K largest estimators of the probability.

The effect of these estimators $\hat{S}_K$ is to reduce a multinomial distribution into a binomial one, in which the outcomes are split in two classes. After splitting, we can estimate $\hat{S}_K$ using Wilson estimator [62]. The reason for Wilson estimator, instead of some other more standard versions such as Wald's [63], is that Wilson produces reliable confidence intervals even in the case of small $\hat{p}_i$ or $\hat{S}_K$. According to Wilson interval

$$S_K = \frac{1}{1 + \frac{z^2}{M}}\left(\hat{S}_K + \frac{z^2}{2M}\right) \pm \frac{z}{1 + \frac{z^2}{M}}\sqrt{\frac{\hat{S}_K(1 - \hat{S}_K)}{M} + \frac{z^2}{4M}}, \tag{I2}$$

with confidence probability $1 - \delta$, and z is the $\delta/2$-quantile of the standard normal distribution.

We consider now the estimators $\hat{S}_K$, $\hat{S}_{K+1}$. Each estimator gives rise to a probability distributions of compatible values of S_K/S_{K+1}. We need to obtain the requirements in $\hat{p}_{K+1}$ to avoid overlapping distributions of S_K/S_{K+1}. In the extremal case, since $\hat{S}_K < \hat{S}_{K+1}$, we have with probability at least $1 - \delta/2$

$$S_K \leq \quad \frac{1}{1 + \frac{z^2}{M}} \left(\hat{S}_K + \frac{z^2}{2M} \right) + \frac{z}{1 + \frac{z^2}{M}} \sqrt{\frac{\hat{S}_K(1 - \hat{S}_K)}{M} + \frac{z^2}{4M}}, \tag{I3}$$

$$S_{K+1} \leq \quad \frac{1}{1 + \frac{z^2}{M}} \left(\hat{S}_{K+1} + \frac{z^2}{2M} \right) - \frac{z}{1 + \frac{z^2}{M}} \sqrt{\frac{\hat{S}_{K+1}(1 - \hat{S}_{K+1})}{M} + \frac{z^2}{4M}}. \tag{I4}$$

Imposing $S_K < S_{K+1}$,

$$\hat{p}_{K+1} > z \left(\sqrt{\frac{\hat{S}_{K+1}(1 - \hat{S}_{K+1})}{M} + \frac{z^2}{4M}} + \sqrt{\frac{\hat{S}_K(1 - \hat{S}_K)}{M} + \frac{z^2}{4M}} \right). \tag{I5}$$

We can further simplify the previous equation with $0 \leq \hat{S}_{K+1} \leq 1$ to obtain

$$\hat{p}_{K+1} > \frac{1}{\sqrt{M}} \left(z\sqrt{4 + z^2} \right) > \frac{z^2}{\sqrt{M}}. \tag{I6}$$

The last step is to relate z and δ in a simpler way. First, we recall standard definitions

$$\frac{\delta}{2} = 1 - \text{erf}(z/\sqrt{2}). \tag{I7}$$

For sufficiently large z, e.g. $z \geq 2$, which corresponds to small $\delta \leq 0.1$, it is possible to bound

$$1 - \text{erf}(z/\sqrt{2}) \geq e^{-z^2}, \tag{I8}$$

which readily transforms into

$$z^2 \geq \log(2/\delta). \tag{I9}$$

Rearranging the equations, we arrive to the desired result

$$\hat{p}_{K+1} \geq \frac{\log(2/\delta)}{\sqrt{M}}. \tag{I10}$$

$\square$

Appendix J: Proof of Lemma 9

We begin by computing the distance between the true and estimated values of the a coefficients for one basis, namely

$$\Delta a_{i_b}^{(b)} \equiv |a_{i_b}^{(b)} - \hat{a}_{i_b}^{(b)}|^2 . \|\boldsymbol{a}^{(b)} - \boldsymbol{a}^{\hat{(b)}}\|_2^2 \leq \sum_{b \in B} \sum_{i \in S_b} |a_{i_b}^{(b)} - \hat{a}_{i_b}^{(b)}|^2. \tag{J1}$$

Assume that for every i we use a different number of measurements M_i, evenly dedicated to real and imaginary parts via Hadamard tests. Using Wald estimators [63] we obtain with probability $1 - \delta$

$$\text{Re}(a) = \text{Re}(\hat{a}) \pm z\sqrt{\frac{1 - \text{Re}(\hat{\alpha})^2}{2M_i}}, \tag{J2}$$

with z being the $1 - \delta/2$ quantile of the normal distribution. This readily implies

$$\Delta \text{Re}(a) \leq \frac{z}{\sqrt{2M_i}}, \tag{J3}$$

and equivalently for the imaginary part.

We recall now standard error propagation to bound the joint error of real and imaginary parts as

$$\Delta |a|^2 = 2|\operatorname{Re}(\hat{a})|\Delta \operatorname{Re}(\hat{a}) + 2|\operatorname{Im}(\hat{a})|\Delta \operatorname{Im}(\hat{a}) \leq (|\operatorname{Re}(\hat{a})| + |\operatorname{Im}(\hat{a})|)\frac{\sqrt{2}z}{\sqrt{M_i}}. \tag{J4}$$

The Chauchy-Schwarz inequality for d-dimensional vectors implies $\|a\|_1 \leq \sqrt{d}\|a\|_2$, which immediately implies $(|\operatorname{Re}(\hat{a})| + |\operatorname{Im}(\hat{a})|) \leq \sqrt{2}\sqrt{\hat{p}_i}$. We can exploit the fact that we have the estimation $\sqrt{\hat{p}_i} = m_i/\hat{S}_k$ from previous sampling, thus implying

$$\Delta p \leq \sqrt{\frac{\hat{p}_i}{S_K}}\frac{2z}{\sqrt{M_i}}. \tag{J5}$$

By selecting $M_i = M\hat{p}_i/\hat{S}_K$, we obtain a bound

$$\Delta |a|^2 \leq \frac{2z}{\sqrt{M}}. \tag{J6}$$

We can readily bound

$$\|\boldsymbol{a}^{(b)} - \boldsymbol{a}^{\hat{(b)}}\|_2^2 \leq \frac{2zK}{\sqrt{M}}. \tag{J7}$$

Identifying

$$\delta \leq 2e^{-z^2/2}, \tag{J8}$$

we can rewrite

$$\|\boldsymbol{a}^{(b)} - \boldsymbol{a}^{\hat{(b)}}\|_2^2 \leq 2\sqrt{2}KB\sqrt{\frac{\log 2/\delta}{M}}, \tag{J9}$$

obtaining the desired result. $\qquad \square$