Probabilistic unitary and state synthesis with optimal accuracy

Seiseki Akibue[1] * Go Kato[2] Seiichiro Tani[3]

[1] *NTT Communication Science Labs., NTT Corporation, 3–1, Morinosato-Wakamiya, Atsugi, Kanagawa 243-0198, Japan.*
[2] *Advanced ICT Research Institute, NICT, 4–2–1, Nukui-Kitamachi, Koganei, Tokyo 184-8795, Japan.*
[3] *Department of Mathematics, Waseda University, 1–6–1, Nishi-Waseda, Shinjuku, Tokyo 169-8050, Japan.*

Abstract. A new circuit synthesis approach, called probabilistic synthesis, samples a gate sequence to suppress the approximate error for realizing a target unitary transformation or preparing a target pure state. When we consider the fault-tolerant circuit synthesis, the description of the approximation error caused by each gate sequence is completely known. By exploiting this description, several probabilistic synthesis algorithms have been developed to reduce the approximation error for specific types of target unitary transformations; however, their fundamental limitations were unknown. In our research, we have revealed the tight inequalities that govern the error reduction achieved by the optimal probabilistic synthesis for realizing any unitary transformations and preparing pure states. In contrast to the existing results, our result shows that the approximation error can be reduced not only quadratically, but also proportionally to $1/d$ for some target unitary transformations acting on d-dimensional Hilbert space. From a computational point of view, we have shown that an optimal probabilistic synthesis algorithm can be constructed based on a semidefinite program (SDP). We have also developed a technique to exponentially reduce the running time of the SDP by exploiting the spherical nature of unitary transformations and pure states. As a result, we have constructed efficient probabilistic synthesis algorithms for realizing arbitrary unitary transformations and preparing pure states, rigorously estimated their time complexity, and demonstrated their performance through numerical experiments.

Keywords: unitary synthesis, state synthesis, circuit compilation, convex approximation, probabilistic approximation

1 Background

To realize information processing in a gate model quantum computer, we need to prepare an initial state and perform unitary transformations on a fixed-size system with the desired accuracy. This is possible by exploiting quantum error correction [1] or the nature of the system [2]. These techniques allow for fault-tolerant implementation of a finite gate set, such as $\{H, T, CNOT\}$, with negligible implementation errors. However, this requires us to prepare a target pure state ϕ and realize a target unitary transformation Υ using a circuit formed from the finite gate set. As a result of the discretization, we can only prepare an approximated pure state $\hat{\phi}$ or realize an approximated unitary transformation $\hat{\Upsilon}$ in general.

To suppress the effect of the overhead caused by the fault-tolerant implementation of each gate, various *synthesis* algorithms [3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15] have been proposed for minimizing the error or the circuit size. Following the celebrated Solovay-Kitaev algorithm [3], the final goal of conventional synthesis algorithms is to deterministically find one of the best circuits for the approximate realization of a unitary transformation or preparation of a pure state. Thus, the minimum approximation error obtained by such *deterministic* unitary (or state) synthesis is given by $\min_{x\in X} \frac{1}{2}\left\|\Upsilon - \hat{\Upsilon}_x\right\|_\diamond$ (or $\min_{x\in X}\left\|\phi - \hat{\phi}_x\right\|_{\mathrm{tr}}$), where X is the label set of unitary transformations (or pure states) realized (or prepared) by circuits with a certain cost, e.g., the circuit size, depth, or the number of T gates.

While it makes sense to approximate a target unitary transformation (or pure state) by utilizing an approximated unitary (or pure state) generated by a single circuit, a recently proposed approach called *probabilistic synthesis* probabilistically samples a circuit for the approximation. Suppose that the probabilistic algorithm independently samples a circuit $\mathcal{C}_x$ implementing $\hat{\Upsilon}_x$ (or preparing $\hat{\phi}_x$) in accordance with a probability distribution $p(x)$ each time Υ (or ϕ) is required in quantum information processing. Then, each realized physical transformation (or prepared state) is described by $\sum_x p(x)\hat{\Upsilon}_x$ (or $\sum_x p(x)\hat{\phi}_x$). This can be interpreted as the transition from coherent errors to incoherent errors [16, 17, 18], and recent studies have experimentally demonstrated that this transition reduces the approximation error of pure states [19]. Moreover, Hastings [16], Campbell [20] and Vadym et al. [21] constructed synthesis algorithms that reduce the approximation error of unitary transformations into $\frac{1}{2}\left\|\Upsilon - \sum_x p(x)\hat{\Upsilon}_x\right\|_\diamond = O(\epsilon^2)$ by choosing $p(x)$ appropriately, where $\{\hat{\Upsilon}_x\}_x$ causes the worst approximation error $\epsilon := \max_\Upsilon \min_x \frac{1}{2}\left\|\Upsilon - \hat{\Upsilon}_x\right\|_\diamond$ if the deterministic synthesis is used. It is important to note that the probabilistic synthesis does not need multiple executions of $\mathcal{C}_x$ for a single realization of Υ (or a single preparation of ϕ), unlike quantum error mitigation [22]. This is because the reduction of the diamond norm guarantees that the probability distribution of the outputs of a quantum information processing task becomes

*seiseki.akibue@ntt.com

closer to the ideal one with just a single sample of $\mathcal{C}_x$.

Despite its potential in a wide range of applications, the limitation of probabilistic synthesis, especially the minimum approximation error $\min_p \frac{1}{2}\left\|\Upsilon - \sum_x p(x)\hat{\Upsilon}_x\right\|_\diamond$ (or $\min_p \left\|\phi - \sum_x p(x)\hat{\phi}_x\right\|_{\rm tr}$), remains unknown, nor is it clear how to find the optimal probability distribution p. While Campbell [20] has demonstrated that the approximation error can be reduced to $10\epsilon^2$ using probabilistic synthesis for a small ϵ for any target unitary transformation, it has been shown that it can be reduced to ϵ^2 for a unitary transformation corresponding to an axial rotation [21]. These results indicate that the synthesis algorithm for general unitary transformations [20] is not optimal. While a few analytical results about the minimum approximation error are obtained for the case of a qubit transformation [23] or state [24, 25, 26] in the context of the optimal convex approximation of a quantum transformation or state, minimax optimization to compute the minimum approximation error makes analyses quite difficult in general.

Note that the result of optimal probabilistic unitary synthesis does not contain that of state synthesis, and vice versa. This is because the generated state in state synthesis is obtained by applying a unitary transformation to a fixed input state $|0\rangle$ while the approximation error in unitary synthesis is quantified for the worst input state. Moreover, a target state could be approximated by probabilistically mixing two unitary transformations whose behaviors are totally different, except for $|0\rangle$.

2 Our contributions

Before presenting our results, we provide intuitive examples demonstrating the capability of probabilistic synthesis in Fig. 1. As a generalization of the qubit examples, we obtain the fundamental relationship between the minimum approximation errors obtained by the deterministic synthesis and the probabilistic one:

Theorem 1 (simplified version) *[27, Theorem 1]*
For any subset $\{\hat{\phi}_x\}_{x\in X}$ of pure states in a finite-dimensional Hilbert space, it holds that

$$\max_\phi \min_p \left\|\phi - \sum_{x\in X} p(x)\hat{\phi}_x\right\|_{tr} = \max_\phi \min_{x\in X} \left\|\phi - \hat{\phi}_x\right\|_{tr}^2, \tag{1}$$

where the maximization of ϕ is taken over the set of pure states.

This theorem compares the *worst* approximation errors occurring when one synthesizes the target state ϕ that is most difficult to approximate by using $\{\hat{\phi}_x\}_{x\in X}$. It implies that the optimal probabilistic synthesis always quadratically reduces the worst approximation error, moreover, it is impossible to further reduce the approximation error.

In many cases, there is no need to synthesize all possible pure states. Instead, it is more useful to understand the limitations of probabilistic synthesis when a target state is chosen from a subset S_G of pure states. As shown in Fig. 1(b), we can also anticipate the quadratic error reduction in this scenario. This expectation is confirmed in the comprehensive version of Theorem 1, which includes the case of Fig. 1(b).

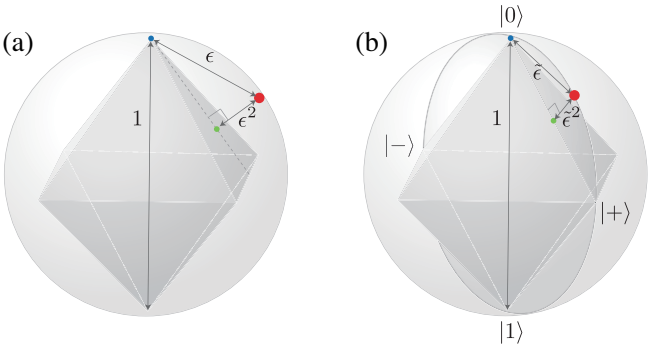

Figure 1: Quadratic reduction of the approximation error by using probabilistic synthesis. We assume that we can exactly generate an eigenstate $\hat{\phi}_x$ of the Pauli operators, represented by the six extreme points of the octahedron. We represent the Bloch sphere by a sphere with radius $\frac{1}{2}$, where the trace distance between two quantum states equals the Euclidean distance between the corresponding points. (a) We can compute $\min_p \left\|\phi - \sum_x p(x)\hat{\phi}_x\right\|_{\rm tr} = \epsilon^2 = \frac{1}{2\sqrt{3}}\left(\sqrt{3}-1\right)$ and $\min_x \left\|\phi - \hat{\phi}_x\right\|_{\rm tr} = \epsilon$, where ϕ is the furthest state from $\{\hat{\phi}_x\}_{x=1}^6$, represented as a large red point. (b) Suppose that the target state is chosen from $S_G := \{\phi : |\phi\rangle = \cos t|0\rangle + \sin t|1\rangle, t \in \mathbb{R}\}$, represented by a meridian. We can compute $\min_p \left\|\phi - \sum_x p(x)\hat{\phi}_x\right\|_{\rm tr} = \tilde{\epsilon}^2 = \frac{1}{2}\left(1 - \frac{1}{\sqrt{2}}\right)$ and $\min_x \left\|\phi - \hat{\phi}_x\right\|_{\rm tr} = \tilde{\epsilon}$, where ϕ is the furthest state in S_G from $\{\hat{\phi}_x\}_{x=1}^6$, represented as a large red point.

In the case of unitary synthesis, we obtain the following theorem.

Theorem 2 (simplified version) *[28, Theorem 4.3]*
For an integer $d \geq 2$ specified below, let Υ and $\{\hat{\Upsilon}_x\}_{x\in X}$ be a target unitary transformation and a finite set of unitary transformations on $\mathbb{C}^d$, respectively. It then holds that

$$\frac{4\delta}{d}\left(1 - \frac{\delta}{d}\right) \leq \max_\Upsilon \min_p \frac{1}{2}\left\|\Upsilon - \sum_{x\in X} p(x)\hat{\Upsilon}_x\right\|_\diamond \leq \epsilon^2$$
$$\text{with} \begin{cases} \delta = 1 - \sqrt{1-\epsilon^2} & \text{and} \\ \epsilon = \max_\Upsilon \min_{x\in X} \frac{1}{2}\left\|\Upsilon - \hat{\Upsilon}_x\right\|_\diamond . \end{cases} \tag{2}$$

This theorem provides bounds on the worst approximation error caused when one probabilistically synthesizes the target unitary that is most difficult to approximate. The gap between the upper and lower bounds exists if and only if $d \geq 3$. We can show that the gap is inevitable by constructing $\{\hat{\Upsilon}_x\}_x$ for achieving the upper bound and that for achieving the lower bound for any d and ϵ. These examples show that Ineqs. (2) are tight.

While the upper bound in this theorem is essentially the same as Theorem 1, the lower bound is totally different. We can easily verify that $\frac{4\delta}{d}\left(1-\frac{\delta}{d}\right) \simeq \frac{2}{d}\epsilon^2$ for small ϵ. This demonstrates that the existing probabilistic synthesis algorithms [16, 20, 21] are far from optimal since the approximation error can be reduced not only quadratically, but also proportionally to $1/d$ by the optimal synthesis for some target unitary transformations.

As a sharper lower bound on the approximation error attained by probabilistic synthesis, we can use the following inequality

$$\min_p \frac{1}{2}\left\|\Upsilon - \sum_x p(x)\hat{\Upsilon}_x\right\|_\diamond \geq 1 - \frac{1}{d^2}\max_{x\in X}\left|\mathrm{tr}\left[U^\dagger \hat{U}_x\right]\right|^2 \tag{3}$$

derived in the proof of Theorem 2, where U (or $\hat{U}_x$) represents a unitary operator realizing Υ (or $\hat{\Upsilon}_x$). This inequality is shaper than the lower bound in Ineqs. (2) since the latter is derived by using the former.

Note that an equivalent expression of this inequality is given by slightly modifying the result [29, Proposition 9] of Wallman et al. as follows:

$$\min_p \frac{1}{2}\left\|\Upsilon - \sum_x p(x)\hat{\Upsilon}_x\right\|_\diamond \geq \min_{x\in X}\frac{d+1}{d}r(\Upsilon^{-1}\circ\hat{\Upsilon}_x), \tag{4}$$

where $r(\mathcal{E}) := 1 - \int d\phi \mathrm{tr}\left[\mathcal{E}(\phi)\phi\right]$ represents an average gate infidelity of a CPTP mapping $\mathcal{E}$. The equivalence between Ineq. (3) and Ineq. (4) can be verified by a straightforward calculation as provided in Appendix A. Thus, in order to achieve large error reduction, the average gate infidelity of $\Upsilon^{-1}\circ\hat{\Upsilon}_x$ must be small for some $x \in X$.

The example below highlights the difference between the lower bounds in Ineq. (2) and Ineq. (3). Suppose the eigenvalues of $U^\dagger\hat{U}_x$ consist of $e^{i\epsilon}$ and $e^{-i\epsilon}$ with small ϵ. We find that $\frac{1}{2}\left\|\Upsilon - \hat{\Upsilon}_x\right\|_\diamond \simeq \epsilon$. Thus, Ineq. (2) implies that the approximation error is lower bounded by $\simeq \frac{2}{d}\epsilon^2$ (regardless of the degeneracy of the eigenvalues). In contrast, Ineq. (3) implies that the approximation error is lower bounded by $\simeq \frac{4n_+n_-}{d^2}\epsilon^2$, where n_+ and n_- represent the degeneracy of $e^{i\epsilon}$ and $e^{-i\epsilon}$, respectively. This indicates that achieving large error reduction requires biased degeneracy. This example and the upper bound in Ineq. (2) demonstrate that the average gate infidelity of $\Upsilon^{-1}\circ\hat{\Upsilon}_x$ and the diamond norm $\frac{1}{2}\left\|\Upsilon - \hat{\Upsilon}_x\right\|_\diamond$ serve as complementary measures for estimating the approximation error attained by the optimal probabilistic synthesis.

From a computational point of view, we show that the optimal probability distribution $\{p(x)\}_{x\in X}$ for approximating Υ (or ϕ) can be computed by the semidefinite program (SDP) when $\{\hat{\Upsilon}_x\}_{x\in X}$ (or $\{\hat{\phi}_x\}_{x\in X}$) realized by using a gate sequence is given. (This set is computable with certain deterministic synthesis algorithms.) The ellipsoid method guarantees that the optimal $\{p(x)\}_{x\in X}$ minimizing $\frac{1}{2}\left\|\Upsilon - \sum_x p(x)\hat{\Upsilon}_x\right\|_\diamond$ (or $\left\|\phi - \sum_x p(x)\hat{\phi}_x\right\|_{tr}$) can be computed in $poly\left(|X|d\right)$-time. However, as explained later, this running time is too long to construct an efficient synthesis algorithm. We resolve this problem by proving the following lemma for reducing the running time of the SDP.

Lemma 3 (simplified version) *[28, Lemma 5.3]*
When a set of single-qubit unitary transformations $\{\hat{\Upsilon}_x\}_{x\in X}$ forms an ϵ-covering, i.e., $\max_\Upsilon \min_{x\in X}\frac{1}{2}\left\|\Upsilon - \hat{\Upsilon}_x\right\|_\diamond \leq \epsilon$, then

$$\min_p \left\|\Upsilon - \sum_{x\in X} p(x)\hat{\Upsilon}_x\right\|_\diamond = \min_{\hat{p}} \left\|\Upsilon - \sum_{x\in\hat{X}} \hat{p}(x)\hat{\Upsilon}_x\right\|_\diamond \tag{5}$$

holds, where $\hat{X} := \left\{x \in X : \frac{1}{2}\left\|\Upsilon - \hat{\Upsilon}_x\right\|_\diamond \leq 2\epsilon\right\}$.

We can prove the similar lemma for the case of convex approximation of pure states [27, Lemma 4]. Note that the lemma for states holds for any dimensional Hilbert space, while we have only proven Lemma 3 for single-qubit unitary transformations.

In the Introduction, it is mentioned that X is the label set of unitary transformations (or pure states) that can be achieved by circuits with a certain cost. If we consider the cost as the number of T gates (known as T-count), then the size of X exponentially grows with respect to the T-count. When we aim to synthesize a single-qubit unitary transformation within an error of ϵ, it has been established that the T-count grows as $O\left(\log\left(\frac{1}{\epsilon}\right)\right)$ [30]. These imply that $\{\hat{\Upsilon}_x\}_{x\in X}$ forms an ϵ-covering when $|X| = poly\left(\frac{1}{\epsilon}\right)$. In this case, the running time of our SDP to compute the optimal $\{p(x)\}_{x\in X}$ becomes $poly\left(\frac{1}{\epsilon}\right)$, which is not efficient as a synthesis algorithm. However, Lemma 3 ensures that we can confine X to $\hat{X}$ in our SDP without increasing the approximation error. Since $|\hat{X}|$ is typically upper bounded by a constant independent of ϵ, the running time of this modified SDP is significantly reduced. Moreover, the quadratic error reduction is guaranteed due to Theorem 2 since $\{\hat{\Upsilon}_x\}_{x\in X}$ forms an ϵ-covering.

Based on this modified SDP, we construct a probabilistic synthesis algorithm for unitary transformations in the following theorems:

Theorem 4 (informal version) *[28, Theorem 5.4]*
For a given gate set, there exists a probabilistic synthesis algorithm for a single-qubit unitary with
INPUT*: a target single-qubit unitary Υ and target approximation error $\epsilon \in (0,1)$*
OUTPUT*: a gate sequence implementing a single-qubit unitary $\hat{\Upsilon}_x$ sampled from a set $\{\hat{\Upsilon}_x\}_x$ in accordance with probability distribution $\hat{p}(x)$.*
such that the algorithm satisfies the following properties:

- Efficiency*: The algorithm calls a deterministic synthesis algorithm constant times and the whole running time is polylog $\left(\frac{1}{\epsilon}\right)$,*
- Quadratic improvement*: The approximation error $\frac{1}{2}\left\|\Upsilon - \sum_x \hat{p}(x)\hat{\Upsilon}_x\right\|_\diamond$ obtained with this algorithm is upper bounded by ϵ^2, whereas the error*

$\min_x \frac{1}{2}\left\|\Upsilon - \hat{\Upsilon}_x\right\|_\diamond$ *obtained by deterministic synthesis using the unitaries in* $\{\hat{\Upsilon}_x\}_x$ *is upper bounded by* ϵ.

It is important to highlight that the algorithm in this theorem works for any dimensional unitary transformations with the same guarantee of efficiency and quadratic error reduction if we can prove Lemma 3 for such unitary transformations. In the next subsection, we perform numerical experiments to support that Lemma 3 holds for higher dimensional unitary transformations.

In the case of state synthesis, we construct a similar algorithm that is applicable to any dimensional pure states and satisfies the same two properties as the unitary synthesis algorithm [27, Theorem 2]. (In the case of state synthesis, the second property is measured by the trace distance.) Moreover, it can be extended into the case when a target pure state is restricted on S_G, as shown in Fig. 1 (b).

2.1 Numerical experiments

First, we show how the probabilistic synthesis algorithm, as provided in the state version of Theorem 4, can reduce the T-count through a numerical experiment. We select a target state ϕ from $S_G = \{\phi \in \mathbf{P}\left(\mathbb{C}^2\right) : |\phi\rangle = \cos t|0\rangle + \sin t|1\rangle, t \in \mathbb{R}\}$, as shown in Fig. 1(b). As demonstrated in Fig. 2, our probabilistic synthesis algorithm reduces the approximation error quadratically, consequently halving the T-count needed to achieve a certain approximation error.

Secondly, we numerically support that Lemma 3 holds for higher dimensional unitary transformations. In the following numerical experiment, we first construct ϵ-coverings $\{\hat{\Upsilon}_x\}_{x\in X}$ of d-dimensional unitary transformations by randomly choosing $|X| = 10^5$ and $|X| = 10^7$ unitary operators for $d = 2$ and $d = 4$, respectively. We interpret $\{\hat{\Upsilon}_x\}_{x\in X}$ as the set of available unitary transformations in probabilistic and deterministic approximation.

Next, we randomly choose a target unitary transformation Υ and compute the approximation error $\epsilon_\Upsilon := \min_{x\in X} \frac{1}{2}\left\|\Upsilon - \hat{\Upsilon}_x\right\|_\diamond$ attained by the deterministic synthesis. We define the approximation error attained by probabilistically mixing restricted available unitary transformations as

$$\epsilon_\Upsilon^{prob}(\epsilon') := \min_p \frac{1}{2}\left\|\Upsilon - \sum_{x\in\hat{X}(\epsilon')} p(x)\hat{\Upsilon}_x\right\|_\diamond \tag{6}$$

where $\hat{X}(\epsilon') := \{x \in X : \frac{1}{2}\left\|\Upsilon - \hat{\Upsilon}_x\right\|_\diamond \leq \epsilon'\}$. Note that $\epsilon_\Upsilon^{prob}(\epsilon')$ is a monotonically decreasing function. Moreover, if Lemma 3 holds for any dimensional unitary transformations, $\epsilon_\Upsilon^{prob}(2\epsilon) = \epsilon_\Upsilon^{prob}(1)$.

In Fig. 3, we draw the graphs of $\epsilon_\Upsilon^{prob}(\epsilon')$ for 10 randomly chosen Υ by using different colors corresponding to Υ. We can observe that $\epsilon_\Upsilon^{prob}(\epsilon')$ is saturated when $\epsilon' \geq 1.4\epsilon$ and ϵ_Υ^{prob} is comparable or smaller than ϵ_Υ^2 since $\frac{\log(\epsilon_\Upsilon^{prob}(\epsilon'))}{\log(\epsilon_\Upsilon)} \geq 2 \Leftrightarrow \epsilon_\Upsilon^{prob}(\epsilon') \leq \epsilon_\Upsilon^2$. Note that some instances satisfying $\epsilon_\Upsilon^{prob}(\epsilon') > \epsilon_\Upsilon^2$ does not violate the right inequality of Theorem 2 since $\epsilon_\Upsilon \leq \max_\Upsilon \epsilon_\Upsilon = \epsilon$.

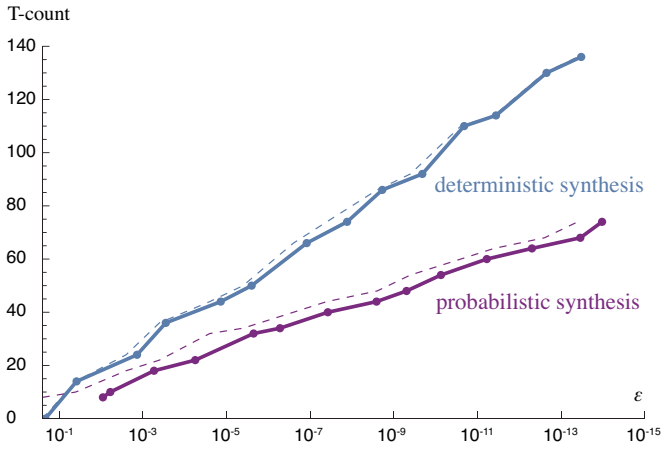

Figure 2: **Relationship between T-count and the approximation error for synthesizing $|\phi\rangle = \cos t|0\rangle + \sin t|1\rangle$ with $t = 1$.** For each target approximation error, we run the Ross-Selinger algorithm [30] to obtain a gate sequence to approximate ϕ. The blue dashed line interpolates points, each of which represents a target approximation error and the T-count of the gate sequence. The actual approximation error and the T-count achieved by the gate sequence are plotted by blue dots. Note that both the target and actual approximation errors are represented by ϵ. For each of the target approximation errors, we run the probabilistic synthesis algorithm and obtain a list of six gate sequences to be probabilistically sampled. The purple dashed line interpolates points, each of which represents a target approximation error and the maximum T-count of gate sequences in the list. The actual approximation error and the maximum T-count achieved by optimally mixing the gate sequence are plotted by purple dots.

3 Technical Outline

In the proof of Theorem 1, we analyze the minimum approximation error

$$\min_p \left\|\rho - \sum_x p(x)\hat{\rho}_x\right\|_{\mathrm{tr}} = \min_p \max_{0\leq M\leq\mathbb{I}} \mathrm{tr}\left[M(\rho - \sum_x p(x)\hat{\rho}_x)\right] \tag{7}$$

for general mixed states ρ and ρ_x, which contains minimax optimization by definition. The main tool for the analysis is the strong duality of semidefinite programming. This enables us to formulate the minimum approximation error as an SDP. This analysis establishes a general lemma [27, Lemma 3] about the optimal convex approximation of a quantum state by using a restricted subset of states [24, 25, 26]. While the optimal convex approximation and state synthesis have been studied in different contexts, our lemma has made an important contribution to both topics. First, it has revealed the tight inequalities on the optimal approximation error in

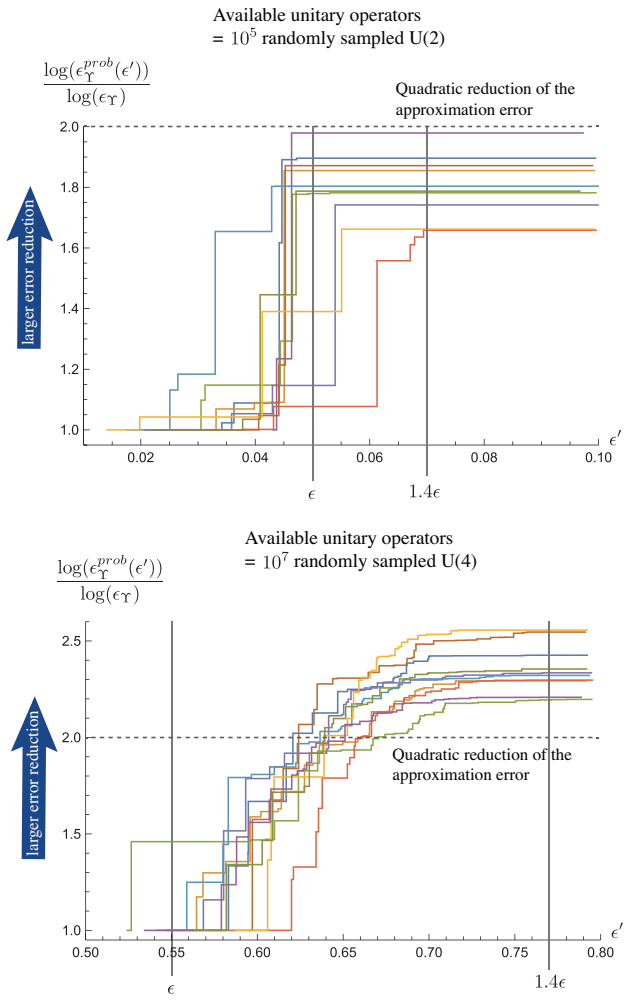

Figure 3: The comparison between the approximation error $\epsilon_\Upsilon^{prob}(\epsilon')$ attained by the probabilistic approximation and that ϵ_Υ attained by the deterministic approximation for 10 randomly sampled target unitary transformations Υ. For both approximations, we use the set of available unitary transformations induced by an ϵ-covering of the set $U(d)$ of d-dimensional unitary operators.

probabilistic synthesis as shown in Theorem 1. Secondly, it is applicable to compute a resource measure in convex resource theories [31, 32, 33] such as the resource theory of entanglement and coherence, as shown in [27].

Such reformulation based on an SDP becomes more complicated but also crucial in the proof of Theorem 2. In this case, we construct an SDP to compute the minimum approximation error $\min_p \|\mathcal{A} - \sum_x p(x)\mathcal{B}_x\|_\diamond$ for any CPTP mappings $\mathcal{A}$ and $\mathcal{B}_x$. In this research, we focus on the case when both $\mathcal{A}$ and $\mathcal{B}_x$ are unitary transformations. However, an SDP for general CPTP mappings is valuable for optimizing probabilistic synthesis including measurement feedbacks [34, 21, 35]. Additionally, such an SDP might be beneficial for improving the simulation of non-unitary dynamics.

As preiously mentioned, Lemma 3 is crucial for constructing efficient synthesis algorithms. This property can be intuitively understood by using the examples shown in Fig. 1. If the goal is to optimally approximate a target state ϕ depicted by the red point in (a) (or (b)), it is sufficient to mix three (or two) Pauli eigenstates that are 2ϵ (or $2\bar{\epsilon}$) close to ϕ. As suggested by this example, the spherical nature of the set of pure states with respect to the trace distance is essential. For the case of pure states, we can prove this property for any dimensional pure states.

However, in the case of unitary transformation, the metric space induced by the diamond norm is more complicated. We have only succeeded in proving this property for single-qubit unitary transformations by exploiting the *magic basis* [36] representation of single-qubit unitary operators. The magic basis representation enables us to embed the metric space of single-qubit unitary transformations into that of S^3 with respect to the angle. While numerical experiments indicate the same fact holds for qudit unitary transformations, the rigorous proof is a subject for future work.

References

[1] B. M. Terhal, "Quantum error correction for quantum memories," *Rev. Mod. Phys.*, vol. 87, pp. 307–346, Apr 2015.

[2] A. Kitaev, "Fault-tolerant quantum computation by anyons," *Annals of Physics*, vol. 303, no. 1, pp. 2–30, 2003.

[3] A. Y. Kitaev, A. H. Shen, and M. N. Vyalyi, *Classical and Quantum Computation*. American Mathematical Society, 2002.

[4] A. W. Harrow, B. Recht, and I. L. Chuang, "Efficient discrete approximations of quantum gates," *Journal of Mathematical Physics*, vol. 43, no. 9, pp. 4445–4451, 2002.

[5] V. Kliuchnikov, D. Maslov, and M. Mosca, "Practical approximation of single-qubit unitaries by single-qubit quantum clifford and t circuits," *IEEE Transactions on Computers*, vol. 65, no. 1, pp. 161–172, 2016.

[6] V. Kliuchnikov, D. Maslov, and M. Mosca, "Asymptotically optimal approximation of single qubit unitaries by clifford and t circuits using a constant number of ancillary qubits," *Phys. Rev. Lett.*, vol. 110, p. 190502, May 2013.

[7] N. J. Ross, "Optimal ancilla-free clifford+v approximation of z-rotations," *Quantum Info. Comput.*, vol. 15, pp. 932–950, sep 2015.

[8] A. Bocharov, M. Roetteler, and K. M. Svore, "Efficient synthesis of universal repeat-until-success quantum circuits," *Phys. Rev. Lett.*, vol. 114, p. 080502, Feb 2015.

[9] A. G. Fowler, "Constructing arbitrary steane code single logical qubit fault-tolerant gates," *Quantum Info. Comput.*, vol. 11, pp. 867–873, sep 2011.

[10] A. Bouland and T. Giurgica-Tiron, "Efficient universal quantum compilation: An inverse-free solovay-kitaev algorithm." arXiv:2112.02040.

[11] N. Mahmud, A. MacGillivray, M. Chaudhary, and E. El-Araby, "Optimizing quantum circuits for arbitrary state synthesis and initialization," in *2021 IEEE 34th International System-on-Chip Conference (SOCC)*, pp. 19–24, 2021.

[12] Z. Zhang, Q. Wang, and M. Ying, "Parallel quantum algorithm for hamiltonian simulation," 2023. arXiv:2105.11889.

[13] X.-M. Zhang, M.-H. Yung, and X. Yuan, "Low-depth quantum state preparation," *Phys. Rev. Res.*, vol. 3, p. 043200, Dec 2021.

[14] K. Fukui, S. Takeda, M. Endo, W. Asavanant, J.-i. Yoshikawa, P. van Loock, and A. Furusawa, "Efficient backcasting search for optical quantum state synthesis," *Phys. Rev. Lett.*, vol. 128, p. 240503, Jun 2022.

[15] S. Ashhab, N. Yamamoto, F. Yoshihara, and K. Semba, "Numerical analysis of quantum circuits for state preparation and unitary operator synthesis," *Phys. Rev. A*, vol. 106, p. 022426, Aug 2022.

[16] M. B. Hastings, "Turning gate synthesis errors into incoherent errors," *Quantum Info. Comput.*, vol. 17, pp. 488–494, mar 2017.

[17] J. J. Wallman and J. Emerson, "Noise tailoring for scalable quantum computation via randomized compiling," *Phys. Rev. A*, vol. 94, p. 052325, Nov 2016.

[18] K. Kuroiwa and Y. O. Nakagawa, "Clifford+t-gate decomposition with limited number of t gates, its error analysis, and performance of unitary coupled cluster ansatz in pre-ftqc era," 2023. arXiv:2301.04150.

[19] A. Hashim, R. K. Naik, A. Morvan, J.-L. Ville, B. Mitchell, J. M. Kreikebaum, M. Davis, E. Smith, C. Iancu, K. P. O'Brien, I. Hincks, J. J. Wallman, J. Emerson, and I. Siddiqi, "Randomized compiling for scalable quantum computing on a noisy superconducting quantum processor," *Phys. Rev. X*, vol. 11, p. 041039, Nov 2021.

[20] E. Campbell, "Shorter gate sequences for quantum computing by mixing unitaries," *Phys. Rev. A*, vol. 95, p. 042306, Apr 2017.

[21] V. Kliuchnikov, K. Lauter, R. Minko, A. Paetznick, and C. Petit, "Shorter quantum circuits," 2022. arXiv:2203.10064.

[22] K. Temme, S. Bravyi, and J. M. Gambetta, "Error mitigation for short-depth quantum circuits," *Phys. Rev. Lett.*, vol. 119, p. 180509, Nov 2017.

[23] M. F. Sacchi and T. Sacchi, "Convex approximations of quantum channels," *Phys. Rev. A*, vol. 96, p. 032311, Sep 2017.

[24] M. F. Sacchi, "Optimal convex approximations of quantum states," *Phys. Rev. A*, vol. 96, p. 042325, Oct 2017.

[25] X.-B. Liang, B. Li, L. Huang, B.-L. Ye, S.-M. Fei, and S.-X. Huang, "Optimal approximations of available states and a triple uncertainty relation," *Phys. Rev. A*, vol. 101, p. 062106, Jun 2020.

[26] L. qiang Zhang, D. hui Yu, and C. shui Yu, "The optimal approximation of qubit states with limited quantum states," *Physics Letters A*, vol. 398, p. 127286, 2021.

[27] S. Akibue, G. Kato, and S. Tani, "Probabilistic state synthesis based on optimal convex approximation," *npj Quantum Information*, vol. 10, no. 1, p. 3, 2024.

[28] S. Akibue, G. Kato, and S. Tani, "Probabilistic unitary synthesis with optimal accuracy," *ACM Transactions on Quantum Computing*, may 2024. Just Accepted.

[29] J. J. Wallman and S. T. Flammia, "Randomized benchmarking with confidence," *New Journal of Physics*, vol. 16, p. 103032, oct 2014.

[30] N. J. Ross and P. Selinger, "Optimal ancilla-free clifford+t approximation of z-rotations," *Quantum Info. Comput.*, vol. 16, pp. 901–953, sep 2016.

[31] M. HORODECKI and J. OPPENHEIM, "(quantumness in the context of) resource theories," *International Journal of Modern Physics B*, vol. 27, no. 01n03, p. 1345019, 2013.

[32] F. G. S. L. Brandão and G. Gour, "Reversible framework for quantum resource theories," *Phys. Rev. Lett.*, vol. 115, p. 070503, Aug 2015.

[33] E. Chitambar and G. Gour, "Quantum resource theories," *Reviews of modern physics*, vol. 91, no. 2, p. 025001, 2019.

[34] A. Paetznick and K. M. Svore, "Repeat-until-success: non-deterministic decomposition of single-qubit unitaries," *Quantum Info. Comput.*, vol. 14, pp. 1277–1301, nov 2014.

[35] N. Yoshioka, S. Akibue, H. Morisaki, K. Tsubouchi, and Y. Suzuki, "Error crafting in probabilistic quantum gate synthesis," 2024.

[36] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters, "Mixed-state entanglement and quantum error correction," *Phys. Rev. A*, vol. 54, pp. 3824–3851, Nov 1996.

A Proof

In this section, we show the equivalence between Ineq. (3) and Ineq. (4). Concretely we show

$$1-\frac{1}{d^2}\left|\mathrm{tr}\left[U^\dagger \hat{U}_x\right]\right|^2=\frac{d+1}{d}r(\Upsilon^{-1}\circ\hat{\Upsilon}_x), \tag{8}$$

where U (or $\hat{U}_x$) represents a unitary operator realizing Υ (or $\hat{\Upsilon}_x$) and $r(\mathcal{E}):=1-\int d\phi \mathrm{tr}\left[\mathcal{E}(\phi)\phi\right]$ represents an average gate infidelity of a CPTP mapping $\mathcal{E}$.

Proof. By straightforward calculation, we obtain

$$\begin{aligned}
r(\Upsilon^{-1}\circ\hat{\Upsilon}_x) &= 1-\int d\phi \mathrm{tr}\left[U^\dagger\hat{U}_x\phi(U^\dagger\hat{U}_x)^\dagger\phi\right] &(9)\\
&= 1-\int d\phi \mathrm{tr}\left[((U^\dagger\hat{U}_x)\otimes(U^\dagger\hat{U}_x)^\dagger)(\phi\otimes\phi)SWAP\right] &(10)\\
&= 1-\frac{1}{d(d+1)}\mathrm{tr}\left[((U^\dagger\hat{U}_x)\otimes(U^\dagger\hat{U}_x)^\dagger)(\mathbb{I}+SWAP)SWAP\right] &(11)\\
&= 1-\frac{1}{d(d+1)}\left(d+\left|\mathrm{tr}\left[U^\dagger\hat{U}_x\right]\right|^2\right) &(12)\\
&= \frac{d}{d+1}\left(1-\frac{1}{d^2}\left|\mathrm{tr}\left[U^\dagger\hat{U}_x\right]\right|^2\right), &(13)
\end{aligned}$$

where we use $\mathrm{tr}\left[(A\otimes B)SWAP\right]=\mathrm{tr}\left[AB\right]$ in the second and fourth equality. This completes the proof. □