

# Adaptively Secure Unitary Designs with Constant Non-Clifford Cost

Lennart Bittel<sup>1,\*</sup> and Lorenzo Leone<sup>1,2,\*</sup>

<sup>1</sup>*Dahlem Center for Complex Quantum Systems, Freie Universität Berlin, 14195 Berlin, Germany*

<sup>2</sup>*Dipartimento di Ingegneria Industriale, Università degli Studi di Salerno,  
Via Giovanni Paolo II, 132, 84084 Fisciano (SA), Italy*



(Received 13 November 2025; revised 23 February 2026; accepted 3 April 2026; published 28 May 2026)

Randomness is a fundamental resource in quantum information, with crucial applications in cryptography, algorithms, and error correction. A central challenge is to construct unitary  $k$  designs that closely approximate Haar-random unitaries while minimizing the costly use of non-Clifford operations. In this Letter, we present a protocol able to generate unitary  $k$  designs on  $n$  qubits, secure against any adversarial quantum measurement, with a system-size-independent number of non-Clifford gates. Our construction applies a  $k$  design only to a subsystem of size  $\Theta(k)$ , independent of  $n$ . This “seed” design is then “diluted” across the entire  $n$ -qubit system by sandwiching it between two random Clifford operators. The resulting ensemble forms an  $\varepsilon$ -approximate unitary  $k$  design on  $n$  qubits. We prove that this construction achieves full quantum security against adaptive adversaries using only  $\tilde{O}(k^2 \log \varepsilon^{-1})$  non-Clifford gates. If one requires security only against polynomial-time adaptive adversaries, the non-Clifford cost decreases to  $\tilde{O}(k + \log^{1+c} \varepsilon^{-1})$ . This is optimal, since we show that at least  $\Omega(k)$  non-Clifford gates are required in this setting. Compared to existing approaches, our method significantly reduces non-Clifford overhead while strengthening security guarantees to adaptive security as well as removing artificial assumptions between  $n$  and  $k$ . These results make high-order unitary designs practically attainable in near-term fault-tolerant quantum architectures.

DOI: [10.1103/PhysRevLett.136.210802](https://doi.org/10.1103/PhysRevLett.136.210802)

Random unitary operations play a central role in modern quantum information science, with applications ranging from cryptography [1,2] and algorithms [3,4] to quantum sensing [5], verification methods [6], and communication protocols [7,8]. They are equally important as conceptual tools: ensembles of random unitaries often provide effective coarse-grained models for the dynamics of highly complex quantum systems. This perspective has revealed deep insights into spectral statistics of many-body Hamiltonians, mechanisms of thermalization [9], holographic dualities [10], the spreading of quantum information [11], and principles of quantum thermodynamics [12]. With such broad relevance, both practical and foundational, a natural question is how randomness arises in quantum theory and how it can be used efficiently.

The most direct construction of quantum randomness is to sample unitary matrices uniformly from the Haar

measure over the unitary group on  $n$  qubits. While mathematically elegant, this approach quickly becomes unfeasible: generating Haar-random unitaries requires circuits of exponential depth in  $n$ , a fact already evident from basic dimensional counting [13]. Fortunately, many applications do not require “full Haar randomness”; it is often sufficient to match only the first few moments, say, up to order  $k$ . A more efficient approach is therefore to use ensembles of unitaries that reproduce the first  $k$  moments of the Haar statistics, known as *unitary  $k$  designs* [14,15]. An even weaker requirement, motivated from complexity theory, is computational pseudorandomness: *pseudorandom unitaries* are ensembles of operators that cannot be distinguished from Haar random unitaries by any *efficient* quantum measurement [16,17]. Both notions have become powerful tools in their own right and have been developed into a vibrant field of research [16,18–21]. Strikingly, very recent advances [19,22] show that unitary  $k$  designs and pseudorandom unitaries can now be implemented in circuits of logarithmic depth in the system size, a dramatic improvement over the exponential-depth constructions required for Haar-random unitaries.

Even though extremely shallow circuits can, in principle, generate ensembles of random unitaries, the true cost in many quantum hardware platforms—particularly early fault-tolerant architectures—depends not only on the *number* of gates but also on the *type* of gates. A central

\*These authors contributed equally to this work.

<sup>†</sup>Contact author: [l.bittel@fu-berlin.de](mailto:l.bittel@fu-berlin.de)

<sup>‡</sup>Contact author: [loleone@unisa.it](mailto:loleone@unisa.it)

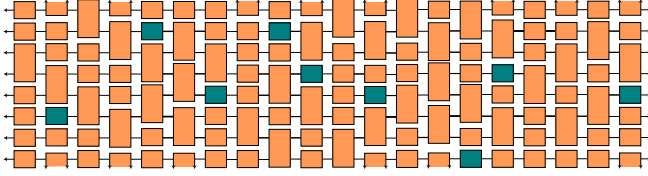


FIG. 1. Clifford circuit (orange boxes) with few interspersed non-Clifford gates (green squares).

limitation arises from the Eastin-Knill theorem [23], which establishes that no quantum error correcting code can implement a universal gate set transversally, i.e., without time or space overhead. As a consequence, fault-tolerant gate-based architectures (based on a single code) are necessarily limited to schemes that require the implementation of at least one non-transversal gate, which then constitutes the most expensive component of the architecture. A widely used approach is to treat Clifford operations as having negligible cost [24] and to supplement them with a non-Clifford operation. The latter, implemented fault-tolerantly via *magic-state injection* [25], extends the Clifford group to a universal gate set. Within this framework, the focus naturally shifts from reducing circuit depth to minimizing the number of non-Clifford gates required to realize unitary  $k$  designs (see Fig. 1), since the dominant cost typically arises from the preparation of *high-fidelity* magic-states. Indeed, despite remarkable progress in recent years [26–30], *magic-state distillation* remains the most resource-intensive component of fault-tolerant quantum computation.

Because of the practical relevance of this problem, substantial effort has been devoted to understanding how to construct random unitaries with minimal *non-Clifford cost*. This line of work has produced a number of important results, summarized in Table I, and reviewed

TABLE I. Summary of the results on unitary design constructions and their non-Clifford cost. Here,  $\times$  indicates that the construction is not fully secure, while  $\checkmark$  denotes full security. In particular,  $\checkmark$  indicates security against all quantum experiments, and  $\hat{\checkmark}$  indicates security against polynomial-time quantum experiments. Quantum-secure and quantum-polynomially-secure designs are defined in Definitions 1 and 2, respectively. The symbol  $[\star]$  marks the results proven in this Letter, while  $[\diamond]$  denotes the non-Clifford cost expressed not in terms of gate count, but in terms of the support size of the diluted  $k$  design in our construction (see Fig. 2 and Theorem 1).

|                                     | Non-Clifford cost                              | Security           |
|-------------------------------------|------------------------------------------------|--------------------|
| Additive $k$ design [32,34]         | $O(k^4 + k \log \epsilon^{-1})$                | $\times$           |
| Relative $k$ designs [32,33]        | $\Theta(nk)$                                   | $\checkmark$       |
| Pseudorandom unitaries [33]         | $\Theta(n)$                                    | $\hat{\checkmark}$ |
| Secure-diluted $k$ design $[\star]$ | $\Theta(k + \log \epsilon^{-1})[\diamond]$     | $\checkmark$       |
| Quantum-secure $k$ design $[\star]$ | $\tilde{O}(k^2 \log \epsilon^{-1})$            | $\checkmark$       |
| Poly-secure $k$ design $[\star]$    | $\tilde{\Theta}(k + \log^{1+c} \epsilon^{-1})$ | $\hat{\checkmark}$ |

in Appendix S.1 [31]. Most notably, a seminal paper [32] established the existence of a scheme capable of generating unitary  $k$  designs with a system-size independent number of non-Clifford gates, namely  $\tilde{O}(k^4)$  [56]. This was a striking result: for moderate values of  $k$ , it implied that random unitary operators could, in principle, be implemented on near-term fault-tolerant quantum platforms with only a constant (system-size independent) resource cost.

However, this excitement was soon tempered. The notion of unitary  $k$  designs used in that result—*additive-error* unitary  $k$  designs—only guarantees indistinguishability from Haar-random unitaries against a restricted class of quantum measurements, namely *nonadaptive* quantum algorithms. Meanwhile, the stronger notion of *relative-error* unitary designs gained popularity and is now widely regarded as the gold standard [19]. Relative-error designs ensure closeness in expectation values under relative error, thereby guaranteeing indistinguishability against the most general (adaptive) quantum measurement. Unfortunately, a later result [33] demonstrated that constructing relative-error  $k$  designs necessarily requires an extensive non-Clifford cost of  $\Omega(nk)$ . This result appeared to rule out the efficient realization of ensembles that are operationally indistinguishable from Haar-random unitaries in fully adaptive quantum experiments. But did it, really?

In this Letter, we introduce a protocol which dilutes a unitary  $k$  design on  $\Theta(k)$  qubits to generate unitary  $k$  designs on  $n$  qubits using only Clifford operations (see Fig. 2). The resulting designs are secure against *any* quantum measurement while requiring, by construction, only a system-size-independent number of non-Clifford gates. In addition, the construction can be implemented in logarithmic depth with the use of extra qubits. This result, which we explain in detail below, has two key implications: (i) *practical*: it enables the implementation of random unitaries on near-term fault-tolerant quantum platforms, since the construction requires *shallow* unitaries with *few* non-Clifford gates; and (ii) *conceptual*: it shows that the notion of relative-error designs is not only unnecessary for

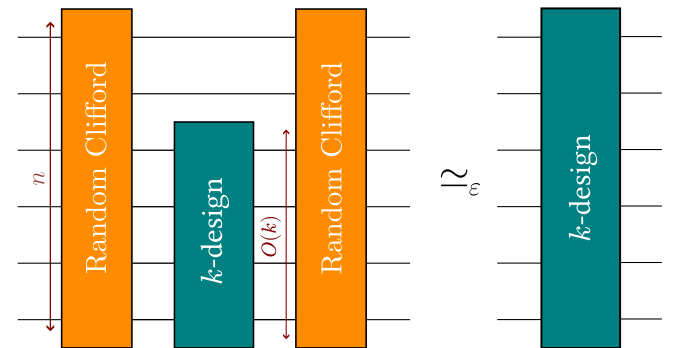


FIG. 2. Our construction: mixing power of Clifford circuits “dilute” the design “seed” into the full  $n$ -qubit space. The figure provides a graphical representation of Theorem 1.

physically motivated applications, but also overly demanding, requiring an extensive number of non-Clifford gates compared to the constant amount sufficient for security against *any* quantum measurement.

*Setup and notation*—We consider a system of  $n$  qubits with Hilbert space  $\mathbb{C}^{2^n}$  and operator basis given by the Pauli group  $\mathbb{P}_n$ . We denote by  $\|\cdot\|_1$  and  $\|\cdot\|_\infty$  the trace norm and the operator norm of operators and by  $\|\cdot\|_\diamond$  the diamond norm for quantum channels. For an ensemble  $\mathcal{E}$  on the unitary group  $\mathcal{U}_n$ ,  $U \sim \mathcal{E}$  indicates that  $U$  is drawn uniformly at random from  $\mathcal{E}$ ; in particular, Haar denotes the Haar (uniform) measure. We will mostly focus on the Clifford group  $\mathcal{C}_n$ , the subgroup of unitaries that maps Pauli operators to Pauli operators.

*Unitary  $k$  designs*—We begin with an overview of the rich landscape of unitary  $k$  designs, unraveling the various concepts and nuances behind them. Let  $\mathcal{E}$  be an ensemble of unitary operators on  $n$  qubits, and define its  $k$ -moment operator as  $\Phi_{\mathcal{E}}(\cdot) := \mathbb{E}_{U \sim \mathcal{E}} U^{\otimes k}(\cdot) U^{\dagger \otimes k}$ . The ensemble  $\mathcal{E}$  forms an *additive error  $\varepsilon$ -approximate unitary  $k$  design* if [35,57]

$$\|\Phi_{\mathcal{E}} - \Phi_{\text{Haar}}\|_\diamond \leq \varepsilon, \quad (1)$$

where  $\|\cdot\|_\diamond$  denotes the diamond norm. Using its variational definition, i.e.,  $\max_\rho \|I \otimes \Phi_{\mathcal{E}}(\rho) - I \otimes \Phi_{\text{Haar}}(\rho)\|_1$ , one can deduce its operational meaning: given a quantum state  $\rho$  with arbitrary additional qubits  $n'$ , an  $\varepsilon$ -approximate  $k$  design in additive error cannot be distinguished from the Haar distribution by any quantum measurement, with resolution greater than  $\varepsilon$ , making  $k$  *parallel queries* to the unitary. This, however, does not cover the most general scenario: the most general quantum measurement may *adaptively* choose subsequent operations based on intermediate measurement outcomes.

The notion of *relative error  $\varepsilon$ -approximate unitary  $k$  designs* is instead defined by the closeness of the quantum channels  $\Phi_{\mathcal{E}}$  and  $\Phi_{\text{Haar}}$  in relative error [35]:

$$\Phi_{\text{Haar}}(1 - \varepsilon) \leq \Phi_{\mathcal{E}} \leq (1 + \varepsilon)\Phi_{\text{Haar}}, \quad (2)$$

where  $\leq$  denotes the ordering of completely positive maps, i.e.  $\Phi \leq \Phi'$  if and only if  $(\Phi' - \Phi)$  is completely positive. From Eq. (2), one deduces the operational meaning: for any positive  $(n + n')$ -qubit matrix  $\tilde{\rho}$ , the expectation value of any operator  $O$  differs by at most  $2\varepsilon$  in relative error when either  $I \otimes \Phi_{\mathcal{E}}$  or  $I \otimes \Phi_{\text{Haar}}$  is applied. Therefore, Eq. (2) implies Eq. (1), but is strictly stronger. In fact, it can be shown [19] that the state produced by any quantum algorithm making  $k$  arbitrary queries to  $U$  can be written as the partial overlap between one unnormalized state obtained by  $k$  parallel queries to  $U$  and another unnormalized state. Since relative error is insensitive to normalization, the outputs of any measurement when  $U$  is sampled from  $\mathcal{E}$  or from Haar differ by at most  $2\varepsilon$ . This establishes

that relative error  $k$  designs are secure against all kind of quantum measurements.

Thus, using relative error designs that require  $\Theta(nk)$  non-Clifford gates [33], the best known non-Clifford cost scaling to guarantee security against general quantum measurements is  $O(nk)$ . In contrast, as anticipated, our construction yields an  $\varepsilon$ -approximate unitary  $k$  design secure against all quantum measurement with a system-size independent number of non-Clifford gates. The key step is to move beyond relative designs and instead focus exclusively on the strictly weaker—yet fully sufficient for all physically motivated purposes—requirement of security against the most general quantum algorithm.

To formalize this, we consider the most general scenario where an adversary has query access to  $k$  copies of an  $n$ -qubit unitary  $U$ . The adversary may operate on  $n + n'$  qubits (including auxiliary workspace) and adaptively choose operations depending on intermediate measurement outcomes.

We capture this idea with the following definition of a *quantum-secure unitary  $k$ -design* [58].

*Definition 1 (Quantum-secure unitary designs)*—Let  $\mathcal{E}$  be an ensemble of  $n$ -qubit unitaries and let Haar denote the Haar measure on  $n$  qubits. Let  $V$  be a collection of  $k$  arbitrary unitary operators acting on  $n + n'$  qubits, and define  $|\Psi_U(V)\rangle := UV_k UV_{k-1} \cdots UV_1 |0\rangle$ .  $\mathcal{E}$  is an  $\varepsilon$ -approximate quantum secure unitary  $k$  design if

$$\sup_{n'} \max_V \left\| \mathbb{E}_{U \sim \mathcal{E}} \Psi_U(V) - \mathbb{E}_{U \sim \text{Haar}} \Psi_U(V) \right\|_1 \leq \varepsilon. \quad (3)$$

Informally, the above norm measures the maximum distinguishing advantage as perceived by any quantum experiment making at most  $k$  queries to a unitary  $U$  sampled either from  $\mathcal{E}$  or from the Haar measure. By construction, quantum-secure designs imply additive error designs and, as explained above, are implied by relative error designs:

$$\text{relative-error} \begin{matrix} \Rightarrow \\ \nLeftarrow \end{matrix} \text{quantum-secure} \begin{matrix} \Rightarrow \\ \nLeftarrow \end{matrix} \text{additive-error}. \quad (4)$$

Pseudorandom unitaries differ conceptually from unitary  $k$  designs [21,36,59]. An ensemble  $\mathcal{E}$  is pseudorandom if no *efficient quantum experiment* can distinguish it, up to polynomially small resolution, from Haar-random unitaries. Here, efficient experiments are those that use at most polynomially many queries  $k$ , with polynomially many auxiliary qubits  $n'$ , and implement unitaries  $V$  and POVMs of polynomial complexity in the total system size  $n$ . This makes pseudorandom unitaries the most practically motivated notion of randomness generation. Yet, known constructions of pseudorandom unitaries are *conditional* on computational assumptions [36].

As in the case of relative-error designs, Ref. [33] shows that generating pseudorandom unitaries requires  $\Omega(n)$  non-Clifford gates. In practice, however, many applications only

demand indistinguishability up to the first  $k$  moments of the Haar distribution. The following definition combines the concepts of unitary  $k$  designs and pseudorandom unitaries into what we refer to as *quantum polynomially secure unitary  $k$  designs*.

**Definition 2** (*Quantum polynomially secure unitary  $k$  designs*)—Let  $V$  be a collection of  $k$  arbitrary unitary operators acting on  $n + n'$  qubits, and define  $|\Psi_U(V)\rangle := UV_k UV_{k-1} \cdots UV_1 |0\rangle$ . Let  $O$  be any operator such that  $\|O\|_\infty \leq 1$ . The ensemble  $\mathcal{E}$  forms a quantum polynomially secure  $\varepsilon$ -approximate unitary  $k$  design if

$$\sup_{n', V, O \in \text{poly}(n)} \left| \mathbb{E}_{U \sim \mathcal{E}} \text{tr}[O \Psi_U(V)] - \mathbb{E}_{U \sim \text{Haar}} \text{tr}[O \Psi_U(V)] \right| \leq \varepsilon, \quad (5)$$

where  $\text{poly}(n)$  denotes the class of objects that can be implemented with resources polynomial in the total number of qubits  $n$ .

Definition 2 is similar in spirit to Definition 1, but it requires security only against polynomial-time quantum measurements that make at most  $k$  queries to  $U$ , making it more relevant for practical applications.

Before concluding this section, we note for the interested reader that the Pauli group ( $k = 1$ ) and the Clifford group ( $k = 2, 3$ ) provide (trivial) examples of quantum-secure designs, as they are also exact unitary designs. The first nontrivial examples of quantum-secure designs are presented in the next section.

*Our construction: Quantum secure designs with constant non-Clifford cost*—The resulting global operation forms an approximate quantum-secure unitary  $k$  design on  $n$  qubits (though not a relative-error design), while the non-Clifford cost remains confined to  $\Theta(k)$  qubits.

Concretely, we consider the ensemble of unitaries  $\mathcal{E}_t := \{C_1 U_t C_2\}$ , where  $C_1, C_2 \sim \mathcal{C}_n$  are uniformly sampled from the Clifford group  $\mathcal{C}_n$  and  $U_t$  from the Haar measure, or from an *exact*  $k$  design, on  $t$  qubits; see Fig. 2. It is worth noting that the Clifford group alone already generates exact unitary 3 designs [60,61], which implies that Clifford unitaries are powerful scramblers of information, the property that is primarily responsible for our construction to work. The next theorem, which constitutes the main result of this Letter, employs state-of-the-art techniques on the *commutant* of the Clifford group [37,38] to show that this ensemble forms an  $\varepsilon$ -approximate quantum-secure unitary  $k$  design.

**Theorem 1**—The ensemble of unitaries  $\mathcal{E}_t$  forms a  $\varepsilon$ -approximate quantum-secure unitary  $k$  design provided that  $t \geq 2k + \log \varepsilon^{-1} + 6$ .

*Proof sketch*—We derive an explicit tight bound on the norm defined in Definition 1. The key step in the proof relies on a functional analysis lemma of independent interest. Specifically, consider a function of the form  $f(x) = \sum_{i=1}^k c_i 2^{-ix}$ . If  $|f(x)| \leq 1$  for every integer

$x \leq k$ , then the coefficients satisfy  $|c_i| \leq 2^{ki}$ . The full proof is deferred to Appendix S.3 1 [31]. ■

The above result shows that the *dilution* of a  $k$  design across  $O(k)$  qubits through the application of random Clifford operators suffice to spread the property around the macroscopic system of  $n$  qubits. Moreover, we show that  $\Omega(k)$  is the minimal support required to dilute an exact  $k$  design into a quantum-secure design on  $n$  qubits, thereby establishing the optimality of Theorem 1.

An immediate question is how our construction translates into bounds on the number of non-Clifford gates required for quantum-secure unitary  $k$  designs. Since the only step involving non-Clifford operations is the injected  $k$  design on  $\Theta(k)$  qubits, the resulting non-Clifford cost is *automatically system-size independent*, establishing the central result of this Letter. The exact number of non-Clifford gates depends on the best available constructions of unitary  $k$  designs. As an example, in the following corollary we apply the result of Ref. [19] and use a hybrid argument, replacing the exact  $k$  design on  $\Theta(k)$  qubits with a relative-error approximate  $k$  design [see Eq. (2)], to derive an explicit bound on the non-Clifford gate count *sufficient* for our construction.

**Corollary 1**—Quantum-secure unitary  $k$  designs can be constructed using  $O(k^2 \log^2 \varepsilon^{-1})$  many non-Clifford gates [62].

The proof has to be found in Appendix S.3 3 [31]. A natural question for the curious reader is how many non-Clifford gates are *necessary* to construct quantum-secure unitary designs, and whether Corollary 1 achieves optimality in this regard. The following theorem establishes a lower bound on the non-Clifford cost required for unitary designs. Specifically, we present a *computationally efficient* quantum algorithm which, given  $k$  parallel queries to a unitary sampled either from an ensemble  $\mathcal{E}$  or from the Haar measure, can distinguish the two unless the unitaries in  $\mathcal{E}$  contain at least  $\Omega(k)$  non-Clifford gates.

**Theorem 2**—Let  $\mathcal{E}$  be an ensemble of unitaries containing at most  $t < n$  non-Clifford gates. If  $12t + 10 \leq k$ , then

$$\|\Phi_{\mathcal{E}_t} - \Phi_{\text{Haar}}\|_\diamond \geq \frac{1}{8}. \quad (6)$$

*Proof sketch*—The distinguishing algorithm proceeds as follows. We construct  $k$  copies of  $|\psi\rangle = U|0\rangle$ . We demonstrate that  $O(k)$  measurements on  $|\psi\rangle$  are able to decide whether: Case A)  $|\psi\rangle$  contains at least one nontrivial Pauli operator  $P$  such that  $P|\psi\rangle = |\psi\rangle$  or, Case B), there is no such Pauli operator. We then notice that for  $U$  being Haar random there is no such Pauli operator with probability  $\sim 1$ , while for  $U$  constructed by  $O(k)$  Clifford gates that are at least  $2^{n-O(k)}$  [63]. The full proof is deferred to Appendix S.3 2 [31]. ■

Although Theorem 1 is *optimal* in terms of the support of the diluted  $k$  design, Theorem 2 shows that the minimal

number of non-Clifford gates required to construct a quantum-secure design is  $\Omega(k)$ . This suggests that the non-Clifford gate count in Corollary 1 may be loose by a quadratic factor. However, we note that the lower bound is derived from the existence of an *efficient* quantum algorithm that can distinguish the two ensembles, whereas Corollary 1 guarantees security even against unbounded measurements. We accordingly defined quantum polynomially secure  $\varepsilon$ -approximate unitary  $k$  designs (Definition 2) as those that cannot be distinguished from Haar-random unitaries by any  $k$ -query computationally efficient quantum algorithm with resolution greater than  $\varepsilon$ . The following corollary establishes, up to computational assumptions detailed in Appendix S.3.3 [31], the optimality of our construction for polynomially secure designs.

**Corollary 2**—Let  $k = \omega(\log n)$ ,  $\varepsilon = o(n)$ . For any  $c > 0$ , quantum polynomially secure unitary  $k$  design can be constructed using  $\tilde{O}((k + \log^{1+c} \varepsilon^{-1}))$  non-Clifford gates. Combined with Theorem 2, this construction is optimal.

This result establishes that constructing unitaries that reproduce the first  $k$  moments of Haar-random unitaries for any efficiently implementable quantum measurement requires—and suffices with—only  $\tilde{\Theta}(k)$  non-Clifford gates, probing the fundamental limits for unitary designs with minimal non-Clifford resources. As detailed in Appendix S.3.3 [31], the key idea in proving Corollary 2 is to replace the exact  $k$  design in our construction with a pseudorandom unitary, which functions as a quantum polynomially secure design. However, under standard computational assumptions [19,36], such a pseudorandom unitary is secure against polynomial-time quantum adversaries only when implemented on  $\omega(\log n)$  qubits. This requirement explains the assumption on  $k$  in Corollary 2.

As a final remark before concluding this section, we note that our construction is nontrivial only when  $k \leq n$ . However, this is precisely the regime of interest for constructing designs with a system-size-independent number of non-Clifford gates. Indeed, by Theorem 2, any unitary  $k$  design with  $k \geq n$  necessarily requires  $\Omega(n)$  non-Clifford gates. In this case, one can instead employ constructions of random unitaries in shallow depth [19,39] to minimize the use of non-Clifford gates, achieving  $O(nk)$  gates for quantum-secure  $k$  designs and  $O(n)$  gates for quantum polynomially secure designs for any  $k = O(\text{polyn})$  [33].

**Low depth and few non-Clifford gates**—Our construction realizes a quantum-unitary  $k$  design with very few non-Clifford gates, representing an important step toward the practical implementation of random unitaries on fault-tolerant platforms. Reducing the number of non-Clifford gates is essential because their logical implementation requires either deep physical circuits or complex error-detection schemes. Without such a reduction, the shallow constructions of Ref. [19] remain shallow only at the logical level—the depth at the physical level increases significantly once non-Clifford operations are made fault

tolerant. That said, besides limiting non-Clifford gates, it is also desirable to keep the overall depth shallow; otherwise, the total number of physical operations still becomes too large. Our construction can achieve both goals: building on previous results [40,41], we can generate the two random Clifford circuits in logarithmic depth while keeping the number of non-Clifford gates low, as formalized in the next corollary and explicitly proved in Appendix S.3.3 [31].

**Corollary 3**—Quantum-secure unitary  $k$  designs on  $n$  qubits can be realized with  $O(k^2 \log \varepsilon^{-1})$  non-Clifford gates,  $O(n^2 + k^2)$  extra qubits, and an overall depth  $O(\log n + \log k)$  under all-to-all connectivity.

As a final remark, the use of extra qubits and all-to-all connectivity seem to be necessary to achieve a shallow-depth version of our construction, as Ref. [42] shows a linear lower bound for implementing a Clifford  $k$  designs with  $k \geq 4$ .

**Outlook**—Our results show that higher-order unitary designs can be implemented with a system-size independent number of non-Clifford gates and in low depth, while remaining indistinguishable from exact designs by any quantum protocol. Unlike previous works, which either only guarantee security against nonadaptive adversaries [32] or rely on relative-error designs requiring  $\Omega(nk)$  non-Clifford gates [33], our construction achieves *full quantum security* with a system-size independent and at a substantially reduced non-Clifford cost:  $\tilde{O}(k^2)$  non-Clifford gates, compared to the  $\tilde{O}(k^4)$  scaling of the original proposal [32]. Moreover, according to Theorem 2 and Corollary 2, our construction also employs an *optimal* number  $\tilde{\Theta}(k)$  of non-Clifford gates if one is only concerned with security against polynomial-time quantum adversaries.

We therefore deduce two important implications: on the conceptual level, the results provide advances in the theory of unitary  $k$  designs by showing that quantum-secure designs embody the minimal and most natural notion of security whereas relative-error designs impose unnecessarily strong requirements. On the practical level, our construction is well suited for near-term implementation on early fault-tolerant devices for two main reasons: (1) all non-Clifford resources are concentrated on a constant, system-size-independent set of qubits, and (2) the overall circuit can be implemented in shallow depth.

Finally, we remark that to dilute an exact  $k$  design into a quantum-secure unitary  $k$  design via Clifford operations, it is both necessary and sufficient to place the  $k$  design on  $\tilde{\Theta}(k)$  qubits. This makes our construction automatically implementable with a system-size-independent number of non-Clifford gates. While this is optimal in terms of support, there is still room for improvement in the state-of-the-art constructions of (polynomially) quantum-secure designs on  $\tilde{\Theta}(k)$  qubits. Combining our approach with the latest unitary designs constructions could further tighten the results of Corollaries 1 and 2, improving the exact count and scaling of non-Clifford gates, as well as extending their regime of validity.

**Acknowledgments**—The authors thank Salvatore F.E. Oliviero for creating the images of the paper, and Armanda Quintavalle for important discussions about Clifford + T fault-tolerant architectures. This work has been supported by the DFG (CRC 183, FOR 2724), by the BMBF (Hybrid++, QuSol), the BMWK (EniQmA), the Munich Quantum Valley (K-8), the QuantERA (HQCC), the Alexander-von-Humboldt Foundation, Berlin Quantum and the European Research Council (ERC AdG DebuQC). This work has also been funded by the DFG under Germany’s Excellence Strategy—The Berlin Mathematics Research Center MATH+ (EXC-2046/1, Project ID: 390685689). L. L. acknowledges funding from the Italian Ministry of University and Research, PRIN PNRR 2022, project “Harnessing Topological Phases for Quantum Technologies,” Code P202253RLY, CUP D53D23016250001, and PNRR-NQSTI project “ECoN: End-to-end Long-Distance Entanglement in Quantum Networks,” CUP J13C22000680006.

**Data availability**—No data were created or analyzed in this study.

- [1] A. Ambainis and A. Smith, in *Proceedings of RANDOM’04*, Lecture Notes in Computer Science Vol. 3122 (Springer, New York, 2004), pp. 249–260.
- [2] W. Kretschmer, in TQC 2021 (Leibniz International Proceedings in Informatics (LIPIcs), Latvia, 2021), Vol. **197**, pp. 2:1–2:20.
- [3] P. Sen, in *21st Annual IEEE Conference on Computational Complexity (CCC’06)* (IEEE, Prague, 2006), pp. 14–287.
- [4] F. G. S. L. Brandão and M. Horodecki, *Quantum Inf. Comput.* **13**, 901 (2013).
- [5] R. Kueng, H. Zhu, and D. Gross, [arXiv:1609.08595](https://arxiv.org/abs/1609.08595).
- [6] J. Eisert, D. Hangleiter, N. Walk, I. Roth, D. Markham, R. Parekh, U. Chabaud, and E. Kashefi, *Nat. Rev. Phys.* **2**, 382 (2020).
- [7] I. Devetak and A. Winter, *Phys. Rev. Lett.* **93**, 080501 (2004).
- [8] B. Groisman, S. Popescu, and A. Winter, *Phys. Rev. A* **72**, 032317 (2005).
- [9] S. Popescu, A. J. Short, and A. Winter, *Nat. Phys.* **2**, 754 (2006).
- [10] P. Hayden and J. Preskill, *J. High Energy Phys.* **09** (2007) 120.
- [11] Y. Sekino and L. Susskind, *J. High Energy Phys.* **10** (2008) 065.
- [12] A. Munson, N. B. T. Kothakonda, J. Haferkamp, N. Yunger Halpern, J. Eisert, and P. Faist, *PRX Quantum* **6**, 010346 (2025).
- [13] E. Knill, Approximation by quantum circuits, [arXiv:quant-ph/9508006](https://arxiv.org/abs/quant-ph/9508006).
- [14] J. Emerson, Y. S. Weinstein, M. Saraceno, S. Lloyd, and D. G. Cory, *Science* **302**, 2098 (2003).
- [15] D. Gross, K. Audenaert, and J. Eisert, *J. Math. Phys. (N.Y.)* **48**, 052104 (2007).
- [16] Z. Ji, Y.-K. Liu, and F. Song, Pseudorandom quantum states, in *Advances in Cryptology—CRYPTO 2018* (Springer International Publishing, New York, 2018), pp. 126–152.
- [17] T. Haug, K. Bharti, and D. E. Koh, *Quantum* **9**, 1759 (2025).
- [18] T. Metger, A. Poremba, M. Sinha, and H. Yuen, [arXiv:2402.14803](https://arxiv.org/abs/2402.14803).
- [19] T. Schuster, J. Haferkamp, and H.-Y. Huang, [arXiv:2407.07754](https://arxiv.org/abs/2407.07754).
- [20] F. Ma and H.-Y. Huang, [arXiv:2410.10116](https://arxiv.org/abs/2410.10116).
- [21] T. Haug, K. Bharti, and D. E. Koh, [arXiv:2306.11677](https://arxiv.org/abs/2306.11677).
- [22] N. LaRacuenta and F. Leditzky, [arXiv:2407.07876](https://arxiv.org/abs/2407.07876).
- [23] B. Eastin and E. Knill, *Phys. Rev. Lett.* **102**, 110502 (2009).
- [24] D. Gottesman, Stabilizer codes and quantum error correction, Ph.D. thesis, California Institute of Technology, 1997.
- [25] S. Bravyi and J. Haah, *Phys. Rev. A* **86**, 052329 (2012).
- [26] A. Krishna and J.-P. Tillich, *Phys. Rev. Lett.* **123**, 070507 (2019).
- [27] K. Fang and Z.-W. Liu, [arXiv:2410.14547](https://arxiv.org/abs/2410.14547).
- [28] Q. T. Nguyen, in *Proceedings of the 57th Annual ACM Symposium on Theory of Computing* (Association for Computing Machinery, New York, 2025), pp. 697–706.
- [29] L. Golowich and V. Guruswami, in *Proceedings of the 57th Annual ACM Symposium on Theory of Computing* (2025), pp. 707–717.
- [30] A. Wills, M.-H. Hsieh, and H. Yamasaki, *Nat. Phys.* **21**, 1842 (2025).
- [31] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevX.10.010344> for details of the proofs, which includes Refs. [9,14,19,32–55].
- [32] J. Haferkamp, F. Montealegre-Mora, M. Heinrich, J. Eisert, D. Gross, and I. Roth, *Commun. Math. Phys.* **397**, 995 (2023).
- [33] L. Leone, S. F. E. Oliviero, A. Hamma, J. Eisert, and L. Bittel, [arXiv:2505.10110](https://arxiv.org/abs/2505.10110).
- [34] Y. Zhang, S. Vijay, Y. Gu, and Y. Bao, *PRX Quantum* **7**, 010344 (2026).
- [35] F. G. S. L. Brandão, A. W. Harrow, and M. Horodecki, *Commun. Math. Phys.* **346**, 397 (2016).
- [36] F. Ma and H.-Y. Huang, [arXiv:2410.10116](https://arxiv.org/abs/2410.10116).
- [37] L. Bittel, J. Eisert, L. Leone, A. A. Mele, and S. F. E. Oliviero, [arXiv:2504.12263](https://arxiv.org/abs/2504.12263).
- [38] L. Bittel and L. Leone, *Quantum* **10**, 2069 (2026).
- [39] L. Cui, T. Schuster, F. Brandao, and H.-Y. Huang, [arXiv:2507.06216](https://arxiv.org/abs/2507.06216).
- [40] C. Moore and M. Nilsson, [arXiv:quant-ph/9808027](https://arxiv.org/abs/quant-ph/9808027).
- [41] J. Jiang, X. Sun, S.-H. Teng, B. Wu, K. Wu, and J. Zhang, [arXiv:1907.05087](https://arxiv.org/abs/1907.05087).
- [42] L. Grevink, J. Haferkamp, M. Heinrich, J. Helsen, M. Hinsche, T. Schuster, and Z. Zimborás, [arXiv:2506.23925](https://arxiv.org/abs/2506.23925).
- [43] J. Emerson, R. Alicki, and K. Życzkowski, *J. Opt. B* **7**, S347 (2005).
- [44] A. W. Harrow and R. A. Low, *Commun. Math. Phys.* **291**, 257 (2009).
- [45] J. Haferkamp, [arXiv:2203.16571](https://arxiv.org/abs/2203.16571).
- [46] D. Gross, K. Audenaert, and J. Eisert, *J. Math. Phys. (N.Y.)* **48**, 052104 (2007).
- [47] B. Magni, A. Christopoulos, A. De Luca, and X. Turkeshi, *Phys. Rev. X* **15**, 031071 (2025).
- [48] D. Weingarten, *J. Math. Phys. (N.Y.)* **19**, 999 (1978).
- [49] A. F. Mello, A. Santini, G. Lami, J. De Nardis, and M. Collura, *Phys. Rev. Lett.* **134**, 150403 (2025).
- [50] R. A. Low, *Phys. Rev. A* **80**, 052314 (2009).

- 
- [51] D. E. Knuth *et al.*, *The Art of Computer Programming* (Pearson Education India, Upper Saddle River, 1997), Vol. 1, 3/E.
- [52] S. Aaronson and D. Gottesman, *Phys. Rev. A* **70**, 052328 (2004).
- [53] A. Gu, L. Leone, K. Goodenough, and S. Khatri, [arXiv: 2502.09483](#).
- [54] L. Leone, S. F. E. Oliviero, S. Lloyd, and A. Hamma, *Phys. Rev. A* **109**, 022429 (2024).
- [55] S. F. E. Oliviero, L. Leone, S. Lloyd, and A. Hamma, *Phys. Rev. Lett.* **132**, 080402 (2024).
- [56] The notation  $\tilde{O}$  denotes  $\tilde{O}(f(n)) = O(f(n)\text{poly log } f(n))$ , i.e. the usual big- $O$  (or Landau) notation up to poly-logarithmic factors.
- [57] A. W. Harrow and R. A. Low, *Commun. Math. Phys.* **291**, 302 (2009).
- [58] A very recent work [39] introduced the notion of quantum-secure designs (referred to as *measurable-error designs*), with the primary goal of minimizing circuit depth. However, with respect to gate count, their results offer no asymptotic improvement over relative-error designs, as both require  $\Theta(nk)$  gates. In sharp contrast, our results show that when it comes to non-Clifford cost, relative-error designs impose an unnecessary burden: they require  $\Omega(n)$  non-Clifford gates, whereas the practically motivated quantum-secure designs of Definition 1 need only  $O(1)$ .
- [59] Z. Ji, Y.-K. Liu, and F. Song, in *Advances in Cryptology—CRYPTO 2018*, Lecture Notes in Computer Science, edited by H. Shacham and A. Boldyreva (Springer International Publishing, Cham, 2018), pp. 126–152.
- [60] Z. Webb, *Quantum Inf. Comput.* **16**, 1379 (2016).
- [61] H. Zhu, *Phys. Rev. A* **96**, 062336 (2017).
- [62] The dependence on the resolution error  $\varepsilon$  can be improved to  $\tilde{O}(k^2 \log \varepsilon^{-1})$  by using  $\tilde{O}(k^2)$  extra qubits, using the construction of Ref. [39]; see Supplemental Material [31] for details.
- [63] L. Leone, S. F. E. Oliviero, and A. Hamma, *Quantum* **8**, 1361 (2024).