Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

# The unbearable hardness of deciding about magic

**Jens Eisert**

jenseisert@gmail.com

Freie Universität Berlin https://orcid.org/0000-0003-3033-1292

**Lorenzo Leone**

Universita degli Studi di Salerno

**Salvatore Oliviero**

Scuola Normale Superiore and Istituto Nanoscienze https://orcid.org/0000-0003-3569-085X

**Article**

**Keywords:**

**Posted Date:** March 31st, 2026

**DOI:** https://doi.org/10.21203/rs.3.rs-8988905/v1

**License:** This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

**Additional Declarations:** There is **NO** Competing Interest.

# The unbearable hardness of deciding about magic

Lorenzo Leone,[1,2,*] Jens Eisert,[2,†] and Salvatore F. E. Oliviero[2,‡]

[1]*Dipartimento di Ingegneria Industriale, Università degli Studi di Salerno, Via Giovanni Paolo II, 132, 84084 Fisciano (SA), Italy*

[2]*Dahlem Center for Complex Quantum Systems, Freie Universität Berlin, 14195 Berlin, Germany*

Identifying the boundary between classical and quantum computation is a central challenge in quantum information. In multi-qubit systems, entanglement and magic are the key resources underlying genuinely quantum behaviour. While entanglement is well understood, magic—essential for universal quantum computation—remains relatively poorly characterised. Here we show that determining membership in the stabilizer polytope, which defines the free states of magic-state resource theory, requires super-exponential time $\exp(n^2)$ in the number of qubits $n$, even approximately. We reduce the problem to solving a 3-SAT instance on $n^2$ variables and, by invoking the exponential time hypothesis, the result follows. As a consequence, both quantifying and certifying magic are fundamentally intractable: any magic monotone for general states must be super-exponentially hard to compute, and deciding whether an operator is a valid magic witness is equally difficult. As a corollary, we establish the robustness of magic as computationally optimal among monotones. This barrier extends even to classically simulable regimes: deciding whether a state lies in the convex hull of states generated by a logarithmic number of non-Clifford gates is also super-exponentially hard. Together, these results reveal intrinsic computational limits on assessing classical simulability, distilling pathological magic states, and ultimately probing and exploiting magic as a quantum resource.

## I. INTRODUCTION

Quantum computers promise the efficient solution of some computational problems that are out of reach for efficient classical computers [1, 2]. For a long time, this idea was mainly seen as an inspiring theoretical possibility. In recent years, however—largely driven by major efforts in the quantum industry—it has started to become an experimental reality. The quantum computers available today are still limited in size and remain noisy, with devices operating at more than 1000 qubits. Even so, machines of this scale, with two-qubit error rates close to a tenth of a percent, would have been hard to imagine not long ago. This rapid progress has created an exciting situation for the field [3–5]. The devices we have today, and those expected in the near future, are commonly referred to as *noisy intermediate-scale quantum* (NISQ) devices [6]. These devices may enable useful quantum applications, and this question is currently the focus of intense research. At the same time, many quantum protocols implemented on such hardware can still be simulated efficiently on classical computers using modern methods. This process, often called "dequantization" [7–10], removes the expected quantum advantage. One of the central questions in the field today is what can truly be achieved with these devices, and to what extent they still operate in a regime that is effectively classical. At a more conceptual level, the key issue is to understand where the boundary lies between the quantum regime—where one may expect a genuine quantum advantage [3–5]—and the classical regime, which still allows for efficient classical description. Identifying this boundary is one of the main challenges of the field.

A particularly clean and rigorous way to formalise this boundary is provided by quantum resource theories [11–13]. In the context of multi-qubit quantum computation, two key resources emerge as necessary for a state to lie on the quantum side of the elusive classical–quantum boundary: entanglement and magic [13, 14]. However, while the resource theory of entanglement [13] can be regarded as being extremely well understood, the same cannot quite be said for the resource theory of magic, largely because it has only recently attracted serious attention [14–16]. Magic-state resource theory formalises the observation that Clifford operations and stabilizer states, although genuinely quantum, are efficiently classically simulable [17, 18], using the rigorous language of quantum resource theory. Within this framework, states confined to the stabilizer polytope—the region of Hilbert space reachable by classical mixtures of stabilizer states—constitute the *free states*, while quantum operations that map the stabilizer polytope onto itself define the *free operations* of the theory [14, 15]. Exiting the stabilizer polytope requires *magic*, the ingredient that pushes quantum states outside the classically simulable regime [18–24]. Magic is, in a sense, the conceptual feature that makes quantum systems "quantum." In more technological terms, it is the ingredient that uplifts schemes based on transversal Clifford circuits to fully universal fault-tolerant quantum computing schemes [24].

In this work, we bring this line of research in a striking new direction: we show that deciding whether a quantum state contains magic necessarily requires algorithms whose runtime grows super-exponentially with the number of qubits $n$. More precisely, we formulate a mathematically well-defined decision problem —namely, a weak membership problem for the stabilizer polytope— and reduce it to a 3-SAT instance on $n^2$ variables. Invoking the Exponential Time Hypothesis,

* loleone@unisa.it
† jense@zedat.fu-berlin.de
‡ s.oliviero@fu-berlin.de

this reduction implies that solving the problem requires $\exp(n^2)$ time. While the technical argument is involved, the upshot is that we can relate membership in the stabilizer polytope to testing properties of *graph states* [25, 26].

This result opens an exciting new direction for this line of research. It is no surprise that identifying the subtle boundary of the stabilizer polytope is difficult: it is a provably computationally hard task that adheres to the rigour of theoretical computer science. While this insight is primarily of conceptual value, we will later show that it also has important practical implications for magic-state resource theory, including quantifying and detecting magic, classical simulation of noisy quantum circuits, as well as resource distillation in modern quantum error correction.

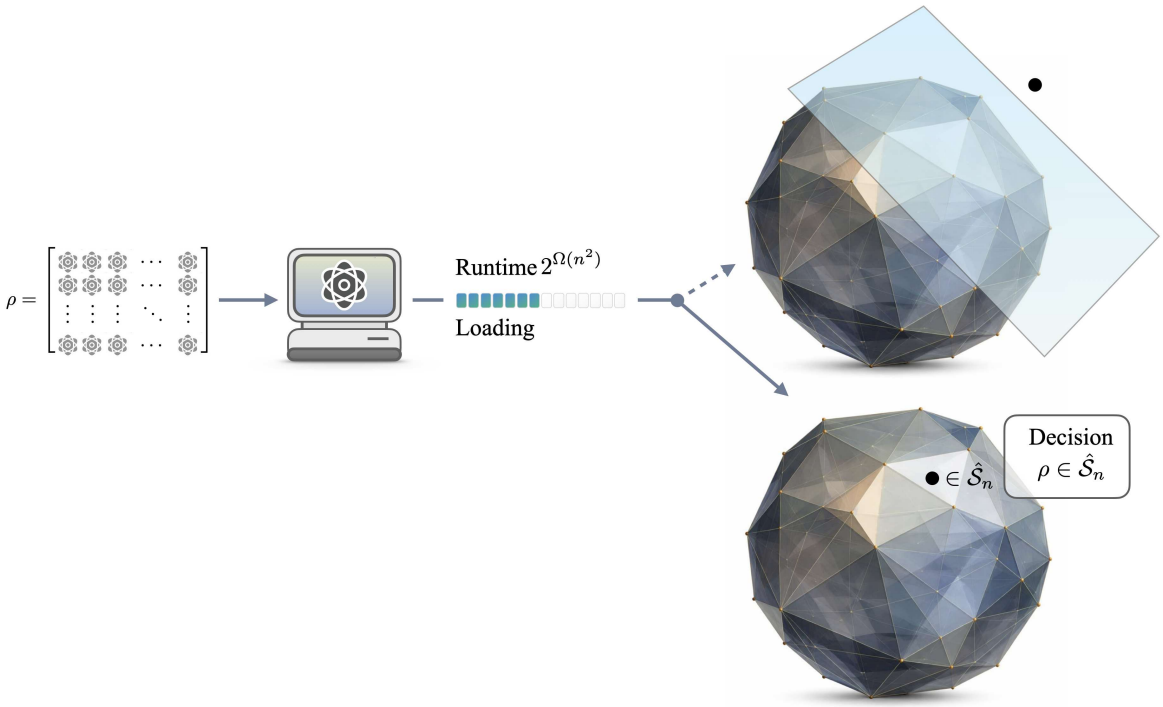



FIG. 1. In this work, we show that any algorithm aimed at determining whether a $d \times d$ density matrix on $n$ qubits, $\rho$, belongs to the stabilizer polytope must necessarily run in super-exponential time. Consequently, any magic monotone also requires super-exponential time to compute. Concretely, we prove that the membership problem in the stabilizer polytope, formulated as a decision problem, lies in the complexity class $\mathsf{QP}^2$. Also depicted in the upper picture is a schematic representation of a witness as a hyperplane separating a quantum state from the stabilizer polytope. At a higher level, it demonstrates that delineating the "classical" setting—where efficient classical simulation is possible, inside the polytope—from the "quantum" setting—where one can hope for computational quantum advantages—is itself a computationally difficult task.

## II. RESULTS

We now state the main conclusion of this work and highlight its implications: deciding whether a quantum state contains magic—whether by quantifying it, detecting it, or determining its membership in the classically simulable region defined by the stabilizer formalism—is computationally hard.

To set the stage in more formal terms, we take the Hilbert space dimension $d = 2^n$ as the input size of our problem, since we are given a $d \times d$ density matrix $\rho$ describing a (possibly mixed) quantum state. We denote by $\mathsf{P}$ the class of decision problems that can be solved in time polynomial in $d$ [27]. For instance, deciding the purity of $\rho$ clearly lies in $\mathsf{P}$. We denote by $\mathsf{NP}$ the class of problems for which there exists a polynomial-time deterministic Turing machine that can verify $\mathsf{YES}$ instance of the decision problem. For our purposes, it is useful to introduce *quasi-polynomial* complexity classes. For $k \in \mathbb{R}_+$, a decision problem with input size $d$ is said to lie in $\mathsf{QP}^k$ if there exists a deterministic Turing machine solving it in time $\exp(\log^k d)$. Note that problems in $\mathsf{QP}^k$ with $k > 1$ have a strictly *super-exponential* runtime with respect to the number of qubits $n$; nevertheless, we will refer to them as *quasi-polynomial* because the scaling is considered relative to the Hilbert space dimension $d$. These classes can be separated from one another under one of the most widely believed assumptions in computational complexity: the *exponential time hypothesis* (ETH) [28] which we will assume throughout this work. ETH asserts that $\mathsf{NP}$-complete problems cannot be solved in subexponential time. While the hierarchy $\mathsf{QP}^k$ for different values of $k$ offers a fine-grained view of quasi-polynomial behaviour, the union

$\bigcup_{k\in\mathbb{N}} \mathsf{QP}^k$ captures exactly all problems that can be solved in strictly subexponential time with respect to the input size $d$.

### A. Quantifying magic

In close analogy with entanglement theory [29], there are two complementary approaches to accessing magic: quantification and witnessing. Both aim to characterise the boundary of the stabilizer polytope. Quantifying magic is the more demanding task: it seeks to distinguish magic states exactly from stabilizer states—i.e., to identify the precise boundary of the stabilizer polytope in Hilbert space—and to assign a meaningful measure of how much magic a state contains. This is formalised through magic monotones, scalar functions that vanish exactly on states inside the stabilizer polytope [21–23]. The scarcity of computable monotones is the main reason for the slow progress in magic-state resource theory. Although many candidates have been proposed, they are notoriously difficult to evaluate [15], typically requiring minimisation over all pure stabilizer states—an inherently super-polynomial task in the Hilbert space dimension. Stabilizer entropies [16, 30–32] partially alleviate this by avoiding such minimisations and being computable from expectation values of Hermitian operators in polynomial time. However, they behave reliably only for pure states; extending them to mixed states again requires an optimisation procedure. These difficulties suggest intrinsic computational hardness in magic-state resource theory and naturally raise the question of whether any efficiently computable magic monotone for mixed states can exist at all—or, more modestly, whether the boundary of the stabilizer polytope can be determined exactly and efficiently.

The main result of this work answers both questions negatively: we show that there is no efficient method to precisely characterise the boundary of the stabilizer polytope, and consequently, any magic monotone must necessarily require super-polynomial time to compute. To formalise these questions, we define the weak membership problem for the stabilizer polytope $\hat{\mathcal{S}}$: given a state $\rho$, decide whether it is a free state, i.e., $\rho \in \hat{\mathcal{S}}$, or whether it is $\varepsilon$-far from any free state. In the following informal theorem, we show that this membership problem lies in $\mathsf{QP}^2$.

**Theorem 1** (Membership in the stabilizer polytope is hard. Informal of Theorem 6 and Lemma 11)**.** *Deciding whether a state $\rho$ lies in the stabilizer polytope $\hat{\mathcal{S}}$, or is $\varepsilon$-far from every state in $\hat{\mathcal{S}}$, belongs to the complexity class $\mathsf{QP}^2$ for any $\varepsilon = 1/\mathsf{poly}(d)$.*

*Proof sketch.* The proof has two parts: showing that the decision problem lies in $\mathsf{QP}^2$, and that it does not belong to $\mathsf{QP}^{2-\eta}$ for any $\eta > 0$. Membership in $\mathsf{QP}^2$ follows from computing the known magic monotone, the *robustness of magic*, which can be evaluated in time $\mathsf{exp}(\log^2 d)$ using the argument in Ref. [21], rigorously refined in Lemma 11. The second, more technical part encodes an instance of the polytope membership problem as a 3-$\mathsf{SAT}$ instance in $O(\log^2 d)$ variables. By the ETH, this cannot be solved in time $\mathsf{exp}(\log^{2-\eta} d)$ for any $\eta > 0$. □

Therefore, even the simple task of classifying a state as free or resourceful in magic-state resource theory requires super-polynomial resources, highlighting the inherent hardness of the theory. Since the golden rule of a magic monotone is to *faithfully* identify the set of free states—i.e., those inside the stabilizer polytope $\hat{\mathcal{S}}$—it follows as a simple corollary that the computational difficulty of every magic monotone valid for mixed states is not accidental, but an intrinsic feature of magic-state resource theory.

**Corollary 1** (No efficient magic monotone)**.** *The computation of any magic monotone $\mathcal{M}$ requires time* $\mathsf{exp}(\log^2 d)$.

At this point, some remarks are in order. Many magic monotones have been proposed so far, and as anticipated, some of them require minimisation over the set of pure stabilizer states. Since the set of stabilizer states contains $\mathsf{exp}(\log^2 d)$ elements, such an optimisation—even when convex—can only be performed by an algorithm scaling as $\mathsf{exp}(\log^2 d)$. From Corollary 1, we see that these measures are indeed *optimal*, as one cannot provably do better within magic-state resource theory. One example is the *robustness of magic* [21], defined as the minimum negativity of the coefficients when expressing $\rho$ as a linear combination of pure stabilizer states.

The computational difficulty of magic monotones established by Corollary 1 applies to monotones designed to detect magic in *arbitrary mixed states*. Indeed, stabilizer entropies [16] are efficiently computable, but they faithfully identify free states only when these states are pure. In Ref. [30], it has been shown that stabilizer entropies can be extended to general mixed states via the standard convex roof construction. While this extension involves an optimisation over all convex decompositions of a mixed state $\rho$ into pure states—a task even more computationally demanding than optimising over all stabilizer states—restricting the rank of $\rho$ can substantially reduce this complexity. Specifically, Ref. [30] observes that computing the extended stabilizer entropy requires an $\mathsf{exp}(\log d\,\mathrm{rank}(\rho)^2)$-time algorithm, placing it within $\mathsf{P}$ for states with constant rank. This shows that, although magic-state resource theory is super-polynomially hard for general mixed states, restricting to low-rank states—common, for example, in matrix product states in many-body physics—allows stabilizer entropies to be practically useful for exploring moderately large systems, far beyond the $n \simeq 5$ limit implied by the full super-polynomial hardness of $\mathsf{QP}^2$.

Finally, it is instructive to compare our results with other resource theories, highlighting that this intrinsic computational hardness is a distinctive feature of magic-state resource theory. In the resource theory of coherence, the structure is simple enough that monotones are efficiently computable [33]. The same holds true for Gaussian continuous-variable entanglement, where even all tangent hyperplanes to the convex set of covariance matrices of bosonic Gaussian separable states can be classified [34]. Similarly, despite the non-convex nature of fermionic non-Gaussianity, monotones remain computable. By contrast, entanglement theory exhibits a much higher computational barrier, with complexity scaling exponentially in the dimension of the Hilbert space [35–37].

### B. Witnessing magic

Faced with this intrinsic barrier, it is natural to consider a weaker—but far more accessible—approach: witnessing magic. Rather than fully quantifying the resource, one may simply ask whether a given state possesses magic at all. A *witness* is a Hermitian operator that, when yielding a positive expectation value, certifies the presence of magic in a quantum state. More precisely, a witness $W$ is such that $\mathrm{tr}(W\rho) \leq 0$ for all $\sigma \in \hat{\mathcal{S}}$. Therefore, if a state $\rho$ gives $\mathrm{tr}(W\rho) > 0$, it is guaranteed to be a magic state. The converse does not hold: a negative expectation value does not imply that $\rho$ is a stabilizer state. Although this limitation may seem unsatisfactory, two points put witnessing in proper context. First, witnesses correspond to directly measurable observables, providing a conceptually and experimentally transparent way to test for magic. Second, witnesses are *complete*: for every non-stabilizer state $\rho$, there exists at least one witness $W_\rho$ that detects it.

By definition, evaluating a witness on a given state can be done in polynomial time. The main drawback, however, is that a single witness cannot detect all magic states. If it could, it would be faithful on $\hat{\mathcal{S}}$, which is impossible according to Theorem 1. Consequently, effectively probing the boundary of the stabilizer polytope and detecting magic across different states requires multiple witnesses. This naturally leads to the question: how can we systematically construct witnesses in magic-state resource theory? In recent years, several approaches have been proposed, ranging from witnesses derived from stabilizer entropies [38] to criteria [39] reminiscent of the celebrated PPT test in entanglement theory [40, 41]. Yet, as we show here, deciding whether a given observable is a valid witness is itself as hard as faithfully identifying the stabilizer polytope—once again requiring super-polynomial computational resources.

**Theorem 2** (Finding witnesses is hard. Informal of Theorem 7 and Lemma 12)**.** *Given a Hermitian operator $W$, deciding whether $\mathrm{tr}(W\sigma) \leq 0$ for all $\sigma \in \hat{\mathcal{S}}$ (i.e., $W$ is a valid witness) or there exists $\tau \in \hat{\mathcal{S}}$ such that $\mathrm{tr}(W\tau) > 0$ (i.e., $W$ it is not a witness) belongs to the complexity class* $\mathsf{QP}^2$.

Thus, magic-state resource theory presents a no-free-lunch scenario: it is computationally hard both to faithfully identify the stabilizer polytope and to determine whether a given operator can serve as a witness. As a result, capturing magic in states not detected by the relatively few known witnesses remains inherently difficult.

### C. The classical-quantum boundary

To come to the point and picking up the core theme of the introduction, we now discuss the implications of our results for delineating the boundary between classical and quantum computation. The classical-quantum boundary is notoriously elusive. Even if a state contains a large amount of magic and cannot be simulated by the stabilizer formalism, it might still be efficiently simulable using other classical methods, such as tensor networks or covariance matrix techniques [42–46]. What is well-defined, however, is the classical-quantum boundary *as seen by the stabilizer formalism*. This boundary separates states that can be efficiently simulated by stabilizer methods from those that cannot.

Although Theorem 1 establishes a fundamental bottleneck in tracing the boundary between free and resourceful states, we can go one step further and consider the classical-quantum boundary from the perspective of magic: given a quantum state, can we determine whether it is classically simulable using the stabilizer formalism? This question goes beyond the simple dichotomy between free and resourceful states, as even states containing a small amount of magic can sometimes be classically simulated [47–49]. For example, consider a state prepared by a Clifford circuit doped with a finite number $t$ of local non-Clifford gates [50–53]. It is well known that such circuits can be simulated in time scaling exponentially in $t$ [47–49], which implies polynomial-time simulation in the number of qubits when $t = O(\log n)$. Hence, the classical-quantum boundary, from the perspective of magic-state resource theory, is broader than the stabilizer polytope; it includes *$t$-doped stabilizer states* and their convex combinations. Therefore, we ask: how hard is it to identify the boundary of the polytope formed by $t$-doped stabilizer states hereby denoted by $\hat{\mathcal{S}}_t$? Identifying this boundary is a crucial step toward understanding the complete picture: states inside the polytope of $t$-doped states are classically simulable, while states outside—and even very close to the boundary—are, in general, hard to simulate. This sharp transition in computational power makes the polytope of $t$-doped states the natural classical–quantum boundary within the stabilizer framework.

**Theorem 3** (Tracing the classical-quantum boundary is hard. Informal of Theorem 8)**.** *For any $\varepsilon = 1/\mathsf{poly}(d)$, the problem of deciding whether a state $\rho$ lies in $\hat{\mathcal{S}}_t$ or is $\varepsilon$-far from every state in $\hat{\mathcal{S}}_t$ satisfies:*

- *It belongs to the complexity class $\mathsf{QP}^2$ for any $t < \log n$.*
- *For any $t = O(\log n)$ lies at least in $\mathsf{QP}^2$ and at most in $\mathsf{QP} = \bigcup_k \mathsf{QP}^k$.*

Unsurprisingly, at this stage, the above theorem shows that this task turns out to be super-polynomially hard. Interestingly, however, as long as $t$ scales logarithmically with the number of qubits, the problem never becomes exponentially hard in the Hilbert space dimension, but instead remains quasi-polynomial.

### D. Summary of conceptual results

Altogether, we reveal an "unbearable hardness" of magic: not only its quantification and reliable certification are intractable in principle, but even deciding membership in the set of classically simulable states is computationally hard in magic-state resource theory. Moreover, this intractability manifests in a way that is significant in practice. Indeed, from the very beginning, nobody expected these problems to be polynomial-time tractable in the number of qubits; exponential complexity already seemed like the most we could reasonably hope for. Yet even this hope would have been enough: exponential-time tools still allow meaningful large-scale simulations. For example, suppose an optimised classical solver can execute about $2^{40} \sim 10^{12}$ operations. A $2^n$-time algorithm can then handle instances of size $n \approx 25$, large enough to capture genuine asymptotic behaviour. Super-exponential scaling completely changes this picture. Under the same computational budget, a $2^{n^2}$-time algorithm becomes infeasible already around $n \approx 5$, restricting us to toy models dominated by finite-size effects. In this sense, replacing $2^n$ by $2^{n^2}$ does not merely worsen constant factors; it fundamentally changes what we can hope to analyse in practice.

To conclude, we extend the discussion to highlight the practical consequences of our results. In particular, we show that for noisy quantum circuits, determining whether noise renders them classically tractable is even harder—super-exponential compared to their brute-force simulation. Similarly, in the context of magic-state distillation, we reveal a striking structural feature of magic-state resource theory: there exist pathological magic states whose distillation, although possible in principle, requires prohibitively super-exponential time, making it useless in practice.

## III. IMPLICATIONS OF OUR RESULTS

While our results are rigorous and framed in the language of theoretical computer science, revealing the intrinsic computational complexity of magic-state resource theory, in this section, we discuss some of their implications in more practical, application-orientated scenarios. In particular, despite its foundational goal of distinguishing classical from quantum physics, magic-state resource theory has two main practical purposes. First, states or quantum operations containing little magic can often be *simulated efficiently by classical means* [18, 19, 48, 54]. This is highly valuable, as it provides quantum theorists with an additional powerful simulation tool: alongside tensor network methods, it enables the study of processes implemented on early quantum devices and supports ongoing experimental efforts. The second major practical application of magic-state resource theory lies in *quantum error correction* and fault-tolerant quantum computing, specifically in the task of *magic-state distillation* [21]. This important subroutine, combined with transversal Clifford gates, allows for efficient fault-tolerant quantum computation [55, 56]. While our results are primarily conceptual, they nevertheless have important implications for both of these practical directions.

### A. Classical simulation of noisy quantum circuits

Quantum computers hold the promise of outperforming classical machines on certain tasks, but this promise remains largely theoretical until fully fault-tolerant hardware is available. In the current noisy intermediate-scale regime, a central goal has been to demonstrate a *quantum advantage*—running a quantum circuit that performs a task beyond the reach of the best classical computers. A leading proposal for such demonstrations has been *random circuit sampling* [57], in which a randomly chosen quantum circuit is executed on a device, and one samples from its output distribution. However, the inherent noise of current devices makes these processes partially classically simulable, and improvements in classical simulation techniques—including tensor network methods and massive GPU implementations—have repeatedly challenged the claimed advantage in an interesting way, at times matching or even surpassing benchmarks [58, 59]. As a result, experimental demonstrations of quantum advantage remain in constant tension with the evolving capabilities of classical simulation.

In this context, a natural systematic question arises: which noisy quantum circuits are efficiently simulable by stabilizer-based classical methods? A noisy quantum circuit is, in general, a quantum channel $\mathcal{E}$, that is, a completely positive trace-preserving linear map on quantum states. Based on our discussion of classical simulability via the stabilizer formalism in Section II C, we say that a channel $\mathcal{E}$ is classically simulable if it never generates states outside the tractable region defined by the $t$-doped stabilizer polytope $\hat{\mathcal{S}}_t$ for $t = O(\log n)$. Concretely, the noisy quantum circuit $\mathcal{E}$ is simulable if $\mathcal{E}(\sigma) \in \hat{\mathcal{S}}_t$ for every input stabilizer state $\sigma$. Equivalently, $\mathcal{E}$ injects at most $t$ non-Clifford operations into any input state, and we call such a map a *completely $t$-doped stabilizer preserving channel*.

Building on the hardness of stabilizer and $t$-doped stabilizer membership problems established in Theorems 1 and 3, we obtain the following result regarding the classical simulation of noisy quantum circuits.

**Corollary 2** (Deciding classical simulation of noisy circuits is hard. Informal of Corollaries 4 and 5)**.** *Given a quantum channel $\mathcal{E}$, deciding whether $\mathcal{E}$ is completely $t$-doped stabilizer preserving or not does not belong to the complexity class $\mathsf{QP}^{2-\eta}$ for any $\eta > 0$. In particular, this problem is much harder than brute-force simulating the process $\mathcal{E}$ on a classical device.*

The content of the above corollary implies, somewhat ironically, that determining whether a noisy quantum circuit is classically simulable in the stabilizer sense is significantly harder than the brute-force simulation of the circuit itself. This highlights the intrinsic computational hardness of deciding magic resources in quantum states and channels, even when considering weaker notions of simulability.

### B. Super-polynomial time magic-state distillation

Magic-state distillation is a key component of universal fault-tolerant quantum computation. In most physical architectures, Clifford gates can be implemented reliably and, crucially, transversally. On their own, however, Clifford operations are not computationally universal. Universality can be achieved by supplementing them with suitable non-Clifford gates. But then the hen-and-egg problem arises: how to realise those non-Clifford gates in the first place. *Magic-state distillation* [20, 21, 55, 56, 60–64] addresses this challenge by starting from multiple copies of noisy non-stabilizer states and processing them into fewer, higher-fidelity non-stabilizer states. These improved resource states are then injected into Clifford circuits via gate teleportation, enabling non-Clifford operations and, ultimately, universal quantum computation. Operationally, one typically has access to only a limited number of noisy magic states on a small number of qubits and aims to convert them into high-fidelity resources suitable for implementing non-Clifford gates. A central issue in this setting is the computational cost of carrying out such distillation procedures.

While Theorem 1 does not directly provide constructive guidance for designing magic-state distillation protocols, it does offer important insights. First, distillation schemes based on the canonical $|T\rangle \propto |0\rangle + e^{i\pi/4}|1\rangle$ resource state are not unique, and identifying alternative resource states that may lead to improved efficiency is a meaningful research direction. Indeed, resource states involving one-, two-, and three-qubit systems, among others, have been explored [20, 21, 56, 62, 63]. However, a first consequence of Theorem 1 is that the very search for new magic resource states is hindered by the intrinsic difficulty of deciding whether a given state is a magic state at all; no efficient optimisation procedure can accomplish this task.

A further implication of Theorem 1 is the existence—presumably in rather pathological cases, though this is hard to assess—of magic states for which distillation, while not ruled out in principle, cannot be achieved by any protocol running in less than super-polynomial time in the Hilbert space dimension, as the next corollary rigorously establishes.

**Corollary 3** (Super-polynomial time magic-state distillation. Informal of Corollary 6)**.** *There exists a (possibly mixed) non-stabilizer state $\rho$ on $n$ qubits such that any stabilizer-based distillation protocol that can be found and executed within polynomial time in $d$ produces at most a super-polynomially small $\mathsf{exp}(-\log^2 d)$ fraction of high-fidelity magic states.*

## IV. DISCUSSION AND OPEN QUESTIONS

In this work, we have shown that the geometry of stabilizer states conceals an unexpectedly severe computational barrier: both the recognition of classicality and the quantification of genuinely quantum resources require $\mathsf{exp}(n^2)$ time. This result fundamentally reframes the practical and conceptual status of the stabilizer polytope, of magic resource theory, and of the broader task of delineating classical from quantum computational power. If one identifies the class of systems that can be efficiently simulated by Pauli propagation methods—and that therefore admit no quantum advantage—with a "classical world," and its complement with a "quantum world," then our results imply that the very act of separating these two regimes is itself computationally intractable. In this sense, the classical–quantum boundary is not merely subtle or poorly understood, but provably hard to determine. This adds a striking and somewhat ironic twist to an already challenging foundational problem.

Many questions remain open. Can these hardness results be circumvented for restricted families of states or channels? Are there physically natural approximations that avoid the $\exp(n^2)$ barrier? Can the promise gap discussed in the present proofs be made larger? Similar exponentially small promise gaps are known from studies of the computational complexity of the separability problem [36], but is there room for improvement? Do analogous phenomena arise in other resource theories? And, ultimately, what concrete lessons do these insights offer for the search for new quantum algorithms that achieve genuine advantages over classical computation? It is hoped that the present work invites the solution of some of these open problems.

## V. ACKNOWLEDGMENTS

We thank Scott Aaronson for pointing out a typographic error to us. This work resulted from presentations and discussions at the Second International Workshop on Many-Body Quantum Magic, held by the InQubator for Quantum Simulation (IQuS) at the University of Washington, 18-29 August 2025. The workshop has been supported by the U.S. Department of Energy, Office of Science, Office of Nuclear Physics Award Number DOE (NP) Award DE-SC0020970 via the program on Quantum Horizons: QIS Research and Innovation for Nuclear Science. 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. This work has been supported by the BMFTR (DAQC, MuniQC-Atoms, QuSol, Hybrid++, PasQuops), Clusters of Excellence (ML4Q, MATH+), the Munich Quantum Valley, Berlin Quantum, the Quantum Flagship (Millenion, Pasquans2), the DFG (CRC 183, SPP 2541), the European Research Council (DebuQC), PraktiQOM, and the Alexander-von-Humboldt Foundation.

[1] P. W. Shor, Proc. 50th Ann. Symp. Found. Comp. Sc. , 124 (1994).
[2] A. Montanaro, npj Quant. Inf. **2**, 15023 (2016).
[3] J. Eisert and J. Preskill, (2025), arXiv:2510.19928.
[4] Z. Zimboras *et al.*, (2024), arXiv:2501.05694.
[5] R. Babbush, R. King, S. Boixo, W. Huggins, T. Khattar, G. H. Low, J. R. McClean, T. O'Brien, and N. C. Rubin, (2025), arXiv:2511.09124.
[6] J. Preskill, (2018), arxiv:1801.00862.
[7] A. Angrisani, A. A. Mele, M. S. Rudolph, M. Cerezo, and Z. Holmes, (2025), arXiv:2501.13101.
[8] A. A. Mele, A. Angrisani, S. Ghosh, S. Khatri, J. Eisert, D. S. Franca, and Y. Quek, (2024), arXiv:2403.13927.
[9] M. S. Rudolph, T. Jones, Y. Teng, A. Angrisani, and Z. Holmes, (2025), arXiv:2505.21606.
[10] Y. Zhou, E. M. Stoudenmire, and X. Waintal, Phys. Rev. X **10**, 041038 (2020).
[11] E. Chitambar and G. Gour, Rev. Mod. Phys. **91**, 1806.06107 (2019).
[12] G. Gour, *Resources of the quantum world* (Cambridge University Press, Cambridge, USA, 2024) p. 2402.05474.
[13] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, Rev. Mod. Phys. **81**, 865 (2009).
[14] V. Veitch, S. A. H. Mousavian, D. Gottesman, and J. Emerson, New J. Phys. **16**, 013009 (2014).
[15] Z.-W. Liu and A. Winter, PRX Quantum **3**, 020333 (2022).
[16] L. Leone, S. F. E. Oliviero, and A. Hamma, Phys. Rev. Lett. **128**, 050402 (2022).
[17] D. Gottesman, (1998), arXiv:quant-ph/9807006.
[18] S. Aaronson and D. Gottesman, Phys. Rev. A **70**, 052328 (2004).
[19] V. Veitch, C. Ferrie, D. Gross, and J. Emerson, New J. Phys. **14**, 113011 (2012).
[20] E. T. Campbell and M. Howard, Phys. Rev. A **95**, 022316 (2017).
[21] M. Howard and E. T. Campbell, Phys. Rev. Lett. **118**, 090501 (2017).
[22] B. Regula, J. Phys. A **51**, 045303 (2018).
[23] J. R. Seddon and E. T. Campbell, Proc. Roy. Soc. A **475**, 20190251 (2019).
[24] A. Mari and J. Eisert, Phys. Rev. Lett. **109**, 230503 (2012).
[25] M. Hein, J. Eisert, and H. J. Briegel, Phys. Rev. A **69**, 062311 (2004).
[26] M. Hein, W. Dür, J. Eisert, R. Raussendorf, M. Van den Nest, and H.-J. Briegel, in *Proceedings of the International School of Physics 'Enrico Fermi' on 'Quantum Computers, Algorithms and Chaos'* (IOS Press, Amsterdam, 2005) pp. 115 – 218.
[27] S. Arora and B. Barak, *Computational complexity: A modern approach*, 1st ed. (Cambridge University Press, Cambridge, USA, 2009).
[28] R. Impagliazzo and R. Paturi, J. Comp. Syst. Sc. **62**, 367 (2001).
[29] O. Gühne and G. Toth, Phys. Rep. **474**, 1 (2009).
[30] L. Leone and L. Bittel, Phys. Rev. A **110**, L040403 (2024).
[31] L. Bittel and L. Leone, (2025), arXiv:2507.22883.
[32] S. F. E. Oliviero, L. Leone, A. Hamma, and S. Lloyd, npj Quant. Inf. **8**, 148 (2022).
[33] T. Baumgratz, M. Cramer, and M. B. Plenio, Phys. Rev. Lett. **113**, 140401 (2014).

[34] P. Hyllus and J. Eisert, New J. Phys. **8**, 51 (2006).
[35] L. Gurvits, (2003), arXiv:quant-ph/0303055.
[36] L. M. Ioannou, Quant. Inf. Comp. **7**, 335 (2007).
[37] S. Gharibian, (2009), arXiv:0810.4507.
[38] T. Haug and P. S. Tarabunga, (2025), arXiv:2504.18098.
[39] Z. Liu, T. Haug, Q. Ye, Z.-W. Liu, and I. Roth, (2025), arXiv:2512.16777.
[40] A. Peres, Phys. Rev. Lett. **77**, 1413 (1996).
[41] M. Horodecki, P. Horodecki, and R. Horodecki, Phys. Lett. A **223**, 1 (1996).
[42] R. Orús, Ann. Phys. **349**, 117 (2014).
[43] J. I. Cirac, D. Pérez-García, N. Schuch, and F. Verstraete, Rev. Mod. Phys. **93**, 045003 (2021).
[44] P. Corboz, R. Orus, B. Bauer, and G. Vidal, Phys. Rev. B **81**, 165104 (2010).
[45] T. Barthel, C. Pineda, and J. Eisert, Phys. Rev. A **80**, 042333 (2009).
[46] J. Eisert, M. Cramer, and M. B. Plenio, Rev. Mod. Phys. **82**, 277 (2010).
[47] S. Aaronson and D. Gottesman, Phys. Rev. A **70**, 052328 (2004).
[48] S. Bravyi and D. Gosset, Phys. Rev. Lett. **116**, 250501 (2016).
[49] S. Bravyi, D. Browne, P. Calpin, E. Campbell, D. Gosset, and M. Howard, Quantum **3**, 181 (2019).
[50] M. Beverland, E. Campbell, M. Howard, and V. Kliuchnikov, Quant. Sc. Tech. **5**, 035009 (2020).
[51] L. Leone, S. F. E. Oliviero, and A. Hamma, Quantum **8**, 1361 (2024).
[52] A. Gu, S. F. E. Oliviero, and L. Leone, Phys. Rev. A **110**, 062427 (2024).
[53] A. Gu, S. F. E. Oliviero, and L. Leone, (2024), arXiv:2403.19610.
[54] X. Wang, M. M. Wilde, and Y. Su, New J. Phys. **21**, 103002 (2019).
[55] S. Bravyi and A. Kitaev, Phys. Rev. A **71**, 022316 (2005).
[56] A. Paetznick and B. W. Reichardt, Phys. Rev. Lett. **111**, 090505 (2013).
[57] A. Morvan *et al.*, Nature **634**, 328 (2024).
[58] X.-H. Zhao, H.-S. Zhong, F. Pan, Z.-H. Chen, R. Fu, Z. Su, X. Xie, C. Zhao, P. Zhang, W. Ouyang, C.-Y. Lu, J.-W. Pan, and M.-C. Chen, (2024), arXiv:2406.18889.
[59] R. Fu *et al.*, (2024), arXiv:2407.00769.
[60] E. T. Campbell, H. Anwar, and D. E. Browne, Phys. Rev. X **2**, 041021 (2012).
[61] D. Litinski, Quantum **3**, 205 (2019).
[62] A. J. Landahl and C. Cesare, (2013), arXiv:1302.3240.
[63] M. B. Hastings and J. Haah, Phys. Rev. Lett. **120**, 050504 (2018).
[64] A. Wills, M.-H. Hsieh, and H. Yamasaki, (2024), arXiv:2408.07764.
[65] M. Grötschel, *Geometric algorithms and combinatorial optimization*, Algorithms and Combinatorics Ser. (Springer, Berlin, 1988).
[66] Y.-K. Liu, (2007), arXiv:0712.3041.
[67] R. Vershynin, *High-dimensional probability: An introduction with applications in data science* (Cambridge University Press, Cambridge, USA, 2018).
[68] F. A. Mele, A. A. Mele, L. Bittel, J. Eisert, V. Giovannetti, L. Lami, L. Leone, and S. F. E. Oliviero, Nature Phys. **21**, 2002 (2025).
[69] S. Bravyi and A. Kitaev, Phys. Rev. A **71**, 022316 (2005).
[70] S. F. E. Oliviero, L. Leone, and A. Hamma, Phys. Rev. A **106**, 2205.02247 (2022), arXiv:2205.02247 [quant-ph].
[71] T. Haug and L. Piroli, Phys. Rev. B **107**, 035148 (2023).
[72] S. Arunachalam and A. Dutt, in *Proc. 57th Ann. ACM Symp. Th. Comp.*, STOC '25 (Association for Computing Machinery, New York, NY, USA, 2025) pp. 1234–1241.
[73] Z. Bao, P. van Dordrecht, and J. Helsen, in *Proceedings of the 57th Ann. ACM Symp. Th. Comp.*, STOC '25 (Association for Computing Machinery, New York, NY, USA, 2025) pp. 1254–1262.
[74] A. Wills, M.-H. Hsieh, and H. Yamasaki, Nature Phys. **21**, 1842 (2025).
[75] S. Rodriguez *et al.*, Nature **645**, 620 (2025).
[76] A. Heimendahl, M. Heinrich, and D. Gross, J. Math. Phys. **63**, 112201 (2022).
[77] A. Gu, S. F. Oliviero, and L. Leone, PRX Quantum **6**, 020324 (2025).
[78] L. Leone, S. F. E. Oliviero, S. Lloyd, and A. Hamma, Physical Review A **109** (2024), 10.1103/physreva.109.022429.
[79] I. Bengtsson and K. Zyczkowski, *Geometry of quantum states: An introduction to quantum entanglement* (Cambridge University Press, Cambridge, USA, 2006).
[80] M. M. Wilde, (2013), 10.1017/CBO9781139525343.
[81] Y. O. Nakagawa and Y. Lee, (2025), arXiv:2508.14546.
[82] M. Beverland, E. Campbell, M. Howard, and V. Kliuchnikov, Quantum Science and Technology **5**, 035009 (2020).

# Supplemental Material

## CONTENTS


S.1. Preliminaries 9
1. Notation and basic notions 9
2. Basics of complexity theory: 3-SAT, quasi-polynomial classes and ETH 10
3. Convex sets and related computational problems 11
4. Preliminaries on $\varepsilon$-nets 13
5. Magic-state resource theory 14
6. Summary of results on $t$-doped stabilizer states 16

S.2. Computational problems in magic-state resource theory 17
1. Weak membership for the stabilizer polytope 20
2. Weak witness detection 22
3. Weak membership in the convex hull of $t$-doped stabilizer states 27

S.3. Additional results and applications 33
1. Completely $t$-doped stabilizer preserving channels and their characterisation 34
2. Channel classification computational problem 36
3. Existence of super-polynomial magic-state distillation 37


## Appendix S.1: Preliminaries

This section is devoted to notation and preliminaries. In Appendix S.1 1 we introduce basic notation and fundamental concepts from quantum information. In Appendix S.1 2 we review standard notions from complexity theory. In Appendix S.1 4 we recall preliminary results on $\varepsilon$-nets. In Appendix S.1 5 we introduce the resource theory of magic states.

### 1. Notation and basic notions

We fix the notation and recall basic notions from quantum information theory that will be used throughout the paper.

We denote the Hilbert space of $n$ qubits by $\mathcal{H}_n$, with dimension $d = 2^n$. The space of linear operators acting on $\mathcal{H}_n$ is denoted by $\mathcal{B}(\mathcal{H}_n)$, and the set of density operators on $\mathcal{H}_n$ by $\mathcal{D}(\mathcal{H}_n) \subset \mathcal{B}(\mathcal{H}_n)$. Many of the quantities considered in this work are naturally expressed in terms of operator norms. In particular, we will make extensive use of Schatten norms, which we now recall.

**Definition 1** (Schatten $p$-norms)**.** *Let $A \in \mathcal{B}(\mathcal{H}_n)$. The Schatten $p$-norm of $A$ is defined as*

$$\|A\|_p \coloneqq \left[\mathrm{tr}(|A|^p)\right]^{1/p}, \tag{S1.1}$$

*where $|A| \coloneqq \sqrt{A^\dagger A}$ and $p \geq 1$. In particular, with $p = 2$ we have the Frobenius norm $\|A\|_2 \coloneqq \sqrt{\mathrm{tr}(A^\dagger A)}$*

The Hilbert–Schmidt inner product on $\mathcal{B}(\mathcal{H}_n)$ is given by $\langle A, B\rangle \coloneqq \mathrm{tr}\big(A^\dagger B\big)$. Hölder's inequality, which generalises the Cauchy–Schwarz inequality, states that

$$|\,\mathrm{tr}\big(A^\dagger B\big)| \leq \|A\|_p \, \|B\|_q, \tag{S1.2}$$

for $p, q \in [1, \infty]$ such that $p^{-1} + q^{-1} = 1$. As a consequence, the Schatten $p$-norm can be shown to admit the variational characterisation

$$\|A\|_p = \max_{\|B\|_q \leq 1} \mathrm{tr}\big(A^\dagger B\big), \tag{S1.3}$$

where again $p^{-1} + q^{-1} = 1$.

### 2. Basics of complexity theory: 3-SAT, quasi-polynomial classes and ETH

This section reviews the basic notion of complexity theory, with an emphasis on the 3-SAT problem. For a comprehensive treatment of complexity theory, we refer the reader to standard references such as Ref. [27]. Throughout this section, let $D$ denote the input size of the computational problems under consideration. For the remainder of the manuscript, we restrict attention to the decision problems defined below.

**Definition 2** (Decision problem)**.** *A decision problem is a computational problem defined by a set $L \subseteq \{0,1\}^D$ (a formal language) such that for every input string $x \in \{0,1\}^D$ the answer is either* YES *or* NO*. The task of a decision problem is to determine whether a given input $x \in L$. A Turing machine that always halts and outputs* YES *exactly when $x \in L$ and* NO *otherwise is said to solve the decision problem.*

We define two complexity classes of primary interest in the remainder of the manuscript.

**Definition 3** (P and NP complexity classes)**.** *The class* P *consists of all decision problems that can be decided in polynomial time by a deterministic Turing machine. The class* NP *consists of all decision problems for which* YES *instances can be verified in polynomial time by a deterministic Turing machine.*

The 3-SAT problem is a fundamental computational problem in complexity theory. It belongs to the class of problems known as NP-complete, which implies that it is both in the complexity class NP and is at least as hard as the hardest problems in NP. In 3-SAT, we are given a Boolean formula consisting of a *conjunction* (AND) of clauses. Each clause is a *disjunction* (OR) of exactly three literals. A literal can be either a variable or its negation. For example, a clause may take the form $(x_1 \vee \neg x_2 \vee x_3)$, where $x_i$ represents a Boolean variable and $\neg$ denotes negation. The goal of the 3-SAT problem is to determine if there exists an assignment of truth values (true or false) to the variables such that the entire Boolean formula evaluates to true. This assignment is often referred to as a satisfying assignment. Given a 3-SAT formula $\phi$ with $n$ variables and $m$ clauses, the question is whether there exists an assignment of truth values $x_1, x_2, \ldots, x_n$ such that $\phi(x_1, x_2, \ldots, x_n)$ evaluates to true. Subsequently, we provide a rigorous definition of the problem.

**Problem 1** (3-SAT)**.** *Problem:*

- *Input: A list of indices $\vec{i} \in [D]^{3,n_c}$ and a negation table $s \in \{\cdot, \neg\}^{3,n_c}$.*
- *Output: Determine, if a Boolean variable $x \in \{0,1\}^D$ exists which satisfies all clauses.*

$$\forall c \in [n_c]: \quad s_{1,c}x_{i_1(c)} \vee s_{2,c}x_{i_2(c)} \vee s_{3,c}x_{i_3(c)}. \tag{S1.4}$$

Here, $[D]^{3,n_c}$ means that we take $n_c$ subsets of 3 elements belonging to $[D]$. To establish the NP-hardness of a problem, a commonly employed approach involves demonstrating a polynomial-time reduction from 3-SAT to the specific problem at hand. This process entails the transformation of 3-SAT instances into instances of the target problem in a manner ensuring that a solution to the transformed instance directly implies a solution to the original 3-SAT instance. The key insight lies in recognising that efficient resolution of the target problem would also imply an efficient resolution of 3-SAT, which is known to be NP-hard. To show the legitimacy of the reduction, one must prove that it operates in polynomial time and satisfies the conditions of "completeness" and "soundness". "Completeness" asserts that if a solution exists for 3-SAT, then a solution also exists for the transformed version of the problem of interest. "Soundness" states that if a solution does not exist for 3-SAT, then there is no solution for the transformed version of the problem of interest. Thus, the NP-hardness of 3-SAT serves as a reference point for proving the computational intractability of other problems, including the ones currently under consideration in our work.

We can give an analogous formulation of the 3-SAT problem as a ground-state problem, which may be more familiar to physicists. Consider spin (Boolean) variables $x_\alpha \in \{0,1\}$, and a sign vector $s$ with components $s_\alpha \in \{0,1\}$, for $\alpha \in \{1, \ldots, D\}$, and define the Hamiltonian

$$H_s(x) = \sum_{\alpha<\beta<\gamma} (x_\alpha \oplus s_\alpha)(x_\beta \oplus s_\beta)(x_\gamma \oplus s_\gamma). \tag{S1.5}$$

where $\oplus$ denotes the addition modulo 2 such that $x_\alpha \oplus 1 = \neg x_\alpha$. The goal is to determine whether there exists an assignment $\bar{x}$ of the variables $x_\alpha$ such that $H(\bar{x}) = 0$, or whether $H(x) \geq 1$ for all possible assignments $x$.

Besides the canonical complexity classes of P and NP, we define the intermediate regime of quasi-polynomial complexity classes.

**Definition 4** (Quasipolynomial time and $\mathsf{QP}^k$)**.** *A decision problem is said to be solvable in* quasipolynomial time *if there exists a deterministic Turing machine that decides it in time*

$$2^{(\log D)^{O(1)}}, \tag{S1.6}$$

*where $D$ is the size of the input. The corresponding complexity class is*

$$\mathsf{QP} := \bigcup_{k\geq 1} \mathsf{TIME}\big(2^{(\log D)^k}\big), \tag{S1.7}$$

*where $\mathsf{TIME}(f(D))$ denotes the class of problems solvable by a deterministic Turing machine in $O(f(D))$ time. For a fixed number $\mathbb{R} \ni k \geq 1$, define*

$$\mathsf{QP}^k := \mathsf{TIME}\big(2^{O((\log D)^k)}\big). \tag{S1.8}$$

*In particular, $\mathsf{QP}^1 = \mathsf{P}$ up to polynomial factors in $D$, and $\mathsf{QP}^2$ corresponds to problems solvable in time $2^{O((\log D)^2)}$. Thus, we have a* hierarchy of quasipolynomial time classes*.*

$$\mathsf{QP}^1 \subseteq \mathsf{QP}^2 \subseteq \mathsf{QP}^3 \subseteq \cdots \subseteq \mathsf{QP} \tag{S1.9}$$

Let us now introduce one of the most widely accepted assumptions in complexity theory, namely the *exponential time hypothesis* (ETH). We will assume ETH throughout this work in order to establish the computational hardness of problems arising in the resource theory of magic states.

**Definition 5** (Exponential time hypothesis (ETH))**.** *The ETH posits that there exists a constant $c > 0$ such that 3-$\mathsf{SAT}$ on $D$ variables cannot be solved in time $2^{cD}$.*

Under the ETH assumption, the quasi-polynomial complexity classes are strictly separated from one another. Indeed, consider a 3-$\mathsf{SAT}$ instance on $\log^k D$ variables for $k \geq 2$. Since it can be solved in time $2^{\log^k D}$, it belongs to $\mathsf{QP}^k$. However, under ETH, it cannot be solved in time $2^{c\log^k D}$ for any constant $c > 0$, and therefore it cannot belong to $\mathsf{QP}^{k-\eta}$ for any $\eta > 0$. This assumption will be used to prove the main result of this work.

### 3. Convex sets and related computational problems

In this section, we introduce the computational problems related to convex sets of $\mathbb{R}^D$ for $D \in \mathbb{N}$.

A subset $K \subseteq \mathbb{R}^D$ is called convex if, for any two points $\mathbf{x}, \mathbf{y} \in K$, the entire straight line segment $\{\lambda\mathbf{x} + (1-\lambda)\mathbf{y} : \lambda \in [0,1]\}$ lies in $K$; the convex hull of a set $S \subseteq \mathbb{R}^D$ is the smallest convex set containing $S$.

A central computational question is the *membership* problem: given a point $\mathbf{y} \in \mathbb{R}^D$ and a description of $K$, decide whether $\mathbf{y} \in K$. Because exact membership may be numerically delicate or computationally hard, one often considers the weak membership problem $\mathrm{WMEM}_\varepsilon(K)$, where a tolerance parameter $\varepsilon \geq 0$ is given, and the task is to decide whether $\mathbf{y}$ lies within a distance $\varepsilon$ of $K$ or is at least $\varepsilon$-far from the interior of $K$. This relaxed formulation affords robustness to small perturbations or approximate descriptions and plays a fundamental role in convex-body algorithms and the complexity analysis of convex-set membership [36, 65].

The weak membership framework is particularly useful in high-dimensional settings or when $K$ is given implicitly (e.g., by a family of constraints or an oracle): it permits stable, approximate decisions even under numerical or representational imprecision. In many algorithmic contexts (volume estimation, optimisation, feasibility testing), a weak-membership oracle suffices to implement efficient procedures. To precisely formulate the weak membership problem $\mathsf{WMEM}_\varepsilon(K)$, we must first establish several key concepts that underpin its definition. First, let us lay down some notational conventions used throughout the work. Given a vector $\mathbf{x} \in \mathbb{R}^D$, we define

$$\|\mathbf{x}\|_2 \coloneqq \sqrt{\sum_{i=1}^{D} x_i^2} \tag{S1.10}$$

the usual Euclidean $\ell_2$-norm. Given two vectors $\mathbf{x}, \mathbf{y}$, we denote their scalar product as $\mathbf{y}^T\mathbf{x} \in \mathbb{R}$.

**Definition 6** (Inner and outer core of a convex set)**.** *Let $K \subseteq \mathbb{R}^D$ be a subset of $\mathbb{R}^D$. We define the following.*

- *We define $S(K,\varepsilon) \coloneqq \{\mathbf{x} \in \mathbb{R}^D \mid \exists\, \mathbf{y} \in K : \|\mathbf{x}-\mathbf{y}\|_2 \leq \varepsilon\}$.*
- *We define $S(K,-\varepsilon) \coloneqq \{\mathbf{x} \in \mathbb{R}^D \mid S(\{\mathbf{x}\},\varepsilon) \subseteq K\}$, where $\{\mathbf{x}\}$ denotes the singleton set containing $\mathbf{x}$.*

In words: given a subset $K \subset \mathbb{R}^D$, $S(K,\varepsilon)$ is the *$\varepsilon$-outer neighborhood* of $K$, i.e. the set obtained by expanding $K$ outward by $\varepsilon$; while $S(K,-\varepsilon)$ is the *inner core* of $K$, i.e. the set of points whose $\varepsilon$-ball lies entirely within $K$.

For our purposes, it will also be useful to formalise the following concept.

**Definition 7** (Well-bounded centered set)**.** *Let $K$ be a subset of $\mathbb{R}^D$. Then, $K$ is called* well-bounded and $p$-centered *if there exist rationals $R > 0$, $r > 0$, and a point $\mathbf{p} \in K$ such that, with the origin $\mathbf{0} \in \mathbb{R}^D$,*

$$K \subseteq S(\{\mathbf{0}\}, R), \tag{S1.11}$$

$$K \supseteq S(\{\mathbf{p}\}, r). \tag{S1.12}$$

Intuitively, the condition $K \subseteq S(\{\mathbf{0}\}, R)$ ensures that $K$ is contained within some large ball of radius $R$ around the origin, i.e., $K$ has a bounded outer size. Meanwhile, the condition $K \supseteq S(\{\mathbf{p}\}, r)$ asserts that $K$ contains a full ball of radius $r$ around some point $\mathbf{p} \in K$, i.e., $K$ has a non-trivial *inner core*. Thus, the definition rules out both unboundedness and degeneracy: it ensures that $K$ is a reasonably sized set with some guaranteed volume (or thickness) from the inside, as well as a global bound from the outside.

Given Definitions 6 and 7, for the subset $K$, we can define weak membership as a rigorous decision problem.

**Problem 2** (Weak membership (WMEM))**.** *Let $K \subseteq \mathbb{R}^D$ be a compact, convex, and well-bounded $p$-centered set. Then, given $\mathbf{y} \in \mathbb{Q}^D$ and an error parameter $\varepsilon \in \mathbb{Q}$, s.t., $\varepsilon > 0$, decide:*

- *If $\mathbf{y} \in S(K, -\varepsilon)$, then output* YES.
- *If $\mathbf{y} \notin S(K, \varepsilon)$, then output* NO.

*An instance of* WMEM *is denoted as* $\mathsf{WMEM}_\varepsilon(K)$.

Similarly, it is possible to define a (polynomially) equivalent notion of weak membership called 1-side weak membership problem.

**Problem 3** (1-sided weak membership problem ($\mathsf{WMEM}^1$))**.** *Let $K \subseteq \mathbb{R}^D$ be a compact, convex, and well-bounded $p$-centered set. Then, given $\mathbf{y} \in \mathbb{Q}^D$ and the error parameter $\varepsilon \in \mathbb{Q}$, s.t. $\varepsilon > 0$, decide:*

- *If $\mathbf{y} \in K$, then output* YES.
- *If $\mathbf{y} \notin S(K, \varepsilon)$, then output* NO.

*An instance of* $\mathsf{WMEM}^1$ *is denoted as* $\mathsf{WMEM}^1_\varepsilon(K)$.

In what follows, we restate a theorem—previously established in Refs. [65, 66]—that demonstrates the polynomial equivalence of these problems. For the purposes of this section, we focus on one direction of the equivalence: specifically, we show that oracle access to $\mathsf{WMEM}^1$ suffices to solve WMEM.

**Theorem 4** (Lemma 2.5 and Lemma 2.6 of Refs. [65, 66])**.** *Let $K \subseteq \mathbb{R}^D$ be a compact, convex and well-bounded $p$-centered set with an associated radii $(R, r)$. Given an instance of $\mathsf{WMEM}_\delta(K)$, there exists an algorithm that runs in time $\mathsf{poly}(D, R, \lceil 1/\delta \rceil)$ and solves $\mathsf{WMEM}_\delta(K)$ using an oracle for $\mathsf{WMEM}^1_\varepsilon(K)$ with*

$$\varepsilon := \frac{\delta^3 r^2}{3 \cdot 2^{12} D^5 R^4}. \tag{S1.13}$$

Let us now introduce a notion closely related to that of membership. For a convex set $K$, deciding whether a point belongs to $K$ can be framed in terms of the existence of a linear witness that certifies non-membership. In this viewpoint, one seeks a hyperplane that supports $K$ and separates it from a given query point. This motivates the following definition, which formalises the notion of such a witness in terms of linear functionals bounded over $K$.

**Definition 8** (Witness of a set $K$)**.** *Let $K$ be a compact, convex and well-bounded centered set. Let $\mathbf{w} \in \mathbb{R}^D$ and $\gamma \in \mathbb{Q}$. Then $\mathbf{w}$ is a weak-witness for $K$ if and only if*

$$\mathbf{w}^T \mathbf{x} \leq \gamma, \quad \forall \mathbf{x} \in K. \tag{S1.14}$$

*Equivalently,* $\max_{\mathbf{x} \in K} \mathbf{w}^T \mathbf{x} \leq \gamma$.

For a given convex set $K$, an alternative to the membership problem is to consider its dual: rather than determining whether a point lies in $K$, we seek a separating hyperplane that certifies membership by distinguishing $K$ from its exterior, i.e., a witness for $K$. The precise formulation follows below, building on the *weak optimisation* problem (WOPT).

**Problem 4** (Weak optimisation (WOPT))**.** *Let $K \subseteq \mathbb{R}^D$ be a compact, convex and well-bounded $p$-centered set. Then given $\mathbf{w} \in \mathbb{Q}^D$ with $\|\mathbf{w}\|_2 \leq 1$ and $\gamma, \delta \in \mathbb{Q}$ such that $\delta > 0$ decide:*

- *If there exists $\mathbf{y} \in S(K, -\delta)$ with $\mathbf{w}^T \mathbf{y} \geq \gamma + \delta$, then output* YES.

- *If for all* $\mathbf{y} \in S(K, \delta)$, $\mathbf{w}^T\mathbf{y} \leq \gamma - \delta$, *then output* NO.

*An instance of* WOPT *is denoted as* $\mathsf{WOPT}_\delta(K)$.

Let us explain in words what Problem 4 is about: given a vector $\mathbf{w}$, WOPT is the decision problem that determines whether $\mathbf{w}$ is a witness for the set $K$ or not. As shown in Refs. [65, 66] and restated below, the weak optimisation problem can be solved efficiently given oracle access to WMEM. This implies the existence of a polynomial-time reduction in $D$ from WMEM to WOPT.

**Theorem 5** (Proposition 2.8 of Refs. [65, 66])**.** *Let $K \subseteq \mathbb{R}^D$ be a compact, convex and well-bounded p-centered set with an associated radii $(R, r)$. Given an instance of $\mathsf{WOPT}_\delta(K)$ with parameters $(\mathbf{w}, \gamma)$, there exists an algorithm that runs in time $\mathsf{poly}(D, R, \lceil 1/\delta \rceil)$ and solves $\mathsf{WOPT}_\delta(K)$ using an oracle for $\mathsf{WMEM}_\varepsilon(K)$ with $\varepsilon = r^3\delta^3/[2^{13}3^3D^5R^4(R+r)]$.*

Let us provide some intuition behind the theorem. The proof that having oracle access to $\mathsf{WMEM}_\varepsilon(K)$ solves $\mathsf{WOPT}_\delta(K)$ proceeds in two conceptual steps. First, one shows that the membership oracle can be used to simulate a weak separation oracle: by querying membership on carefully chosen test points, one either confirms that a point is "inside" (approximately), or else obtains evidence that the point lies outside $K$, from which one can infer a separating hyperplane. Secondly, starting from a known outer ball containing $K$, one alternates between querying the separation oracle and using the returned hyperplanes to cut away regions of space that cannot contain an optimal solution, thereby gradually shrinking the feasible region. After a polynomial number of iterations (in the dimension and precision parameters), this procedure yields a point that approximately maximises the objective over $K$, thus solving $\mathsf{WOPT}_\delta(K)$.

## 4. Preliminaries on $\varepsilon$-nets

This section reviews key notions related to $\varepsilon$-nets, which are part of our subsequent complexity-theoretic arguments. The material presented here follows the treatment in Ref. [67]; readers seeking further details are referred to that work. We begin with the concept of a covering net.

**Definition 9** ($\varepsilon$-covering net [67])**.** *Let $(T, \|\cdot\|)$ be a normed space with the metric induced by the norm, let $K \subseteq T$, and let $\varepsilon > 0$. A subset $C \subseteq K$ is called an* $\varepsilon$-covering net *of $K$ if for every $x \in K$ there exists $x_0 \in C$ such that $\|x - x_0\| \leq \varepsilon$. The* $\varepsilon$-covering number *of $K$, denoted by $\mathscr{C}(K, \|\cdot\|, \varepsilon)$, is defined as the minimal cardinality among all $\varepsilon$-covering nets of $K$. An $\varepsilon$-covering net $C$ is said to be* optimal *if $|C| = \mathscr{C}(K, \|\cdot\|, \varepsilon)$.*

Intuitively, an $\varepsilon$-covering net provides a finite approximation of the set $K$, in the sense that every point in $K$ lies within distance $\varepsilon$ of some net point. A dual notion is that of packing, which captures the maximum number of mutually distinguishable points that can be placed in $K$.

**Definition 10** (($\varepsilon$-packing net) [67])**.** *Let $(T, \|\cdot\|)$ be a normed space with the metric induced by the norm, let $K \subseteq T$, and let $\varepsilon > 0$. A subset $P \subseteq K$ is called an* $\varepsilon$-packing net *of $K$ if $\|x - y\| > \varepsilon$ for every pair of distinct points $x, y \in P$. The* $\varepsilon$-packing number *of $K$, denoted by $\mathscr{P}(K, \|\cdot\|, \varepsilon)$, is defined as the maximal cardinality among all $\varepsilon$-packing nets of $K$. An $\varepsilon$-packing net $P$ is said to be* optimal *if $|P| = \mathscr{P}(K, \|\cdot\|, \varepsilon)$.*

While a covering net ensures that all points are *close* to a net point, a packing net ensures that all net points are *far* from each other. This duality is formalised by the following fundamental lemma.

**Lemma 1** (Relation between covering and packing [67])**.** *Let $(T, \|\cdot\|)$ be a normed space, $K \subseteq T$, and let $\varepsilon > 0$. Then*

$$\mathscr{P}(K, \|\cdot\|, \varepsilon) \;\geq\; \mathscr{C}(K, \|\cdot\|, \varepsilon) \;\geq\; \mathscr{P}(K, \|\cdot\|, 2\varepsilon). \tag{S1.15}$$

In what follows, we will primarily bound covering numbers via packing numbers using Lemma 1, since packing bounds are often easier to obtain through volumetric arguments. Building upon these foundational definitions, we now review a series of standard bounds on covering and packing numbers, progressing from classical geometric objects to the quantum state spaces relevant to our analysis. For conciseness, we state these results without proof.

Let us specialise to the Euclidean space $\mathbb{R}^D$ equipped with the $\ell_2$-norm (see Eq. (S1.10)). We define the ball of radius $r$ and its boundary, the sphere, as

$$B_r(D) \coloneqq \{x \in \mathbb{R}^D : \|x\|_2 \leq r\}, \tag{S1.16}$$

$$\partial B_r(D) \coloneqq \{x \in \mathbb{R}^D : \|x\|_2 = r\}. \tag{S1.17}$$

The following proposition provides volumetric bounds on covering and packing numbers.

**Proposition 1** (Proposition 4.2.10 in Ref. [67]). *Let $K \subset \mathbb{R}^D$ and let $\varepsilon > 0$. Then*

$$\frac{\mathrm{Vol}(K)}{\mathrm{Vol}(B_\varepsilon(D))} \;\le\; \mathscr{C}(K, \|\cdot\|_2, \varepsilon) \;\le\; \mathscr{P}(K, \|\cdot\|_2, \varepsilon) \;\le\; \frac{\mathrm{Vol}(K + B_{\varepsilon/2}(D))}{\mathrm{Vol}(B_{\varepsilon/2}(D))}. \tag{S1.18}$$

Since $\partial B_1(D) \subseteq B_1(D)$, any $\varepsilon$-packing of the sphere is also an $\varepsilon$-packing of the ball. Consequently, the packing number of the ball upper bounds that of its boundary as

$$\mathscr{P}(\partial B_1(D), \|\cdot\|_2, \varepsilon) \;\le\; \mathscr{P}(B_1(D), \|\cdot\|_2, \varepsilon) \;\le\; \frac{\mathrm{Vol}(B_1(D)) + \mathrm{Vol}(B_{\varepsilon/2}(D))}{\mathrm{Vol}(B_{\varepsilon/2}(D))}. \tag{S1.19}$$

We now turn to packing numbers for the set of pure quantum states, which will allow us to discretise continuous families of quantum states.

**Definition 11** (Pure-state manifold [68]). *Let $D \in \mathbb{N}$ and $\mathcal{H} = \mathbb{C}^D$ equipped with the norm $\|\cdot\|_2$. Define the set of pure states as*

$$\mathcal{V}_D \coloneqq \{|\psi\rangle \in \mathcal{H} : \langle\psi|\psi\rangle = 1\}. \tag{S1.20}$$

Note that $\mathcal{V}_D$ can be identified with the unit sphere in $\mathbb{R}^{2D}$ via the real and imaginary parts of the complex amplitudes. Explicitly, the map $|\psi\rangle \mapsto (\mathrm{Re}\,\psi, \mathrm{Im}\,\psi)$ is an isometry between $(\mathbb{C}^D, \|\cdot\|_2)$ and $(\mathbb{R}^{2D}, \|\cdot\|_2)$. This observation allows us to transfer volumetric bounds from the Euclidean setting to the quantum state space.

**Lemma 2** (Packing bound for pure states). *Let $D \in \mathbb{N}$, $\mathcal{H} = \mathbb{C}^D$, and $\varepsilon \in [0, 1)$. Then*

$$\mathscr{P}(\mathcal{V}_D, \|\cdot\|_2, \varepsilon) \;\le\; \left(\frac{3}{\varepsilon}\right)^{2D}. \tag{S1.21}$$

*Proof.* The set $\mathcal{V}_D$ can be identified with the boundary $\partial B_1(2D)$ of the unit ball in $\mathbb{R}^{2D}$. Applying the volumetric upper bound in Proposition 1 to $B_1(2D)$ and using (S1.19) yields

$$\mathscr{P}(\partial B_1(2D), \|\cdot\|_2, \varepsilon) \;\le\; \frac{\mathrm{Vol}(B_1(2D)) + \mathrm{Vol}(B_{\varepsilon/2}(2D))}{\mathrm{Vol}(B_{\varepsilon/2}(2D))} \;\le\; \left(\frac{3}{\varepsilon}\right)^{2D}, \tag{S1.22}$$

where the last inequality follows from the standard scaling $\mathrm{Vol}(B_r(2D)) = r^{2D}\mathrm{Vol}(B_1(2D))$. Identifying $\mathcal{V}_D$ with $\partial B_1(2D)$ completes the proof. □

### 5. Magic-state resource theory

In this section, we review the main notions of magic-state resource theory, which form the core of our work.

Magic-state resource theory provides a rigorous and operational framework for quantifying the non-stabilizer features of quantum states and channels that are necessary for universal quantum computation. In this formalism, one identifies a distinguished set of operations and states that are considered *free* because they can be implemented efficiently and simulated classically. Specifically, the stabilizer formalism—comprising Clifford unitaries, preparations of stabilizer states, Pauli measurements, classical feed-forward, and discarding of subsystems—defines a subclass of quantum computations that admit efficient classical simulation via the Gottesman–Knill theorem [17].

Within this resource theory, the *free states* are those confined to the stabilizer polytope $\hat{\mathcal{S}}_n$, i.e. the convex hull of all stabilizer states, which can be prepared using only free operations and therefore do not contribute any computational advantage beyond classical simulation. Free operations are taken to be stabilizer-preserving maps that, by definition, cannot generate non-stabilizer features from free states. A state lying outside this polytope is said to possess *magic* [69]—a resource quantified by magic monotones that capture the extent to which the state enables computational tasks beyond the stabilizer regime.

In what follows, we define $\mathcal{S}_n$ the set of pure stabilizer states and $\hat{\mathcal{S}}_n$ its convex hull, i.e., the stabilizer polytope. We denote by $\mathbb{P}_n$ the Pauli group on $n$ qubits and $P \in \mathbb{P}_n$ Pauli strings. Let us formally define magic monotones below.

**Definition 12** (Magic monotones). *A stabilizer monotone $\mathcal{M}$ is a real-valued function for all $n$ qubit systems (or collections thereof) such that (i) $\mathcal{M}(\rho) = 0$ if and only if $\rho \in \hat{\mathcal{S}}_n$; (ii) $\mathcal{M}$ is non-increasing under free operations $\mathcal{M}(\mathcal{E}(\rho)) \le \mathcal{M}(\rho)$ for any free operation $\mathcal{E}$. A function $\mathcal{M}$ is pure-state stabilizer monotone if condition (i) is satisfied for pure states and (ii) holds for any pair $(|\psi\rangle, \mathcal{E})$, obeying $\mathcal{E}(|\psi\rangle\langle\psi|) = |\phi\rangle\langle\phi|$.*

The arguably most important pure-state magic monotone is the stabilizer entropy, defined below.

**Definition 13** (Stabilizer Rényi entropy [16])**.** *Given a pure state $\psi$, the $\alpha$-stabilizer entropy is defined as*

$$M_\alpha(|\psi\rangle) = \frac{1}{1-\alpha}\log\frac{1}{d}\sum_P \mathrm{tr}^{2\alpha}(P\psi)\,. \tag{S1.23}$$

*$M_\alpha$ obeys the following properties.*

*(i) It is a pure-state magic monotone [30] for any $\alpha \geq 2$.*

*(ii) Its computation lies in $\mathsf{TIME}(2^{O(n)})$ and thus in $\mathsf{P}$.*

*(iii) It provides a family of efficiently computable witnesses [38]: let $\rho$ be a mixed state and $M_\alpha(\rho)$ computed through Eq. (S1.23), then*

$$M_\alpha(\rho) - \frac{1-2\alpha}{1-\alpha}S_2(\rho)\,, \tag{S1.24}$$

*where $S_2(\rho) \coloneqq -\log \mathrm{tr}\,\rho^2$, is a magic witness for any $\alpha \geq \frac{1}{2}$.*

*(iv) It can be extended to mixed states through the "convex roof" or convex hull construction [30]*

$$\tilde{M}_\alpha(\rho) = \frac{1}{1-\alpha}\log\sup_{(q_j,\phi_j)}\sum_j q_j\left(\frac{1}{d}\sum_{P\in\mathbb{P}_n}\mathrm{tr}^{2\alpha}(P\psi)\right) \tag{S1.25}$$

*which can be computed in time $\mathsf{TIME}(2^{O(nr^2)})$ with $r = \mathrm{rank}(\rho)$.*

The key fact that makes the stabilizer entropy relevant in the context of this work is that it can be computed in polynomial-time in the Hilbert space dimension. Indeed, given a quantum state $\rho = |\psi\rangle\langle\psi|$ corresponding to a pure state, decision problems associated with the computation of $M_\alpha(\rho)$ lie in $\mathsf{QP}^1 = \mathsf{P}$ in the input parameter $d = 2^n$. In sharp contrast, the main result of this paper demonstrates that the computation of any magic resource monotone working on mixed states is fundamentally super-polynomial; in particular (according to Corollary 1), it lies in $\mathsf{QP}^k$ for $k \geq 2$ (see Definition 4). Below, we introduce a magic monotone whose computation requires $\mathsf{TIME}(2^{O(\log^2 d)})$ thus lying in $\mathsf{QP}^2$, thus matching the optimal computational time for magic monotones.

**Definition 14** (Robustness of magic [21])**.** *Given a $n$-qubit density matrix $\rho$. The robustness of magic is defined*

$$\mathcal{R}(\rho) \coloneqq \min_x\left\{\sum_i |x_i| \;:\; \rho = \sum_i x_i\sigma_i,\;\; \sigma_i \in \mathcal{S}_n\right\}. \tag{S1.26}$$

*$\mathcal{R}$ is a magic monotone.*

Similarly to magic monotones, one can define witnesses within the magic-state resource theory.

**Definition 15** (Magic witness)**.** *Let $\gamma \in \mathbb{Q}$. A Hermitian operator $W$ is called a* magic witness *if $\mathrm{tr}(W\sigma) \leq \gamma$ for every stabilizer state $\sigma \in \hat{\mathcal{S}}_n$.*

As anticipated in the main text, a clarifying remark is in order. For general mixed states, the robustness of magic—despite its optimality as a resource monotone—incurs a computational cost scaling as $2^{\Theta(n^2)}$, rendering it intractable for systems beyond $n \simeq 5$, as also noted in Ref. [21]. Remarkably, however, the extended stabiliser entropy $\tilde{M}_\alpha$ admits a polynomial-time computation when the rank satisfies $r = O(1)$. This feature is crucial, as it enables the systematic study of magic-state resource theory beyond pure states, extending its applicability to low-rank density matrices. Low rank states are ubiquitous in many-body physics. A prominent example is provided by *matrix product states* (MPS) with constant bond dimension, which naturally arise as ground states of one-dimensional gapped many-body Hamiltonians. For these systems, the magic content of reduced density matrices can be faithfully quantified using the extended stabilizer entropy, with a computational cost that is polynomial in the Hilbert space dimension. As emphasised in the main text, this represents a substantial practical advantage over the super-exponential scaling $2^{\Theta(n^2)}$ associated with robustness-based measures. An intriguing open problem is whether the computational cost of the extended stabilizer entropy for reduced density matrices of MPS states can be further reduced from $2^{\Theta(n)}$ to $\mathsf{poly}(n)$, as is the case for the stabilizer entropy of the MPS state [70, 71].

Another magic monotone that plays a crucial role in magic-state resource theory due to its intuitive operational meaning is the *stabilizer fidelity* [49], i.e., the fidelity of a given pure state with the closest state in the stabilizer polytope.

**Definition 16** (Stabilizer fidelity [49]). *Let $\rho$ be a (possibly mixed) state. Its stabilizer fidelity $F_{stab}$ is defined as*

$$F_{stab}(\rho) = \max_{\sigma \in \hat{\mathcal{S}}_n} \|\sqrt{\sigma}\rho\sqrt{\sigma}\|_1. \tag{S1.27}$$

For pure states, the stabilizer fidelity is closely related to the stabilizer entropy. The following lemma will be useful in the remainder of the manuscript.

**Lemma 3** (Relation between stabilizer entropy and stabilizer fidelity [31, 72, 73]). *Let $\psi$ be a pure quantum state. Let $P_6(\psi) := 2^{-2M_3(\psi)}$ with $M_3$ the $(\alpha = 3)$ stabilizer entropy. Then it holds that*

$$(P_6(\psi))^C \leq F_{stab}(\psi) \leq (P_6(\psi))^{1/6} \tag{S1.28}$$

*for some constant $C > 0$.*

Resource monotones are useful in the central tasks of quantum resource theory, which is resource distillation [21, 50, 69]. Indeed, magic-state distillation [69] is a cornerstone of fault-tolerant quantum computing within the magic resource theory framework. While Clifford gates and stabilizer states—the free resources in this theory—can be made fault-tolerant, they are insufficient for universal quantum computation. Universality requires non-Clifford gates, such as the $T$-gate or the Toffoli gate, which are enabled by high-fidelity resource states like the single-qubit $|T\rangle$ or the multi-qubit $|CCZ\rangle$ state vector. Since these magic states cannot be prepared fault-tolerantly directly, distillation protocols are essential: they consume multiple noisy copies to produce fewer copies of a purified magic state.

A primary historical bottleneck has been the resource overhead, defined as the number of noisy input states required per clean output. Early protocols [69] exhibited overhead scaling as $O(\log^\gamma(1/\varepsilon))$ for a target error $\varepsilon$, with $\gamma > 1$. Recent theoretical breakthroughs have dramatically improved this scaling. Notably, protocols achieving constant overhead [74] ($\gamma = 0$) have been developed by constructing asymptotically good quantum codes with transversal non-Clifford gates and efficient decoders, making them suitable for distilling magic state vectors like $|T\rangle$ and $|CCZ\rangle$. This progress is being matched by experimental advances, including the first logical-level demonstrations [75]. Beyond engineering efficient protocols, the formal framework of magic resource theory [14]—extended to quantum channels [23]—provides the necessary tools to quantify magic and establish fundamental lower bounds on the resource overhead of any distillation protocol, delineating what is possible from what is provably hard.

Besides the task of resource distillation through stabilizer protocols, the magic-state resource theory is also important because it aims to quantify the resources a quantum algorithm requires to achieve some form of advantage over classical computation. In particular, stabilizer states do not yield any computational advantage, as their outcome distributions—when measured in the computational basis—can be efficiently simulated classically in polynomial time [47]. From this fact, within the resource theory of magic, one identifies additional classes of operations that strictly extend the stabilizer operations, known as *completely stabilizer operations*. These are quantum channels that, when applied to stabilizer states, return only stabilizer states. We define them precisely below.

**Definition 17** (Completely stabilizer-preserving channel [23]). *Let $\mathcal{E}$ be a quantum channel acting on an $n$-qubit input state. We say that $\mathcal{E}$ is completely stabilizer-preserving iff $[\mathcal{E} \otimes I](\sigma) \in \hat{\mathcal{S}}_{n+m}$ for all $\sigma \in \hat{\mathcal{S}}_{n+m}$ $\forall m$.*

These channels play a pivotal role because, as they cannot create non-stabilizer states from scratch, their action can be simulated classically in full, and no magic state can be distilled from their output. Moreover, as it has been shown in Ref. [76], the set of completely stabilizer operations is strictly larger than the one obtained by arbitrary combinations of the basic stabilizer protocols introduced above.

Ref. [23] provides a characterisation of completely stabilizer-preserving channels in terms of the Choi state $\rho_{\mathcal{E}}$ uniquely associated with the channel, which we recall below for completeness.

**Lemma 4** (Characterisation of completely stabilizer preserving channels [23]). *Given a channel $\mathcal{E}$, it is stabilizer preserving if and only if the Choi state $\rho_{\mathcal{E}} \in \hat{\mathcal{S}}_{2n}$.*

### 6. Summary of results on $t$-doped stabilizer states

Within the stabilizer formalism, stabilizer states are not the only states that can be efficiently simulated classically. Indeed, small perturbations of stabilizer states can also remain classically simulable. In particular, adding $t$ local non-Clifford gates produces states that can be simulated in time $\mathrm{poly}(n, 2^{O(t)})$ [47, 48]. Below, we define a class of states that is strictly larger than this: the $t$-doped states.

**Definition 18** ($t$-doped states [52, 77]). *A state vector $|\psi\rangle$ on $n$ qubits is called* $t$-doped *if there exists a Clifford unitary $C \in \mathcal{C}_n$ and a $t$-qubit state vector $|\phi\rangle_t$ such that*

$$|\psi\rangle = C\big(|0\rangle^{\otimes(n-t)} \otimes |\phi\rangle_t\big). \tag{S1.29}$$

*We denote by $\mathcal{S}_{n,t}$ the set of all $t$-doped pure states, and by $\hat{\mathcal{S}}_{n,t}$ its convex hull.*

This definition induces a natural hierarchy of quantum states indexed by the parameter $t$. When $t = 0$, the class $\mathcal{S}_{n,0}$ coincides exactly with the set of stabilizer states. More generally, the definition imposes a strong structural constraint on the state. When $C$ is the identity, the reference state vector $|0\rangle^{\otimes(n-t)}$ is stabilised by the operators $Z_1, \ldots, Z_{n-t}$, implying that it admits a stabilizer group of size at least $2^{n-t}$. Since Clifford unitaries preserve stabilizer structure, every $t$-doped state is stabilised by at least $2^{n-t}$ Pauli operators. Consequently, for small values of $t$, such as $t = O(\log n)$, the state differs from a stabilizer state only on a logarithmic number of qubits, up to a Clifford transformation, and all non-stabilizer structure can be localised to a small subsystem. At the opposite extreme, when $t = \Theta(n)$, the class becomes universal, as any $n$-qubit state is trivially $n$-doped. In this sense, $t$ quantifies the effective "non-stabilizer dimension" of the state, namely the number of qubits required to store all its magic once stabilizer correlations have been disentangled. From a computational perspective, $t$-compressibility guarantees efficient classical evaluation of measurement statistics whenever $t$ is small [48]. The parameter $t$ also marks a sharp computational boundary. For $t = O(\log n)$, the associated classical simulation algorithms run in polynomial time. By contrast, when $t$ grows superlogarithmically, there exist $t$-doped stabilizer states that are believed to be classically intractable [77]: i.e., there must exist $t$-doped stabilizer states with $t = \omega(\log n)$ that cannot be efficiently simulated classically, in line with standard expectations for the emergence of quantum advantage.

The class of $t$-doped stabilizer states naturally contains many previously studied families of efficiently simulable states. In particular, any state obtained from a stabilizer state by applying at most $t$ local non-Clifford gates is a $O(t)$-doped stabilizer state [51, 78]. In the following lemma, we bound the number of $t$-doped stabilizer states given a tolerance $\varepsilon$ in trace distance.

**Lemma 5** (Packing number of $t$-doped stabilizer states)**.** *Let $\varepsilon > 0$, and let $\mathcal{S}_{n,t}$ denote the set of all pure $t$-doped stabilizer states on $n$ qubits. The $\varepsilon$-packing number of $\mathcal{S}_{n,t}$ with respect to the trace distance satisfies*

$$\mathscr{P}(\mathcal{S}_{n,t}, \varepsilon) \leq 2^{2n^2+3n} \left(\frac{3}{\varepsilon}\right)^{2^{2t}}. \tag{S1.30}$$

*Proof.* By definition, any $t$-doped stabilizer state can be written, using Definition 18, in the form

$$|\psi\rangle = C\,|0\rangle^{\otimes(n-t)} \otimes |\phi\rangle_t\,, \tag{S1.31}$$

where $|\phi\rangle_t$ is a state vector on $t$ qubits, up to the action of a Clifford unitary on $n$ qubits. Consequently, an $\varepsilon$-packing of $\mathcal{S}_{n,t}$ can be constructed by taking the orbit of an $\varepsilon$-packing of the set $\mathcal{V}_{2^t}$ of pure $t$-qubit states under the action of the $n$-qubit Clifford group. This yields the upper bound $\mathscr{P}(\mathcal{S}_{n,t}, \varepsilon) \leq |\mathcal{C}_n|\, \mathscr{P}(\mathcal{V}_{2^t}, \varepsilon)$, where $\mathcal{C}_n$ denotes the $n$-qubit Clifford group. Using the standard upper bound $|\mathcal{C}_n| \leq 2^{n^2+2n} \prod_{j=1}^{n}(4^j - 1)$ and applying Proposition 1, which gives $\mathscr{P}(\mathcal{V}_{2^t}, \varepsilon) \leq \left(\frac{3}{\varepsilon}\right)^{2^{2t}}$, we obtain

$$\mathscr{P}(\mathcal{S}_{n,t}, \varepsilon) \leq 2^{n^2+2n} \prod_{j=1}^{n}(4^j - 1) \left(\frac{3}{\varepsilon}\right)^{2^{2t}}. \tag{S1.32}$$

Finally, using the bound $\prod_{j=1}^{n}(4^j - 1) \leq 2^{n(n+1)}$, we can conclude that

$$\mathscr{P}(\mathcal{S}_{n,t}, \varepsilon) \leq 2^{2n^2+3n} \left(\frac{3}{\varepsilon}\right)^{2^{2t}}. \tag{S1.33}$$

□

### Appendix S.2: Computational problems in magic-state resource theory

In this section, we define the computational problems addressed in this paper within the framework of magic state resource theory. We begin by restating the definition introduced in Appendix S.1 3, adapting it to the case of convex sets of quantum states. This specialisation is motivated by the fact that the stabilizer polytope, denoted by $\hat{\mathcal{S}}_n$, is a convex set by definition. However, we keep the discussion as general as possible by referring to an arbitrary convex set of states, which we denote by $\mathcal{K}$ throughout.

**Definition 19.** *Let $\mathcal{K} \subseteq \mathcal{H}_n$ be a convex subset of $\mathcal{H}_n$. We define the following.*

- *We define $S(\mathcal{K}, \varepsilon) \coloneqq \{\rho \in \mathcal{H}_n \mid \exists\, \sigma \in \mathcal{K} : \|\rho - \sigma\|_2 \leq \varepsilon\}$.*
- *We define $S(\mathcal{K}, -\varepsilon) \coloneqq \{\rho \in \mathcal{K} \mid S(\{\rho\}, \varepsilon) \subseteq \mathcal{K}\}$, where $\{\rho\}$ denotes the singleton set containing the state $\rho$.*

This definition mimics the one given for convex sets of $\mathbb{R}^D$, introducing a notion of the inner core and outer neighbourhood for convex sets of states. The careful reader will notice that in Definition 19 we use the Frobenius norm instead of the trace distance, which is the metric typically used to characterise the distance between quantum states. The two norms are related by the inequalities $\|\rho-\sigma\|_2 \leq \|\rho-\sigma\|_1 \leq \sqrt{d}\,\|\rho-\sigma\|_2$, that is, they differ by at most a factor polynomial in the dimension of the Hilbert space. Since the access model in the computational problems studied here consists of a $d\times d$ density matrix, the two distances are therefore equivalent up to polynomial factors. We thus work with the Frobenius norm because of its direct connection to the Euclidean $\ell_2$-norm on vectors, as we explain below. We can also generalise the notion of a well-bounded centered set, and we obtain the following.

**Definition 20** (Well-bounded centered set for states)**.** *Let $\mathcal{K}$ be a convex subset of $\mathcal{H}_n$. Then, $\mathcal{K}$ is called* well-bounded and $\overline{\rho}$-centered *if there exist rationals $R>0$, $r>0$, and a state $\overline{\rho}\in\mathcal{K}$ such that, with the state $I/d\in\mathcal{H}_n$,*

$$\mathcal{K}\subseteq S(I/d,R), \tag{S2.1}$$

$$\mathcal{K}\supseteq S(\overline{\rho},r). \tag{S2.2}$$

Again, this notion provides a generalisation of the concept of a *well-bounded centered* set, but for states. We can now proceed to introduce the notion of the weak membership problem for a set of quantum states. We define it for a general convex set $\mathcal{K}$, since in what follows we will consider not only the stabilizer polytope but also the convex hull of $t$-doped states introduced in Appendix S.1 6.

**Problem 5** ($\rho$-weak membership problem ($\rho$-WMEM))**.** *Let $\mathcal{K}\subseteq\mathcal{H}_n$ be a convex and well-bounded $\overline{\rho}$-centered set for states on $n$ qubits. Let $\rho\in\mathcal{D}(\mathcal{H}_n)$ be a quantum state and $\varepsilon>0$ a error parameter. Decide:*

- *If $\rho\in S(\mathcal{K},-\varepsilon)$ output* YES.
- *If $\rho\notin S(\mathcal{K},\varepsilon)$ output* NO.

*An instance of $\rho$-*WMEM *is denoted as $\rho$-*$\mathsf{WMEM}_\varepsilon(\mathcal{K})$.

Let us now define a (polynomially) equivalent version of the membership problem, namely the 1-sided weak membership problem.

**Problem 6** (1-sided $\rho$-weak membership problem $\rho$-$\mathsf{WMEM}^1$)**.** *Let $\mathcal{K}\subseteq\mathcal{H}_n$ be a convex and well-bounded $\overline{\rho}$-centered set for $n$ qubits. Let $\rho\in\mathcal{D}(\mathcal{H}_n)$ be a quantum state and $\varepsilon>0$ a error parameter. Decide:*

- *If $\rho\in\mathcal{K}$ output* YES*;*
- *If $\rho\notin S(\mathcal{K},\varepsilon)$ output* NO.

*An instance of $\rho$-*$\mathsf{WMEM}^1$ *is denoted as $\rho$-*$\mathsf{WMEM}^1_\varepsilon(\mathcal{K})$.

Here, we have presented a generalisation of the membership problem to an arbitrary convex set of quantum states. Since quantum states can be represented as vectors in $\mathbb{R}^D$ with $D=d^2-1$ [79], $\rho$-$\mathsf{WMEM}_\varepsilon(\mathcal{K})$ can be naturally related to $\mathsf{WMEM}_\varepsilon(K)$. The following lemma formalise this relationship by establishing the equivalence of the two problems.

**Lemma 6** (Equivalence between $\rho$-$\mathsf{WMEM}^{(1)}$ and $\mathsf{WMEM}^{(1)}$)**.** *Let $\mathcal{K}$ be a convex and well-bounded $\overline{\rho}$-centered set for quantum states on $\mathcal{H}_n$, and let $K\subseteq\mathbb{R}^D$ with $D=d^2-1$ denote its image under the Pauli-coordinate map $\mathbf{x}:\rho\mapsto\mathbf{x}(\rho)$, where $\mathbf{x}(\rho)=(\mathrm{tr}(\rho P)/\sqrt{d}:P\neq I)$. The set $K$ is well-bounded and centered in the origin with*

$$R=\sqrt{1-1/d}, \tag{S2.3}$$

$$r=\frac{1}{\sqrt{d(d-1)}}. \tag{S2.4}$$

*Moreover an instance of $\rho$-*$\mathsf{WMEM}_\varepsilon(\mathcal{K})$ *is equivalent to an instance of* $\mathsf{WMEM}_\varepsilon(K)$*. Similarly, an instance of $\rho$-*$\mathsf{WMEM}^1_\varepsilon(\mathcal{K})$ *is equivalent to an instance of* $\mathsf{WMEM}^1_\varepsilon(K)$.

*Proof.* Every state has a Pauli expansion $\rho=\frac{1}{d}I+\frac{1}{d}\sum_{P\neq I}x_P P$ with $x_P=\mathrm{tr}(\rho P)$. For pure states, the coefficients satisfy $\sum_{P\neq I}x_P^2=1-1/d$, which shows that the entire state space embeds into the Euclidean ball $B(0,\sqrt{1-1/d})\subset\mathbb{R}^D$. Thus $K=\mathbf{x}(\mathcal{K})$ is compact and convex; moreover, by Ref. [79, Section 1.2], the full state space is a well-bounded centered set with radii $R=\sqrt{(1-1/d)}$ and $r=\sqrt{1/(d^2-d)}$. The map $\mathbf{x}$ maps the Hilbert-Schmidt distance

$$\|\rho-\rho'\|_2=\|\mathbf{x}(\rho)-\mathbf{x}(\rho')\|_2 \tag{S2.5}$$

to the Euclidean distance with no prefactors. Hence, an $\varepsilon$-membership test in state space coincides exactly with an $\varepsilon$-membership test in the image set $K$. This proves the claim. □

After defining the membership problem, one can extend the notion of a separating hyperplane. In this sense, we can discuss *witnesses*, operators that separate the states belonging to the convex set $\mathcal{K}$ from those that do not belong to it. This leads to the following definition:

**Definition 21** (Witness for a convex set of states)**.** *Let $\mathcal{K}$ be a convex set of states on $\mathcal{H}_n$, let $W$ be a Hermitian operator, and let $\gamma \in \mathbb{Q}$. We say that $W$ is a* witness *for $\mathcal{K}$ if*

$$\max_{\sigma\in\mathcal{K}} \operatorname{tr}(W\sigma) \leq \gamma. \tag{S2.6}$$

Some authors define a witness for a convex set by additionally requiring the existence of at least one state $\rho \notin \mathcal{K}$ such that $\operatorname{tr}(W\rho) > \gamma$, thereby ensuring that $W$ is a "useful" witness. We instead adopt the above definition without this extra condition in order to guarantee that the set of witnesses remains convex, which is crucial for the proof of Lemma 11, and centered at the identity, which is, of course, a "useless" witness.

Note that if $\mathcal{V}$ is a set of pure states and $\mathcal{K} = \operatorname{conv}(\mathcal{V})$ is their convex hull, then the condition of Eq. (S2.6) reduces to

$$\max_{\psi\in\mathcal{V}} \langle\psi|W|\psi\rangle \leq \gamma. \tag{S2.7}$$

This follows because the map $\sigma \mapsto \operatorname{tr}(W\sigma)$ is linear, and a linear functional attains its maximum over a convex set at one of its extreme points. Since the extreme points of $\mathcal{K}$ are precisely the pure states in $\mathcal{V}$, it suffices to verify the inequality only on $\mathcal{V}$.

Related to the notion of a witness, we state the following lemma, which explores the relationship between membership and witnesses.

**Lemma 7** (Witness vs. weak membership)**.** *Given a state $\rho$ such that $\|\rho - \sigma\|_2 > 0$ for all $\sigma \in \mathcal{K}$, there exists a witness operator $W$ that detects $\rho$ in the sense that*

$$\operatorname{tr}(W\rho) > \max_{\sigma\in\mathcal{K}} \operatorname{tr}(W\sigma) + \min_{\sigma\in\mathcal{K}} \|\rho - \sigma\|_2^2 . \tag{S2.8}$$

Conversely, if there exists a witness that detects $\rho$, i.e., $\operatorname{tr}(W\rho) > \max_{\sigma\in\mathcal{K}} \operatorname{tr}(W\sigma) + \varepsilon$, then it holds that $\min_{\sigma\in\mathcal{K}} \|\rho - \sigma\|_2 > \frac{\varepsilon}{\|W\|_2}$.

*Proof.* First of all, let us define $\tau = \arg\min_{\sigma\in\mathcal{K}} \|\rho - \sigma\|_2$. We have $\|\rho - \tau\|_2 > \varepsilon$. The state $\tau + t(\sigma - \tau) \in \mathcal{K}$ for $\sigma \in \mathcal{S}_n$ since it is a convex combination of $\tau$ and $\sigma$. The function $f(t) = \|\rho - \tau - t(\sigma - \tau)\|_2^2 = \operatorname{tr}\big[(\rho - \tau - t(\sigma - \tau))^2\big]$ achieves its minimum at $t = 0$ by construction. Differentiating with respect to $t$, we have

$$f'(t) = 2\operatorname{tr}[(\rho - \tau - t(\sigma - \tau))(\tau - \sigma)] \implies f'(0) = \operatorname{tr}((\rho - \tau)(\tau - \sigma)). \tag{S2.9}$$

Since $t = 0$ is a minimum, we have $f'(0) \geq 0$. We define the witness as $W = \rho - \tau$. Indeed, we have

$$\begin{aligned} \operatorname{tr}(W(\rho - \sigma)) &= \operatorname{tr}((\rho - \tau)(\rho - \sigma)) \\ &= \operatorname{tr}((\rho - \tau)(\rho - \tau)) + \operatorname{tr}((\rho - \tau)(\tau - \sigma)) \\ &\geq \|\rho - \tau\|_2^2, \end{aligned} \tag{S2.10}$$

proving the first statement. Conversely, if there exists $W$ such that $\operatorname{tr}(W\rho) > \max_\sigma \operatorname{tr}(W\sigma) + \varepsilon$. It follows that for any $\tau \in \mathcal{K}$, it holds that $\operatorname{tr}(\tau W) \leq \max_\sigma \operatorname{tr}(W\sigma) < \operatorname{tr}(W\rho) - \varepsilon$. Hence, we have $\varepsilon < \operatorname{tr}(W[\rho - \tau]) \leq \|W\|_2 \|\rho - \tau\|_2$, proving the statement. □

Similarly to what was done for convex sets of $\mathbb{R}^D$, we can introduce the problem of deciding whether an operator is a witness or not.

**Problem 7** (Weak witnessing detection (WWD))**.** *Let $\mathcal{K} \subseteq \mathcal{B}(\mathcal{H}_n)$ be a convex set and well-bounded $\overline{\rho}$-centered set for $n$-qubit states, let $W \in \mathcal{B}(\mathcal{H}_n)$ satisfy $\|W\|_2 \leq 1$, and let $\gamma, \delta \in \mathbb{Q}$ with $\delta > 0$. Decide:*

- *If some $\rho \in S(\mathcal{K}, -\delta)$ satisfies $\operatorname{tr}(W\rho) \geq \gamma + \delta$, output* YES.
- *If all $\rho \in S(\mathcal{K}, \delta)$ satisfy $\operatorname{tr}(W\rho) \leq \gamma - \delta$, output* NO.

*An instance is denoted* $\mathsf{WWD}_\delta(\mathcal{K})$.

The careful reader may have noticed that, in the promise of Problem 7, a witness $W$ is required to satisfy $\|W\|_2 \leq 1$, whereas in the definition of a witness given in Definition 21 no such condition is imposed. This discrepancy merely reflects a rescaling of the constant $\gamma \in \mathbb{Q}$ and ultimately only affects the scaling in the reduction used in Lemma 9.

**Lemma 8.** *An instance of* $\mathsf{WWD}_\delta(\mathcal{K})$ *with parameters* $(W, \gamma)$ *corresponds to* $\mathsf{WOPT}_\delta(K)$ *with parameters* $(\mathbf{w}, \gamma - \mathrm{tr}(W)/d)$*, where* $K = \mathbf{x}(\mathcal{K}) \subset \mathbb{R}^D$ *with* $D = d^2 - 1$ *and* $\mathbf{x}(\rho)$ *is the Pauli-coordinate map, and* $\mathbf{w} = \mathbf{x}(W)$*.*

*Proof.* Mapping $W$ to $\mathbf{w} = (\mathrm{tr}(WP) : P \neq I)$ gives $\mathrm{tr}(W\rho) = \mathbf{w}^T\mathbf{x} + \mathrm{tr}(W)/d$ for any $\rho$ with coordinates $\mathbf{x}$. optimising $\mathrm{tr}(W\rho)$ over $\mathcal{K}$ is therefore equivalent to optimising $\mathbf{w}^T\mathbf{x}$ over $K$, up to the constant shift $\mathrm{tr}(W)/d$. For a YES instance of $\mathsf{WOPT}_\delta(K)$, the condition $\mathbf{w}^T\mathbf{x} \geq \gamma - \mathrm{tr}(W)/d + \delta$ implies $\mathrm{tr}(W\rho) \geq \gamma + \delta$, matching the YES case of $\mathsf{WWD}_\delta(\mathcal{K})$. For a NO instance, $\mathbf{w}^T\mathbf{x} \leq \gamma - \mathrm{tr}(W)/d - \delta$ yields $\mathrm{tr}(W\rho) \leq \gamma - \delta$, matching the NO case. □

Given the connection between $\mathsf{WMEM}_\varepsilon$ and $\mathsf{WOPT}_\delta$, one can expect an analogous relation between $\rho$-$\mathsf{WMEM}_\varepsilon$ and $\mathsf{WWD}_\delta$. This correspondence follows directly from Theorem 5 and from the observation that $\rho$-$\mathsf{WMEM}_\varepsilon$ can be embedded into $\mathsf{WMEM}_\varepsilon$, while $\mathsf{WWD}_\delta$ can be embedded into $\mathsf{WOPT}_\delta$. For completeness, we state the following lemma to formalise this connection.

**Lemma 9** (Reduction from $\rho$-$\mathsf{WMEM}_\varepsilon$ to $\mathsf{WWD}_\delta$)**.** *Let* $\mathcal{K}$ *be a convex and well-bounded* $\overline{\rho}$*-centered set for quantum states. Given an instance of* $\mathsf{WWD}_\delta(\mathcal{K})$ *with parameters* $(W, \gamma)$*, where* $W$ *is an operator on* $\mathcal{H}_n$ *with* $\|W\|_2 \leq 1$*,* $\gamma \in \mathbb{Q}$*, and* $\delta \in (0, 1)$*, there exists an algorithm that runs in time* $\mathsf{poly}(d, \delta^{-1})$ *and solves* $\mathsf{WWD}_\delta(\mathcal{K})$ *using an oracle for* $\rho$-$\mathsf{WMEM}_\varepsilon(\mathcal{K})$ *with* $\varepsilon = \frac{\delta^3}{2^{13} \cdot 3^3 \cdot d^{13}}$*.*

*Proof.* The proof follows directly from Lemmas 6 and 8. Given the equivalence between the protocols $\mathsf{WMEM}_\varepsilon$ and $\rho$-$\mathsf{WMEM}_\varepsilon$, as well as between $\mathsf{WOPT}_\delta$ and $\mathsf{WWD}_\delta$, we can construct a polynomial-time algorithm for $\mathsf{WWD}_\delta$ with oracle access to $\rho$-$\mathsf{WMEM}_\delta$. The approach proceeds by first reducing the weak witness problem to its optimisation counterpart $\mathsf{WOPT}_\delta$, then applying Theorem 5 to complete the reduction. □

To conclude this preliminary section, we recall that the two membership problems defined above, namely $\rho$-$\mathsf{WMEM}$ and $\rho$-$\mathsf{WMEM}^1$, are polynomially equivalent. This correspondence follows directly from Theorem 4 and from the observation that $\rho$-$\mathsf{WMEM}_\varepsilon$ can be embedded into $\mathsf{WMEM}_\varepsilon$, while $\rho$-$\mathsf{WMEM}^1$ can be embedded into $\mathsf{WMEM}^1$. For completeness, we state the following lemma to formalise this connection.

**Lemma 10** (Reduction from $\rho$-$\mathsf{WMEM}^1$ to $\rho$-$\mathsf{WMEM}$)**.** *Let* $\mathcal{K}$ *be a convex and well-bounded* $\overline{\rho}$*-centered set for quantum states. Given an instance of* $\rho$-$\mathsf{WMEM}_\varepsilon(\mathcal{K})$*, and* $\varepsilon \in (0, 1)$*, there exists an algorithm that runs in time* $\mathsf{poly}(d, \delta^{-1})$ *and solves* $\rho$-$\mathsf{WMEM}_\varepsilon(\mathcal{K})$ *using an oracle for* $\rho$-$\mathsf{WMEM}^1_\delta(\mathcal{K})$ *with* $\delta = \frac{\varepsilon^3}{2^{12} \cdot 3 \cdot d^{12}}$*.*

*Proof.* The proof follows directly from Lemma 6. Given the equivalence between the protocols $\mathsf{WMEM}_\varepsilon$ and $\rho$-$\mathsf{WMEM}_\varepsilon$, as well as between $\mathsf{WMEM}^1_\delta$ and $\rho$-$\mathsf{WMEM}^1_\delta$, we can construct a polynomial-time algorithm for $\rho$-$\mathsf{WMEM}_\varepsilon$ with oracle access to $\rho$-$\mathsf{WMEM}^1_\delta$. The approach proceeds by first reducing $\rho$-$\mathsf{WMEM}_\varepsilon$ to $\mathsf{WMEM}_\varepsilon$ , then applying Theorem 4 to complete the reduction. □

### 1. Weak membership for the stabilizer polytope

This section investigates the computational problem of determining membership in the convex hull of pure stabilizer states. Formally, letting $\mathcal{S}_n$ denote the set of pure stabilizer states on $n$ qubits and $\hat{\mathcal{S}}_n$ its convex hull, we analyse the complexity of the weak membership problem $\rho$-$\mathsf{WMEM}_\varepsilon(\hat{\mathcal{S}}_n)$. We begin by examining the hardness of the decision problem $\rho$-$\mathsf{WMEM}_\varepsilon(\hat{\mathcal{S}}_n)$.

**Theorem 6** (Superpolynomial complexity of $\rho$-$\mathsf{WMEM}_\varepsilon(\hat{\mathcal{S}}_n)$)**.** *Let* $\hat{\mathcal{S}}_n$ *denote the stabilizer polytope on* $n$ *qubits. Under the* exponential time hypothesis *(ETH) (see Definition 5), for any* $\varepsilon \leq \frac{1}{2^{13} 3^3 d^{53/2} \log^{18} d}$*, the decision problem* $\rho$-$\mathsf{WMEM}_\varepsilon(\hat{\mathcal{S}}_n)$ *is not contained in the complexity class* $\mathsf{QP}^{2-\eta}$ *for any* $\eta > 0$*.*

*Proof.* We prove that $\rho$-$\mathsf{WMEM}_\varepsilon(\hat{\mathcal{S}}_n)$ belongs to $\mathsf{QP}^2$ in Lemma 11. To show that this inclusion is strict, we reduce the membership problem to the weak witness detection problem $\mathsf{WWD}_\delta(\hat{\mathcal{S}}_n)$. In Theorem 7, we show that $\mathsf{WWD}_\delta(\hat{\mathcal{S}}_n)$ does not belong to $\mathsf{QP}^{2-\eta}$ for any $\eta > 0$, assuming the exponential time hypothesis stated in Definition 5, provided that $\delta \leq \frac{1}{d^{9/2} \log^6 d}$. The reduction follows from the existence of a polynomial-time reduction between $\rho$-$\mathsf{WMEM}_\varepsilon(\hat{\mathcal{S}}_n)$ and $\mathsf{WWD}_\delta$, established in Lemma 9. As a consequence, $\rho$-$\mathsf{WMEM}_\varepsilon(\hat{\mathcal{S}}_n)$ with $\varepsilon \leq \frac{1}{2^{13} 3^3 d^{53/2} \log^{18} d}$ requires time $2^{\Omega(\log^2 d)}$, which implies the desired lower bound. This concludes the proof. □

It should be remarked that while the promised gap $\varepsilon$ in Theorem 6 is small, it is "only" polynomially small in the input parameter $d$ and such a situation is common in conceptual hardness proofs of this kind. Original proofs stating the hardness of testing membership in the set of bipartite separable quantum states featured promise gaps that were even inversely exponential in the dimension of the quantum system and hence doubly exponential in the system size [35, 36]. When this was brought down to an inverse polynomial scaling, a situation reminiscent of the one encountered here, this was dubbed the "strong NP-hardness of the quantum separability problem" [37]. The underlying reason for this behavior lies in the structure of the reductions: rather than proving hardness directly, one relies on a chain of polynomial-time reductions between different computational problems. In our case, following the same strategy as in the proof of strong NP-hardness for quantum separability, we exploit the polynomial equivalence between the classes $\mathsf{WMEM}_\varepsilon$ and $\mathsf{WOPT}_\delta$. This equivalence, however, comes at a quantitative cost, as made explicit in Theorem 5, where one must fix $\varepsilon = O(\delta^3/d^{13})$. In principle, these constants could be improved in two ways: either by strengthening the polynomial-time reduction between $\mathsf{WMEM}_\varepsilon$ and $\mathsf{WOPT}_\delta$, or by shortening the overall reduction chain. For instance, a direct reduction showing that oracle access to $\rho$-$\mathsf{WMEM}_\varepsilon$ suffices to solve 3-SAT would lead to significantly improved parameter dependencies.

**Lemma 11** ($\rho$-$\mathsf{WMEM}_\varepsilon(\hat{\mathcal{S}}_n) \in \mathsf{QP}^2$)**.** *For any $0 < \varepsilon < 1$, there exists an algorithm running in time $\mathsf{poly}(2^{\log^2 d}, \log \frac{1}{\varepsilon})$ that solves $\rho$-$\mathsf{WMEM}_\varepsilon(\hat{\mathcal{S}}_n)$.*

*Proof.* As noted in Ref. [21], the optimisation for the robustness of magic (see Definition 14) can be written as the following linear program:

$$\mathcal{R}(\rho) = \max_{\mathbf{w}\,:\,\|A^T\mathbf{w}\|_\infty \le 1} (-\mathbf{x}^T \cdot \mathbf{w}) \tag{S2.11}$$

where $A_{i,j} = \frac{1}{\sqrt{d}} \operatorname{tr}(P_i \sigma_j)$ and $\mathbf{x}_i = \frac{1}{\sqrt{d}} \operatorname{tr}(P_i \rho)$. Eq. (S2.11) states that the dual problem associated with computing the robustness of magic is precisely the maximisation of a magic witness. Indeed, it can be recast as

$$\mathcal{R}(\rho) = \max_{\substack{W \\ \operatorname{tr}(W\sigma)\le 1\ \forall \sigma \in \mathcal{S}_n}} \operatorname{tr}(W\rho). \tag{S2.12}$$

Indeed, the condition $\|A^T\mathbf{w}\|_\infty \le 1$ corresponds to the condition $\max_{\sigma\in\mathcal{S}_n} \operatorname{tr}(W\sigma) \le 1$, which is exactly the definition of a witness for a convex set (see Definition 21 with $\gamma = 1$). We define the set of such witnesses $\mathcal{W} \coloneqq \{W : \max_{\sigma\in\mathcal{S}_n} \operatorname{tr}(\sigma W) \le 1\}$. So far we have implicitly assumed an ideal setting in which the full space of witnesses can be accessed exactly. We now relax this assumption and consider the more realistic case in which only a restricted subset of witnesses is available, due to finite numerical precision. Given that the set of witnesses is a convex set, we can define its inner and outer boundary according to Definition 6. It is immediate to see that we can write the inner boundary of $\mathcal{W}$ as

$$S(\mathcal{W}, -\nu) = \{\, W \in \mathcal{W} \;:\; W + H \in \mathcal{W} \text{ for all } H \text{ with } \|H\|_2 \le \nu \,\}. \tag{S2.13}$$

That is, $S(\mathcal{W}, -\nu)$ consists of witnesses that remain feasible under any Hilbert–Schmidt perturbation of norm at most $\nu$. For the reminder of the proof, we adopt the lighter notation $\mathcal{W}_\nu \equiv S(\mathcal{W}, -\nu)$. The corresponding optimisation problem is

$$\mathcal{R}_\nu(\rho) \coloneqq \max_{W\in\mathcal{W}_\nu} \operatorname{tr}(\rho W), \tag{S2.14}$$

which is a linear program over $\mathcal{W}_\nu$. This problem can be solved up to additive error $\nu$ using the ellipsoid algorithm [65] in time $\mathsf{poly}\big(2^{\log^2 d}, \log \nu^{-1}\big)$ provided that $\mathcal{W}_\nu$ is nonempty. Our goal is to show that the linear program associated with the robustness of magic can be used to solve the one-sided problem $\rho$-$\mathsf{WMEM}^1_\delta(\hat{\mathcal{S}}_n)$ (see Problem 6). The full weak membership problem $\mathsf{WMEM}_\varepsilon(\hat{\mathcal{S}}_n)$ then follows by invoking Lemma 10.

We begin with the YES case. Let $\rho \in \hat{\mathcal{S}}_n$. By construction, any admissible witness $W \in \mathcal{W}_\nu$ satisfies $\operatorname{tr}(W\rho) \le 1$. Since the linear program is solved only up to additive error $\nu$, the algorithm may return a value of at most $1+\nu$. Therefore, no YES instance can be misclassified as NO, provided that for every $\rho \notin S(\hat{\mathcal{S}}_n, \delta)$ the value returned by the algorithm is strictly larger than $1 + \nu$.

We now turn to the NO case. Let $\rho$ be a state such that $\min_{\sigma\in\hat{\mathcal{S}}_n} \|\rho - \sigma\|_2 \ge \delta$ . Let us show that there exists a witness $W$ such that $\operatorname{tr}(W\rho) > 1 + \nu$. By Lemma 7, there exists a Hermitian operator $\tilde{W}$ satisfying

$$\operatorname{tr}\Big(\tilde{W}\rho\Big) > \max_{\sigma\in\hat{\mathcal{S}}_n} \operatorname{tr}\Big(\tilde{W}\sigma\Big) + \delta^2. \tag{S2.15}$$

We now show how to rescale $\tilde{W}$ into a valid element of $\mathcal{W}_\nu$. Recall that for $W \in \mathcal{W}_\nu$ the constraint $\operatorname{tr}[(W + H)\sigma] \le 1$ for all $\sigma \in \hat{\mathcal{S}}_n$ implies, by linearity and Hölder's inequality, that $\operatorname{tr}(W\sigma) \le 1 - \nu$. Define

$$W \coloneqq \frac{\tilde{W}}{\max_{\sigma\in\hat{\mathcal{S}}_n} \operatorname{tr}\Big(\tilde{W}\sigma\Big)} - \nu I. \tag{S2.16}$$

By construction, $W \in \mathcal{W}_\nu$. Evaluating $W$ on $\rho$ yields

$$\begin{aligned} \mathrm{tr}(W\rho) &= \frac{\mathrm{tr}\left(\tilde{W}\rho\right)}{\max_{\sigma\in\hat{\mathcal{S}}_n} \mathrm{tr}\left(\tilde{W}\sigma\right)} - \nu \\ &> 1 + \frac{\delta^2}{\max_{\sigma\in\hat{\mathcal{S}}_n} \mathrm{tr}\left(\tilde{W}\sigma\right)} - \nu. \end{aligned} \tag{S2.17}$$

In order to distinguish YES from NO instances, we require $\mathrm{tr}(W\rho) > 1 + \nu$, which is ensured if

$$\frac{\delta^2}{\max_{\sigma\in\hat{\mathcal{S}}_n} \mathrm{tr}\left(\tilde{W}\sigma\right)} > 3\nu, \tag{S2.18}$$

where we are accounting for the additive error $\nu$. Using the bound

$$\max_{\sigma\in\hat{\mathcal{S}}_n} \mathrm{tr}\left(\tilde{W}\sigma\right) \leq \|\tilde{W}\|_2 = \|\rho - \tau\|_2 \leq 2, \tag{S2.19}$$

we obtain the sufficient condition $\nu \leq \delta^2/6$.

Finally, note that the explicit construction above also shows that the feasible set $\mathcal{W}_\nu$ is nonempty, completing the implications that computing the robustness offers a way to perform 1-sided $\rho$-$\mathsf{WMEM}^1_\delta(\hat{\mathcal{S}}_n)$. By Lemma 6, we know that $\rho$-$\mathsf{WMEM}^1_\delta(\hat{\mathcal{S}}_n)$ and $\rho$-$\mathsf{WMEM}_\varepsilon(\hat{\mathcal{S}}_n)$ are polynomially equivalent, as long $\delta = \frac{\varepsilon^3}{2^{12}\cdot 3\cdot d^{12}}$, meaning that in order to solve $\rho$-$\mathsf{WMEM}_\varepsilon$, $\nu \leq \frac{\varepsilon^6}{2^{25}\cdot 3^3\cdot d^{24}}$, implying the desired runtime for the algorithm.

□

### 2. Weak witness detection

In this section, we prove one of the main results of the paper, namely that determining whether a Hermitian operator $W$ functions as a witness for the stabilizer polytope does not belong to $\mathsf{QP}^{2-\eta}$ for any $\eta > 0$ under the ETH assumption Definition 5.

**Theorem 7** (Superpolynomial complexity of $\mathsf{WWD}_\delta(\hat{\mathcal{S}}_n)$)**.** *Let $\hat{\mathcal{S}}_n$ denote the stabilizer polytope on $n$ qubits. Under the* exponential time hypothesis *(ETH) (see Definition 5), for any $\delta \leq \frac{1}{d^{9/2}\log^6 d}$, the decision problem $\mathsf{WWD}_\delta(\hat{\mathcal{S}}_n)$ is not contained in the complexity class $\mathsf{QP}^{2-\eta}$ for any $\eta > 0$.*

*Proof.* The overall strategy for proving the theorem is to reduce the weak-witness detection problem to another decision problem, whose solution requires resolving a general 3-SAT instance on $O(\log^2 d)$ variables. Under the ETH (see Definition 5), this requires time $2^{\Omega(\log^2 d)}$. Since the input of the problem is a $d \times d$ density matrix, this implies that $\mathsf{WWD}_\delta(\hat{\mathcal{S}}_n)$ does not lie in $\mathsf{QP}^{2-\eta}$ for any $\eta > 0$, see Definition 4. In particular, we construct a Hamiltonian $H$ acting on $\log^2 d$ qubits of the type in Eq. (S1.5).

The proof is technically involved. Its key ingredient is the immediate observation, established below, that optimising a Hamiltonian over two copies of graph states already requires solving a 3-SAT problem with $O(\log^2 d)$ variables. The remainder of the proof introduces suitable penalty terms that encode this optimisation into the problem of optimising a Hamiltonian over a single stabilizer state, thereby completing the argument.

**Local Hamiltonian.** This part of the proof constructs an instance of the local Hamiltonian problem. The key insight exploits the structure of stabilizer states, specifically focusing on graph states [25, 26]. Recall that a graph state vector $|G(A)\rangle$ on $n$ qubits, associated with an adjacency matrix $A$, is defined as

$$|G(A)\rangle \coloneqq \prod_{(i,j):i<j} CZ_{i,j}^{A_{i,j}} |+\rangle^{\otimes n} = \frac{1}{\sqrt{d}} \sum_{\mathbf{x}\in\{0,1\}^n} (-1)^{\frac{1}{2}\mathbf{x}^T A\mathbf{x}} |\mathbf{x}\rangle . \tag{S2.20}$$

For each pair $(i,j)$, let $s_{ij} \in \{0,1\}$ and define the operators

$$\begin{aligned} X^{s_{ij}}(i,j) &\coloneqq d/4 \left(|\mathbf{0}\rangle + (-1)^{s_{ij}} |i,j\rangle\right)\left(\langle\mathbf{0}| + (-1)^{s_{ij}} \langle i,j|\right), \\ Y^{s_{ij},s_{kl}}(i,j,k,l) &\coloneqq d/4 \left(|\mathbf{0}\rangle + (-1)^{s_{ij}} |i,j\rangle\right)\left(\langle\mathbf{0}| + (-1)^{s_{kl}} \langle k,l|\right), \end{aligned} \tag{S2.21}$$

where $|i,j\rangle$ denotes the computational basis state of Hamming weight 2 with ones at positions $i$ and $j$. Their expectation values on the graph state state vector $|G(A)\rangle$ are given by

$$\begin{aligned}\langle G(A)|X^{s_{ij}}(i,j)|G(A)\rangle &= \frac{\left(1+(-1)^{s_{ij}}(-1)^{A_{i,j}}\right)^2}{4},\\ \langle G(A)|Y^{s_{ij},s_{kl}}(i,j,k,l)|G(A)\rangle &= \frac{\left(1+(-1)^{s_{ij}}(-1)^{A_{i,j}}\right)\left(1+(-1)^{s_{kl}}(-1)^{A_{k,l}}\right)}{4}.\end{aligned} \tag{S2.22}$$

Since there is a one-to-one correspondence between the operators $X(i,j)$ and the matrix entries $A_{i,j}$, it suffices to consider $X(i,j)$ with $j>i$. Similarly, for $Y(i,j,k,l)$, we can restrict to $i<j<k<l$. We, therefore, relabel them as

$$\{X_\alpha^{s_\alpha}\}_{\alpha=1}^{n(n-1)/2} \quad \text{and} \quad \{Y_{\beta\gamma}^{s_\beta,s_\gamma}\}_{\beta<\gamma}^{n(n-1)/2}, \tag{S2.23}$$

with each $\alpha,\beta,\gamma$ corresponding to a specific pair $(i,j)$ and $s_\alpha,s_\beta,s_\gamma\in\{0,1\}$ denote the sign of the operator. Now consider the Hamiltonian

$$H=\sum_{\alpha<\beta<\gamma} X_\alpha^{s_\alpha}\otimes\frac{(Y_{\beta,\gamma}^{s_\beta,s_\gamma}+Y_{\beta,\gamma}^{s_\beta,s_\gamma\dagger})}{2} \tag{S2.24}$$

acting on two copies of the Hilbert space $\mathcal{H}^{\otimes 2}$. Its expectation value on two copies of the graph state is

$$\begin{aligned}\langle G(A)|^{\otimes 2}H|G(A)\rangle^{\otimes 2} &= \sum_{\alpha<\beta<\gamma}\left(\frac{1+(-1)^{A_\alpha+s_\alpha}}{2}\right)^2\left(\frac{1+(-1)^{A_\beta+s_\beta}}{2}\right)\left(\frac{1+(-1)^{A_\gamma+s_\gamma}}{2}\right)\\ &= \sum_{\alpha<\beta<\gamma}(x_\alpha\oplus s_\alpha)(x_\beta\oplus s_\beta)(x_\gamma\oplus s_\gamma),\end{aligned}$$

where $x_\alpha^2=x_\alpha$ and

$$x_\alpha \coloneqq \left(\frac{1+(-1)^{A_\alpha}}{2}\right)\in\{0,1\}. \tag{S2.25}$$

Moreover $\oplus$ denotes the binary sum modulo 2. We have successfully encoded a generic 3-SAT instance (see Eq. (S1.5)) in the expectation values of graph states. Let us now define $\mathcal{G}_n^{(2)}\coloneqq\{|G(A)\rangle^{\otimes 2}: A\in\mathbb{F}_2^{\binom{n}{2}\times\binom{n}{2}}\}$ as the set of identical bipartite graph states on $2n$ qubits. Consider a Hamiltonian $H$ as in Eq. (S2.24), with arbitrary $\alpha,\beta,\gamma\in[n(n-1)/2]$. We consider the following decision problem:

- Output YES, if $\max_{|\psi\rangle\in\mathcal{G}_n^{(2)}}\langle\psi|H|\psi\rangle=0$.
- Output NO, if $\max_{|\psi\rangle\in\mathcal{G}_n^{(2)}}\langle\psi|H|\psi\rangle\geq 1$.

This problem is at least as hard as solving a 3-SAT instance with $n(n-1)/2=\Omega(\log^2 d)$ variables. Assuming the ETH (Definition 5) again, such an instance requires time $2^{\Omega(\log^2 d)}$, and therefore the above decision problem cannot belong to the class $\mathsf{QP}^{2-\eta}$ for any $\eta>0$.

**Extension to single copy graph state.** The previous step alone does not suffice to establish the hardness of $\mathsf{WWD}(\hat{\mathcal{S}}_n)$, as it only demonstrates hardness for 2-copy graph states. The next step, therefore, shows that the problem remains hard even for single-copy graph states by reducing the 2-copy case to the single-copy case. The key idea is to extend the original Hamiltonian $H$ to a new operator $H^{(1)}$ such that minimising its expectation value over single-copy graph states is at least as hard as minimising $H$ over the set $\mathcal{G}_n^{(2)}$. To proceed, let us notice that the set

$$\mathcal{G}_n^{(2)}\subset\{|G(A)\rangle\ :\ A\in\mathbb{F}_2^{\binom{2n}{2}\times\binom{2n}{2}}\}=:\mathcal{G}_{2n}. \tag{S2.26}$$

Hence, our strategy is to force, through a penalty Hamiltonian, the low energy states to lie in $\mathcal{G}_n^{(2)}$. Let us label the entries of the adjacency matrix as $i_A$ for the first $n$ columns and $i_B$ for the last $n$ columns, with the same convention applied to rows. At this point, we define the operator

$$\bar{X}(i,j)\coloneqq\frac{d^2}{4}(|\mathbf{0}\rangle-|i,j\rangle)(\langle\mathbf{0}|-\langle i,j|). \tag{S2.27}$$

which coincides with $X^1$ in Eq. (S2.21). Its expectation value over $|G(A)\rangle\in\mathcal{G}_{2n}$ is given by

$$\langle G(A)|\bar{X}(i_A,j_B)|G(A)\rangle=\left(\frac{1-(-1)^{A_{i_A,j_B}}}{2}\right), \tag{S2.28}$$

where we observe that this expectation value vanishes if and only if $A_{i_A,j_B} = 0$. We further introduce the operator

$$W(i_A, j_A, k_B, l_B) \coloneqq d^2/4(|\mathbf{0}\rangle - |i_A, j_A\rangle \otimes |k_B, l_B\rangle)(\langle \mathbf{0}| - \langle i_A, j_A| \otimes \langle k_B, l_B|), \tag{S2.29}$$

whose expectation value over $|G(A)\rangle \in \mathcal{G}_{2n}$ reads

$$\langle G(A)|W(i_A, j_A, k_B, l_B)|G(A)\rangle = \frac{1 - (-1)^{A_{i_A,j_A} + A_{k_B,l_B}}}{2}, \tag{S2.30}$$

which vanishes if and only if $A_{i_A,j_A} = A_{k_B,l_B}$.

It follows that incorporating these terms into the Hamiltonian preserves the minimum obtained on two copies of the graph states while penalising all other states in the minimisation problem. We, therefore, consider the following Hamiltonian

$$H^{(1)} = H + (\|H\|_\infty + 1)\left[\sum_{i,j} \bar{X}(i_A, j_B) + \sum_{i \neq j} W(i_A, j_A, i_B, j_B)\right]. \tag{S2.31}$$

By construction, we have

$$\langle G(A)|H^{(1)}|G(A)\rangle = \langle G(A)|H|G(A)\rangle \quad \forall\, |G(A)\rangle \in \mathcal{G}_n^{(2)} \tag{S2.32}$$

since the added positive operator vanishes if and only if the graph on $2n$ vertices satisfies two conditions: first, there are no edges between the first $n$ vertices (labeled by $A$) and the last $n$ vertices (labeled by $B$); second, the connection pattern among the $A$ vertices exactly matches that among the $B$ vertices.

Let us now show that optimising the energy of $H^{(1)}$ over $\mathcal{G}_{2n}$ is at least as hard as optimising $H$ over $\mathcal{G}_n^{(2)}$.

- YES. Assume there exists a graph state vector $|G(A)\rangle \in \mathcal{G}_n^{(2)}$ such that $\langle G(A)|H|G(A)\rangle = 0$, then, by the argument above, we also have $\langle G(A)|H^{(1)}|G(A)\rangle = 0$.
- NO. Assume that $\min_{|G(A)\rangle \in \mathcal{G}_n^{(2)}} \langle G(A)|H|G(A)\rangle \geq 1$ (i.e., the second-to-last lowest value for $H$ over $\mathcal{G}_n^{(2)}$), then for any $|G(A)\rangle \in \mathcal{G}_{2n} \setminus \mathcal{G}_n^{(2)}$, we have

  $$\langle G(A)|H^{(1)}|G(A)\rangle \geq (\|H\|_\infty + 1) - |\langle G(A)|H|G(A)\rangle| \geq 1\,, \tag{S2.33}$$

  where we have exploited the fact that the expectation value of $H$ is strictly positive (and greater than 1) for graph states $\notin \mathcal{G}_n^{(2)}$. This proves the NO instance.

**Extension to maximally coherent stabilizer states.** At this point, we have established that optimising a Hamiltonian over the set of graph states on $2n$ qubits reduces to solve a 3-SAT with $n(n-1)/2$ variables. We now demonstrate that Hamiltonian optimisation over the set of *maximally coherent stabilizer states* in the computational basis is at least as hard as the corresponding problem for graph states. The set of maximally coherent stabilizer states in the computational basis is formally defined as

$$\mathcal{C}_n \coloneqq \left\{ \bigotimes_{i=1}^n U_i\, |G(A)\rangle : |G(A)\rangle \in \mathcal{G}_n,\, U_i \in \{Z, S, SZ\} \right\}, \tag{S2.34}$$

meaning that each maximally coherent state is locally unitarily equivalent to a graph state under diagonal local unitaries. These unitaries $U_i$ are restricted to the Pauli-$Z$ basis and include the phase gate $S$ and its composition with $Z$. The action of such a diagonal unitary $U_i = \mathrm{diag}(1, \phi_i)$ with $|\phi_i| = 1$ transforms a graph state as follows

$$\bigotimes_i U_i\, |G(A)\rangle = \frac{1}{\sqrt{d}} \sum_x (-1)^{\sum_{i,j} A_{i,j} x_i x_j} \prod_{i=1}^n \phi_i^{x_i}\, |x\rangle\,. \tag{S2.35}$$

Here, each basis vector $|x\rangle$ acquires an additional phase factor $\prod_i \phi_i^{x_i}$ that depends on the unitary operations applied.

To analyse the effect of these unitaries, consider the operator

$$S(i) \coloneqq \frac{d^2}{4}(|\mathbf{0}\rangle - |i\rangle)(\langle \mathbf{0}| - \langle i|), \tag{S2.36}$$

where $|i\rangle$ denotes the basis state with a single 1 at position $i$. The expectation value of $S(i)$ over a transformed state $\bigotimes_i U_i\, |G(A)\rangle$ evaluates to $|1 - \phi_i|^2/4$. Crucially, this expectation value attains its minimum value of 0 if and only if

$\phi_i = 1$, which corresponds to the case where no nontrivial diagonal unitary is applied to the $i$-th qubit. This property will be essential for our hardness reduction, as it allows us to detect and penalise deviations from the original graph state structure. We, therefore, extend the Hamiltonian as

$$H^{(2)} = H^{(1)} + 2(\|H\|_\infty + 1)\sum_{i=1}^{2n} S(i). \tag{S2.37}$$

By construction, we have $\langle\sigma|H^{(2)}|\sigma\rangle = \langle\sigma|H^{(1)}|\sigma\rangle$ for all $|\sigma\rangle \in \mathcal{G}_{2n}$, since $\mathcal{G}_{2n} \subseteq \mathcal{C}_{2n}$ and the operators $S(i)$ vanish when acting on states $\mathcal{G}_{2n}$. We now demonstrate that optimising $H^{(2)}$ over the set $\mathcal{C}_{2n}$ is at least as hard as optimising $H^{(1)}$ over $\mathcal{G}_{2n}$, thereby establishing that the former problem requires time $2^{\Omega(n^2)}$.

- YES. Assume that $\min_{|\sigma\rangle\in\mathcal{G}_{2n}}\langle\sigma|H^{(1)}|\sigma\rangle = 0$; then, by the argument above, we also have $\min_{|\sigma\rangle\in\mathcal{C}_{2n}}\langle\sigma|H^{(2)}|\sigma\rangle = 0$.

- NO. Assume that $\min_{|\sigma\rangle\in\mathcal{G}_{2n}}\langle\sigma|H^{(1)}|\sigma\rangle \geq 1$, then for any $|\sigma\rangle \in \mathcal{C}_{2n} \setminus \mathcal{G}_{2n}$, we have

$$\begin{aligned}
\langle\sigma|H^{(2)}|\sigma\rangle &\geq 2(\|H\|_\infty + 1)\frac{1}{2} + \langle\sigma|H|\sigma\rangle + \langle\sigma|(H^{(1)} - H)|\sigma\rangle && \text{(S2.38)}\\
&\geq (\|H\|_\infty + 1) - |\langle\sigma|H|\sigma\rangle| && \text{(S2.39)}\\
&\geq (\|H\|_\infty + 1) - \|H\|_\infty && \text{(S2.40)}\\
&= 1, && \text{(S2.41)}
\end{aligned}$$

  where we have exploited the fact that $\phi \in \{\pm i, \pm i\}$ for diagonal Clifford operators. Moreover, we noticed that the operator $H^{(1)} - H$ is a positive operator.

**Extension to stabilizer states with exponentially small overlap with $|\mathbf{0}\rangle$.** In the previous step, we showed that minimising $H$ over the set of maximally coherent stabilizer states requires exponential time. We now define the following set:

$$\mathcal{T}_n \coloneqq \left\{|\sigma\rangle \in \mathcal{S}_n : |\langle\sigma|\mathbf{0}\rangle|^2 \geq \frac{1}{d}\right\}, \tag{S2.42}$$

which consists of stabilizer states whose overlap with the $|\mathbf{0}\rangle$ state is at least $1/d$. Since $\mathcal{C}_{2n} \subseteq \mathcal{T}_{2n}$, we aim—as in previous steps—to penalise states outside this set, thereby preserving the hardness of our solution. Let us define the Hamiltonian

$$H^{(3)} \coloneqq H^{(2)} + d^2(\|H\|_\infty + 1)\left(|\mathbf{0}\rangle\langle\mathbf{0}| - \frac{I}{d^2}\right), \tag{S2.43}$$

where $|\mathbf{0}\rangle$ is defined on $2n$ qubits, and show that minimising $H^{(3)}$ over $\mathcal{T}_{2n}$ is at least as hard as optimising $H^{(2)}$ over $\mathcal{C}_{2n}$. The strategy is the same; analysing the YES and NO instances, we have the following:

- YES. Assume that $\min_{|\sigma\rangle\in\mathcal{C}_{2n}}\langle\sigma|H^{(2)}|\sigma\rangle = 0$. Since the $\operatorname{argmin}|\sigma\rangle$ is maximally coherent, it holds that $|\langle\sigma|\mathbf{0}\rangle|^2 = 1/d^2$ and therefore $\langle\sigma|H^{(3)}|\sigma\rangle = 0$.

- NO. Assume that $\min_{\sigma\in\mathcal{C}_{2n}}\langle\sigma|H^{(2)}|\sigma\rangle \geq 1$. Then, for any $|\sigma\rangle \in \mathcal{T}_{2n} \setminus \mathcal{C}_{2n}$, we have

$$\begin{aligned}
\langle\sigma|H^{(3)}|\sigma\rangle &\geq d^2(\|H\|_\infty + 1)\left(|\langle\sigma|\mathbf{0}\rangle|^2 - \frac{1}{d^2}\right) + \langle\sigma|H^{(3)} - H|\sigma\rangle - |\langle\sigma|H|\sigma\rangle| && \text{(S2.44)}\\
&\geq \frac{d^2(\|H\|_\infty + 1)}{d^2} - \|H\|_\infty \geq 1. && \text{(S2.45)}
\end{aligned}$$

  Above, we have used the fact that if $|\sigma\rangle$ is not maximally coherent, then the overlap with $|\mathbf{0}\rangle$ must be $\geq 2/d^2$. We also used the fact that, similarly to above, the operator $H^{(2)} - H$ is a positive operator.

  Hence, optimising $H^{(3)}$ over the set $\mathcal{T}_{2n}$ is at least as hard as optimising $H^{(2)}$ on $\mathcal{C}_{2n}$ and therefore requires solving a 3-SAT with $n(n-1)/2$ variables.

**Extension to stabilizer states.** We now demonstrate that determining the eigenenergy difference of a Hamiltonian over the set of stabilizer states $\mathcal{S}_{2n}$ is at least as hard as optimising it over $\mathcal{T}_{2n}$. The key distinction between these sets is that states in $\mathcal{T}_{2n}$ maintain a non-vanishing overlap with $|\mathbf{0}\rangle$, whereas $\mathcal{S}_{2n}$ includes states orthogonal to $|\mathbf{0}\rangle$. To establish this relationship, we consider the modified Hamiltonian

$$H^{(4)} = H^{(3)} + (\|H\|_\infty + \|H^{(3)} - H^{(2)}\|_\infty + 1)\left[I - d^2\,|\mathbf{0}\rangle\langle\mathbf{0}|\right], \tag{S2.46}$$

and define the following decision problem for $H^{(4)}$:

- $\min_{|\sigma\rangle\in\mathcal{S}_{2n}}\langle\sigma|H|\sigma\rangle \leq 0$ output YES.
- $\min_{|\sigma\rangle\in\mathcal{S}_{2n}}\langle\sigma|H|\sigma\rangle \geq 1$ output NO.

Let us show, in a similar fashion to what was done before, that solving the above decision problem for $H^{(4)}$ is at least as hard as optimising $H^{(3)}$ on $\mathcal{T}_{2n}$.

- YES. Assume $\min_{|\sigma\rangle\in\mathcal{T}_{2n}}\langle\sigma|H^{(3)}|\sigma\rangle = 0$. Let $|\sigma\rangle$ be the argmin. Then we have

  $$\langle\sigma|H^{(4)}|\sigma\rangle = (\|H\|_\infty + \|H^{(3)} - H^{(2)}\|_\infty + 1)(1 - d^2|\langle\mathbf{0}|\tau\rangle|^2) \leq 0, \tag{S2.47}$$

  as an immediate and trivial consequence of the definition of $\mathcal{T}_{2n}$.

- NO. Assume $\min_{|\sigma\rangle\in\mathcal{T}_{2n}}\langle\sigma|H^{(3)}|\sigma\rangle \geq 1$. For any $|\sigma\rangle \in \mathcal{S}_{2n} \setminus \mathcal{T}_{2n}$ we have

  $$\langle\sigma|H^{(4)}|\sigma\rangle \geq (\|H\|_\infty + \|H^{(3)} - H^{(2)}\|_\infty + 1) + \langle\sigma|H^{(2)} - H|\sigma\rangle - |\langle\sigma|H + (H^{(3)} - H^{(2)})|\sigma\rangle| \geq 1, \tag{S2.48}$$

  which shows the claim. Notice that we decomposed $H^{(3)} = H + (H^{(2)} - H) + (H^{(3)} - H^{(2)})$ and noticed that $(H^{(2)} - H) \geq 0$.

**Reduction to $\mathsf{WWD}_\delta$.** To complete the proof, we reduce our problem to an instance of WWD as defined in Problem 7. Let $\hat{\mathcal{S}}_n$ denote the convex hull of stabilizer states $\mathcal{S}_n$, with $S(\hat{\mathcal{S}}_{2n}, -\delta)$ and $S(\hat{\mathcal{S}}_{2n}, \delta)$ representing its inner and outer approximations, respectively. Our WWD instance involves a Hermitian operator $W$ with $\|W\|_2 \leq 1$, where the decision problem is the following:

- If there exists $\rho \in S(\hat{\mathcal{S}}_{2n}, -\delta)$ s.t. $\mathrm{tr}(W\rho) \geq \gamma + \delta$, then output YES.
- If $\forall\rho \in S(\hat{\mathcal{S}}_{2n}, \delta)$ holds that $\mathrm{tr}(W\rho) \leq \gamma - \delta$, then output NO.

We now establish the connection to our Hamiltonian problem by defining

$$W := -\frac{H^{(4)}}{\|H^{(4)}\|_2}. \tag{S2.49}$$

For the YES case, when $\max_{\sigma\in\hat{\mathcal{S}}_{2n}}\mathrm{tr}(W\sigma) \geq 0$, there exists $\rho \in S(\hat{\mathcal{S}}_{2n}, -\delta)$ satisfying $\mathrm{tr}(W\rho) \geq \gamma + \delta$. Conversely, if $\max_{\sigma\in\hat{\mathcal{S}}_{2n}}\mathrm{tr}(W\sigma) \leq -\frac{1}{\|H^{(4)}\|_2}$, then all $\rho \in S(\hat{\mathcal{S}}_{2n}, -\delta)$ satisfy $\mathrm{tr}(W\rho) \leq \gamma - \delta$. Crucially, optimisation over the convex hull $\hat{\mathcal{S}}_{2n}$ reduces to optimisation over its extreme points $\mathcal{S}_{2n}$, meaning $\max_{\sigma\in\hat{\mathcal{S}}_{2n}}\mathrm{tr}(W\sigma) = \max_{\sigma\in\mathcal{S}_{2n}}\langle\sigma|W|\sigma\rangle$. This reduction demonstrates that the Hamiltonian decision problem maps directly to $\mathsf{WWD}_\delta(\hat{\mathcal{S}}_n)$, thereby establishing its hardness. To show it, set $\gamma := -\frac{1}{2\|H^{(4)}\|_2}$.

- YES. Assume $\max_{\sigma\in\mathcal{S}_{2n}}\mathrm{tr}(W\sigma) \geq 0$. Since $\hat{\mathcal{S}}_{2n}$ is a convex set, there exists $\sigma_\varepsilon \in S(\hat{\mathcal{S}}_{2n}, -\delta)$ such that $\|\sigma - \sigma_\delta\|_2 \leq 2\delta R/r$ [79]. We recall that $R \leq 1$, $r \geq 1/d^2$ , because $H^{(4)}$ is defined on $2n$ qubits. This means that

  $$|\,\mathrm{tr}(W\sigma_\delta) - \mathrm{tr}(W\sigma)| \leq \|\sigma_\delta - \sigma\|_2 \leq \delta d^2, \tag{S2.50}$$

  implying that $\mathrm{tr}(W\sigma_\delta) \geq \mathrm{tr}(W\sigma) - \delta d^2 \geq -\delta d^2$, using that $\mathrm{tr}(W\sigma) \geq 0$. Imposing the latter, being $\geq \gamma + \delta$, we find $\delta \leq \frac{1}{(1+d^2)\|H^{(4)}\|_2}$.

- NO. Assume that $\max_{\sigma\in\mathcal{S}_{2n}}\mathrm{tr}(W\sigma) \leq -\frac{1}{\|H^{(4)}\|_2}$. By the definition of $S(\hat{\mathcal{S}}_{2n}, \delta)$, for any $\rho \in S(\hat{\mathcal{S}}_{2n}, \delta)$, there exists $\sigma \in \mathcal{S}_{2n}$ such that $\|\rho - \sigma\|_2 \leq \delta$, implying that $|\,\mathrm{tr}(W\rho) - \mathrm{tr}(W\sigma)| \leq \delta$. It follows that, for any $\rho \in S(\hat{\mathcal{S}}_{2n}, \delta)$

  $$\mathrm{tr}(W\rho) \leq \mathrm{tr}(W\sigma) + \delta \leq \max_{\sigma\in\mathcal{S}_{2n}}\langle\sigma|W|\sigma\rangle + \delta \leq -\frac{1}{\|H^{(4)}\|_2} + \delta = 2\gamma + \delta. \tag{S2.51}$$

  Imposing the latter $\leq \gamma - \delta$, we find $\delta \leq -\gamma/2 = \frac{1}{4\|H^{(4)}\|_2}$.

To accommodate both YES and NO instances, we take $\delta \leq \frac{1}{4d^2\|H^{(4)}\|_2}$. Therefore, we have shown that with oracle access to $\mathsf{WWD}_\delta(\hat{\mathcal{S}}_{2n})$ with parameters $\left(W \equiv H^{(4)}/\|H^{(4)}\|_2, \gamma \equiv -\frac{1}{2\|H^{(2)}\|_2}\right)$ and any $\delta \leq \frac{1}{4d^2\|H^{(4)}\|_2}$, it is possible to solve the optimisation of a Hamiltonian over the set of pure stabilizer states $\mathcal{S}_{2n}$, which we have demonstrated to be at least as hard as a 3-SAT with input size $n(n-1)/2 = \Omega(\log^2 d)$. We are just left to find an upper bound to $\|H^{(4)}\|_2$, which is what we will turn to now.

**Upper bound on $\|H^{(4)}\|_2$.** We then have the following chain of inequalities

$$\begin{aligned}
\|H^{(4)}\|_2 &\leq d\|H^{(4)}\|_\infty \\
&\leq d\Big(\|H^{(3)}\|_\infty + d^2(\|H\|_\infty + \|H^{(3)} - H^{(2)}\|_\infty + 1)\Big) \\
&\leq d\Big[\|H^{(3)}\|_\infty + d^2(d^2+1)(\|H\|_\infty + 1)\Big] \\
&\leq d\Big[\|H^{(2)}\|_\infty + d^2(d^2+2)(\|H\|_\infty + 1)\Big] \\
&\leq d\Big[\|H^{(1)}\|_\infty + 2d^2(\|H\|_\infty + 1)\log d + d^2(d^2+2)(\|H\|_\infty + 1)\Big] \\
&\leq d\big[\|H\|_\infty + 2d^2(\|H\|_\infty + 1)\log^2 d + d^2(d^2 + 4\log d)(\|H\|_\infty + 1)\big] \\
&\leq d(\|H\|_\infty + 1)(d^2(d^2 + 6\log^2 d)) \leq 2^4 d^7 \log^6 d\,,
\end{aligned} \tag{S2.52}$$

where we have used the definitions of $H^{(4)}, H^{(3)}, H^{(2)}, H^{(1)}$ and $H$, and the triangle inequality. Consequently, we obtain that $\|H^{(4)}\|_2 \leq 2^4 d^7 \log^6 d$. Therefore, any $\delta \leq \frac{1}{2^6 d^9 \log^6 d}$ suffices to establish the computational hardness of $\mathsf{WWD}_\delta(\hat{\mathcal{S}}_{2n})$. To conclude, we are just left to map $d^2 \mapsto d$. □

Moreover, it is simple to see that $\mathsf{WWD}(\hat{\mathcal{S}}_n)$ is contained in the complexity class $\mathsf{QP}^2$, as the next lemma points out.

**Lemma 12** ($\mathsf{WWD}_\delta(\hat{\mathcal{S}}_n) \in \mathsf{QP}^2$)**.** *The decision problem $\mathsf{WWD}_\delta(\hat{\mathcal{S}}_n)$ belongs to $\mathsf{QP}^2$.*

*Proof.* Given an operator $W$ with the threshold parameter $\gamma$ and the precision $\delta$, to solve $\mathsf{WWD}_\delta(\hat{\mathcal{S}}_n)$ is sufficient to check the value of $\mathrm{tr}(W\sigma)$ for any $\sigma \in \mathcal{S}_n$ (see Problem 7) up to precision $\delta$. This algorithm takes $\mathsf{poly}(2^{\log^2 d}, \log \delta^{-1})$. Therefore for any $\delta$ reasonably large the problem can be solved within the complexity class $\mathsf{QP}^2$. □

### 3. Weak membership in the convex hull of $t$-doped stabilizer states

In this section, we extend the results obtained regarding the membership of stabilizer states to the more general notion of $t$-doped stabilizer states. Let us first start by examining the hardness of the decision problem $\rho$-$\mathsf{WMEM}_\varepsilon(\hat{\mathcal{S}}_{n,t})$. To do so, we first prove that determining whether a Hermitian operator $W$ functions as a witness for the $t$-doped stabilizer polytope does not belong to $\mathsf{QP}^{2-\eta}$ for any $\eta > 0$ under the ETH assumption (see Definition 5), and for any $t \leq \kappa n$ with $\kappa < 1$.

**Lemma 13** (Superpolynomial complexity of $\mathsf{WWD}_\delta(\hat{\mathcal{S}}_{n,t})$)**.** *Let $\hat{\mathcal{S}}_{n,t}$ denote the $t$-doped stabilizer polytope on $n$ qubits and let $t \leq \kappa n$ with $\kappa < 1$. Under the ETH (Definition 5), for any*

$$\delta \leq \frac{1}{2d^{31/2}\log^{30} d}, \tag{S2.53}$$

*the decision problem $\mathsf{WWD}_\delta(\hat{\mathcal{S}}_{n,t})$ is not contained in the complexity class $\mathsf{QP}^{2-\eta}$ for any $\eta > 0$.*

*Proof.* The overall strategy is to reduce the problem of optimising a Hamiltonian over the $t$-doped stabilizer polytope $\hat{\mathcal{S}}_{n,t}$ to an instance of weak-witness detection. The reduction allows us to embed a decision problem that, under the ETH, requires time $2^{\Omega(\log^2 d)}$ into $\mathsf{WWD}_\delta(\hat{\mathcal{S}}_{n,t})$ while preserving a non-vanishing promise gap. Since the input is a $d \times d$ density matrix with $d = 2^n$, this implies that the problem cannot lie in $\mathsf{QP}^{2-\eta}$ for any $\eta > 0$, see Definition 4.

By Theorem 7, there exists a family of Hamiltonians $H$ acting on $n-t$ qubits such that the following promise problem is at least as hard as solving a 3-SAT instance on $\Omega(\log^2 2^{n-t})$ variables:

- YES. $\max_{\sigma \in \hat{\mathcal{S}}_{n-t}} \mathrm{tr}(H\sigma) \geq 1$.
- NO. $\max_{\sigma \in \hat{\mathcal{S}}_{n-t}} \mathrm{tr}(H\sigma) \leq 0$.

To symmetrise the promise around a tunable threshold, define the shifted Hamiltonian

$$H' \coloneqq \alpha I - H, \tag{S2.54}$$

for a parameter $\alpha > 0$ to be fixed later. The problem becomes the following.

- YES. $\max_\sigma \operatorname{tr}(H'\sigma) \geq \alpha$.
- NO. $\max_\sigma \operatorname{tr}(H'\sigma) \leq \alpha - 1$.

**Embedding into $\hat{\mathcal{S}}_{n,t}$.** Recall that every state vector in $\hat{\mathcal{S}}_{n,t}$ admits a decomposition of the form (Definition 18):

$$|\psi\rangle = C\big(|\phi\rangle \otimes |0\rangle^{\otimes(n-t)}\big), \tag{S2.55}$$

where $|\phi\rangle$ is a $t$-qubit state vector and $C$ is a Clifford unitary. Let $|\phi\rangle$ be a $t$-qubit state vector to be fixed later, and define the lifted Hamiltonian acting on $n$ qubits as

$$H''_n \coloneqq |\phi\rangle\langle\phi| \otimes H'_{n-t}. \tag{S2.56}$$

By construction, $H''_n$ acts as $H'_{n-t}$ on the subspace where the first $t$ qubits are aligned with $|\phi\rangle$ and suppresses contributions from states with small overlap on this subsystem. Let us show that optimising $H''_n$ on $\mathcal{S}_{n,t}$ is at least as hard as optimising $H'_{n-t}$ on $\mathcal{S}_{n-t}$ thus proving the claim for any $(n-t) = \Omega(n)$.

- YES. Assume there exists a state vector $|\sigma\rangle_{n-t}$ such that

$$\operatorname{tr}\big(H'_{n-t}\,|\sigma\rangle\langle\sigma|\big) \geq \alpha. \tag{S2.57}$$

  Then the product state vector $|\phi\rangle \otimes |\sigma\rangle$ satisfies

$$\operatorname{tr}\big(H''_n(|\phi\rangle\langle\phi| \otimes |\sigma\rangle\langle\sigma|)\big) \geq \alpha. \tag{S2.58}$$

  We note that the $n$-qubit state $|\phi\rangle \otimes |\sigma\rangle$ lies in $\hat{\mathcal{S}}_{n,t}$. Hence, the lifted instance is a YES instance.

- NO. Assume instead that

$$\max_{|\sigma_{n-t}\rangle \in \mathcal{S}_{n-t}} \operatorname{tr}\big(H'_{n-t}\,|\sigma\rangle\langle\sigma|\big) \leq \alpha - 1. \tag{S2.59}$$

  Let $|\psi\rangle \in \hat{\mathcal{S}}_{n,t}$ be an arbitrary $t$-doped state vector. For this, we define the overlap

$$F \coloneqq \operatorname{tr}\big(|\psi\rangle\langle\psi|\,(|\phi\rangle\langle\phi| \otimes I)\big). \tag{S2.60}$$

  If $F \geq 1 - \vartheta$, then projecting $|\psi\rangle$ onto the $|\phi\rangle$ subspace produces a normalised $(n-t)$-qubit state vector $|\tilde{\sigma}\rangle$ such that, by the gentle measurement lemma [80], we have that

$$\| \, |\psi\rangle\langle\psi| - |\phi\rangle\langle\phi| \otimes |\tilde{\sigma}\rangle\langle\tilde{\sigma}| \, \|_1 \leq 2\sqrt{\vartheta}. \tag{S2.61}$$

  We now aim to show $|\tilde{\sigma}\rangle$ to be a stabilizer state (see Eq. (S1.23)). Let $P_6$ be the stabilizer purity introduced in Lemma 3. We then choose $|\phi\rangle \coloneqq \operatorname{argmin}_\tau P_6(|\tau\rangle)$, which minimises the stabilizer purity over $t$-qubit states. However, by construction of $t$-doped states (Definition 18), for any $|\psi\rangle \in \mathcal{S}_{n,t}$ we have that $P_6(|\phi\rangle) \leq P_6(|\psi\rangle)$. Let us note that we can bound

$$\| \, |\psi\rangle\langle\psi|^{\otimes 6} - (|\phi\rangle\langle\phi| \otimes |\tilde{\sigma}\rangle\langle\tilde{\sigma}|)^{\otimes 6} \|_1 \leq 6 \, \| \, |\psi\rangle\langle\psi| - |\phi\rangle\langle\phi| \otimes |\tilde{\sigma}\rangle\langle\tilde{\sigma}| \, \|_1 \leq 12\sqrt{\vartheta}. \tag{S2.62}$$

  Then, we find that

$$\begin{aligned} P_6(|\phi\rangle) \leq P_6(|\psi\rangle) &\leq P_6(|\phi\rangle)P_6(|\tilde{\sigma}\rangle) + 12\sqrt{\vartheta} \left\| \frac{1}{d} \sum_P P^{\otimes 6} \right\|_\infty \\ &\leq P_6(|\phi\rangle)P_6(|\tilde{\sigma}\rangle) + 12\sqrt{\vartheta}. \end{aligned} \tag{S2.63}$$

  Thus we obtain the lower bound

$$P_6(|\tilde{\sigma}\rangle) \geq 1 - \frac{12\sqrt{\vartheta}}{P_6(|\phi\rangle)}. \tag{S2.64}$$

Let us now use the relation between stabilizer purity and stabilizer fidelity $F_{\text{stab}}$ in Lemma 3 to lower bound the stabilizer fidelity of $|\tilde{\sigma}\rangle$ as

$$F_{\text{stab}}(|\tilde{\sigma}\rangle) \geq \left(1 - \frac{12\sqrt{\vartheta}}{P_6(|\phi\rangle)}\right)^C \geq 1 - \frac{12C\sqrt{\vartheta}}{P_6(|\phi\rangle)} \geq 1 - \frac{12C\sqrt{\vartheta}}{2^t}. \tag{S2.65}$$

where we have upper bounded $P_6(|\phi\rangle) \leq 2^t$ being $|\phi\rangle$ a $t$-qubit state. Hence, we deduce that there exists a stabilizer state vector $|\sigma\rangle$ on $n-t$ qubits such that

$$\| \, |\sigma\rangle\langle\sigma| - |\tilde{\sigma}\rangle\langle\tilde{\sigma}| \, \|_1 = 2\sqrt{1 - F_{\text{stab}}(|\tilde{\sigma}\rangle)} \leq 2\sqrt{12C2^{-t}\sqrt{\theta}}\,, \tag{S2.66}$$

from which we have

$$\operatorname{tr}\big(|\tilde{\sigma}\rangle\langle\tilde{\sigma}| \, H'_{n-t}\big) \leq \alpha - 1 + 2\sqrt{12C2^{-t}\sqrt{\theta}}\|H_{n-t}\|_\infty \tag{S2.67}$$

where we have used the condition in Eq. (S2.59). Using $F \geq 1-\theta$, we derive

$$\begin{aligned}
\operatorname{tr}\big(|\psi\rangle\langle\psi| \, (|\phi\rangle\langle\phi| \otimes H'_{n-t})\big) &\leq \alpha - \operatorname{tr}(|\psi\rangle\langle\psi| \, (|\phi\rangle\langle\phi| \otimes H_{n-t})) \\
&\leq \alpha - \operatorname{tr}(|\phi\rangle\langle\phi| \otimes |\tilde{\sigma}\rangle\langle\tilde{\sigma}| \, (|\phi\rangle\langle\phi| \otimes H_{n-t})) + 2\sqrt{\vartheta}\|H_{n-t}\|_\infty \\
&\leq \alpha - \operatorname{tr}(|\tilde{\sigma}\rangle\langle\tilde{\sigma}| \, H_{n-t}) + 2\sqrt{\theta}\|H_{n-t}\|_\infty \\
&\leq \operatorname{tr}\big(|\tilde{\sigma}\rangle\langle\tilde{\sigma}| \, H'_{n-t}\big) + 2\sqrt{\theta}\|H_{n-t}\|_\infty \\
&\leq \alpha - 1 + \left(2\sqrt{\theta} + 2\sqrt{12C2^{-t}\sqrt{\theta}}\right)\|H_{n-t}\|_\infty.
\end{aligned} \tag{S2.68}$$

In summary, this yields

$$\operatorname{tr}(|\psi\rangle\langle\psi| \, H''_n) \leq \alpha - 1 + \left(2\sqrt{\theta} + 2\sqrt{12C2^{-t}\sqrt{\theta}}\right)\|H_{n-t}\|_\infty. \tag{S2.69}$$

In the converse case, if $F < 1-\vartheta$, then $|\psi\rangle$ has significant support outside the $|\phi\rangle$ subspace, and by definition of $H''_n$,

$$\operatorname{tr}\big(|\psi\rangle\langle\psi| \, (|\phi\rangle\langle\phi| \otimes H'_{n-t})\big) \leq \| \, \langle\phi| \, \psi \, |\phi\rangle \, \|_1 \|H'_{n-t}\|_\infty = \| \, \langle\phi| \, \psi \, |\phi\rangle \, \|_2 \|H'_{n-t}\|_\infty < (1-\theta)\|H'_{n-t}\|_\infty\,, \tag{S2.70}$$

where we have used the fact that $\|\langle\phi|\psi|\phi\rangle\|_2 = \sqrt{\operatorname{tr}[(|\phi\rangle\langle\phi| \otimes I) \, |\psi\rangle\langle\psi| \, (|\phi\rangle\langle\phi| \otimes I) \, |\psi\rangle\langle\psi|]} = F$. We now impose for both cases, i.e., $F \geq 1-\theta$ and $F < 1-\theta$, that $\operatorname{tr}\big(|\psi\rangle\langle\psi| \, H''_{n-t}\big) \leq \alpha - \frac{1}{2}$ to separate YES and NO instance. For the latter case, we simply impose

$$(1-\theta)\|H'_{n-t}\|_\infty \leq \alpha - \frac{1}{2}. \tag{S2.71}$$

Crudely upper bounding $\|H'_{n-t}\|_\infty = \alpha + \|H_{n-t}\|_\infty$, we find a condition on the value of $\alpha$ as $\alpha \geq \theta^{-1}[\frac{1}{2} + \|H_{n-t}\|_\infty(1-\theta)]$. In the former case ($F \geq 1-\theta$), we impose

$$\alpha - 1 + \left(2\sqrt{\theta} + 2\sqrt{12C2^{-t}\sqrt{\theta}}\right)\|H_{n-t}\|_\infty \leq \alpha - \frac{1}{2} \tag{S2.72}$$

from which we get the condition for $\theta$ given by

$$\theta^{1/4} \leq \frac{1}{2(2 + 2\sqrt{12C2^{-t}})\|H_{n-t}\|_\infty}. \tag{S2.73}$$

By choosing $\alpha$ and $\vartheta$ such that both bounds lie strictly below $\alpha - \frac{1}{2}$, no state in $\hat{\mathcal{S}}_{n,t}$ can achieve expectation value at least $\alpha$ in the NO case. Combining the two, we get

$$\begin{aligned}
\alpha &\geq (1/2 + \|H_{n-t}\|_\infty)\|H_{n-t}\|_\infty^4[2(2 + 2\sqrt{12C2^{-t}})]^4 \\
&\geq 2^{-10}d^{15}\log^{30} d.
\end{aligned} \tag{S2.74}$$

**Reduction to $\mathsf{WWD}_\delta$.** Define the witness

$$W \coloneqq -\frac{H_n''}{\|H_n''\|_2}, \qquad \gamma \coloneqq -\frac{1}{2\|H_n''\|_2}. \tag{S2.75}$$

Since optimisation over the convex set $\hat{\mathcal{S}}_{n,t}$ is attained on its extreme points, the Hamiltonian decision problem reduces to an instance of $\mathsf{WWD}_\delta(\hat{\mathcal{S}}_{n,t})$ with parameters $(W, \gamma)$, provided $\delta$ is chosen sufficiently small. Using submultiplicativity and the definition of $H'$, we obtain

$$\begin{aligned} \|H_n''\|_2 &\leq \|H'_{n-t}\|_2 \\ &\leq \alpha\sqrt{d}2^{-t/2} + \|H_{n-t}\|_2 \\ &\leq \frac{1}{2}d^{31/2}\log^{30} d, \end{aligned} \tag{S2.76}$$

where we have set $\alpha = 2^{-10}d^{15}\log^{30} d$, and the last inequality follows from the norm bounds established in Theorem 7. Combining this with the conditions in the YES and NO cases of the weak-witness formulation yields the constraint

$$\delta \leq \frac{1}{2d^{31/2}\log^{30} d}. \tag{S2.77}$$

With this choice of $\delta$, oracle access to $\mathsf{WWD}_\delta(\hat{\mathcal{S}}_{n,t})$ suffices to solve a Hamiltonian optimisation problem on $n - t \geq (1-\kappa)n$ qubits. Therefore, under the ETH, the problem $\mathsf{WWD}_\delta(\hat{\mathcal{S}}_{n,t})$ cannot belong to $\mathsf{QP}^{2-\eta}$ for any $\eta > 0$. □

**Theorem 8** (Superpolynomial complexity of $\rho$-$\mathsf{WMEM}_\varepsilon(\hat{\mathcal{S}}_{n,t})$)**.** *Let $\hat{\mathcal{S}}_{n,t}$ be the $t$-doped stabilizer polytope on $n$-qubits and $t \leq \kappa n$ with $\kappa < 1$. Under the ETH (Definition 5), and for any $\varepsilon \leq (2^{16}3^3d^{119/2}\log^{90} d)^{-1}$, with $C$ the constant introduced in Lemma 3, the decision problem of $\rho$-$\mathsf{WMEM}_\varepsilon(\hat{\mathcal{S}}_{n,t}) \notin \mathsf{QP}^{2-\eta}$ for any $\eta > 0$.*

*Proof.* To prove the result, we reduce the membership problem to the weak witness detection problem $\mathsf{WWD}_\delta(\hat{\mathcal{S}}_{n,t})$. In Lemma 13 we show that $\mathsf{WWD}_\delta(\hat{\mathcal{S}}_{n,t}) \notin \mathsf{QP}^{2-\eta}$ for any $\eta > 0$, assuming the exponential time hypothesis stated in Definition 5, provided

$$\delta \leq \frac{1}{2d^{31/2}\log^{30} d}. \tag{S2.78}$$

The reduction follows from Lemma 9. Consequently, as long as $\varepsilon \leq (2^{16}3^3d^{119/2}\log^{90} d)^{-1}$, there exists a polynomial-time reduction between $\rho$-$\mathsf{WMEM}_\varepsilon(\hat{\mathcal{S}}_{n,t})$ and $\mathsf{WWD}_\delta\hat{\mathcal{S}}_{n,t})$. This concludes the proof. □

Before introducing the next lemma, we introduce two notions that are relevant to our analysis. The following definition extends the notion of robustness of magic by allowing, as free resources, convex combinations of $t$-doped states, as also done in Ref. [81]. As in the case of stabilizer states (see Definition 14), this quantity serves a dual purpose: (i) it yields an algorithm to decide membership in the convex hull of $t$-doped states $\hat{\mathcal{S}}_{n,t}$, and (ii) it quantifies how far a given state lies outside $\hat{\mathcal{S}}_{n,t}$.

**Definition 22** ($t$-extended robustness of magic)**.** *Given an $n$-qubit density matrix $\rho$, the $t$-extended robustness of magic is defined as*

$$\mathcal{R}_t(\rho) \coloneqq \min_{\{x_i,\psi_i\}} \left\{ \sum_i |x_i| \;:\; \rho = \sum_i x_i\psi_i,\ \psi_i \in \mathcal{S}_{n,t} \right\}. \tag{S2.79}$$

*Moreover, $\mathcal{R}_t(\rho) = 0$ if and only if $\rho \in \hat{\mathcal{S}}_{n,t}$.*

We now establish a dual characterisation of the $t$-extended robustness and derive a uniform upper bound on its value.

**Lemma 14** (Dual formulation of $t$-extended robustness)**.** *The $t$-extended robustness admits the dual formulation*

$$\mathcal{R}_t(\rho) = \max_{\substack{W \\ \mathrm{tr}(W\psi)\leq 1\ \forall\,\psi\in\mathcal{S}_{n,t}}} \mathrm{tr}(W\rho). \tag{S2.80}$$

*Moreover, for any $n$-qubit state $\rho$, it holds that $\mathcal{R}_t(\rho) \leq \sqrt{d(d+1)}$, and any feasible witness $W$ in Eq. (S2.80) satisfies $\|W\|_2 \leq \sqrt{d(d+1)}$.*

*Proof.* The dual formulation (S2.80) follows from the standard dual program associated with $\mathcal{R}_t$, in complete analogy with the robustness of magic (Definition 14). To upper bound both the $t$-extended robustness and the Hilbert–Schmidt norm of feasible witnesses, we proceed as in Ref. [15]. First, we observe that

$$\mathcal{R}_t(\rho) \leq \max_{\substack{W \\ \mathrm{tr}(W\psi)\leq 1 \ \forall\, \psi\in\mathcal{S}_{n,t}}} \|W\|_\infty \leq \max_{\substack{W \\ \mathrm{tr}(W\psi)\leq 1 \ \forall\, \psi\in\mathcal{S}_{n,t}}} \|W\|_2. \tag{S2.81}$$

Expanding $W$ in the Pauli basis, its Hilbert–Schmidt norm can be written as

$$\|W\|_2 = \sqrt{\frac{1}{d}\sum_P \mathrm{tr}^2(WP)}. \tag{S2.82}$$

The constraint $\mathrm{tr}(W\psi) \leq 1$ holds in particular for all stabilizer states $\sigma \in \mathcal{S}_n$. For such states, $\mathrm{tr}(\sigma P) \in \{\pm 1\}$ for all Pauli operators $P$, and hence

$$\mathrm{tr}(W\sigma) = \frac{1}{d}\sum_P \mathrm{tr}(WP)\,\mathrm{tr}(\sigma P) \leq 1. \tag{S2.83}$$

It follows that imposing the constraint only on stabilizer states is already sufficient to bound $\mathcal{R}_t$, so the remainder of the proof coincides with the argument in Ref. [15]. We include it here for completeness.

Fix an unsigned stabilizer group $G$. Define the matrix $V_{P,\sigma} = \frac{1}{\sqrt{d}}\,\mathrm{tr}(\sigma P)$, where $P \in G$ and $\sigma$ ranges over the $d$ stabilizer states with stabilizer group $G$. This matrix is unitary: it is square since $|G| = d$, and its rows and columns are orthonormal. Indeed,

$$\sum_\sigma V_{P,\sigma}V_{Q,\sigma} = \frac{1}{d}\sum_\sigma \mathrm{tr}(\sigma P)\,\mathrm{tr}(\sigma Q) = \frac{1}{d}\sum_\sigma \mathrm{tr}(\sigma PQ) = \delta_{P,Q}, \tag{S2.84}$$

and similarly,

$$\sum_P V_{P,\sigma}V_{P,\sigma'} = \frac{1}{d}\sum_P \mathrm{tr}(\sigma P)\,\mathrm{tr}(\sigma' P) = \mathrm{tr}(\sigma\sigma') = \delta_{\sigma,\sigma'}, \tag{S2.85}$$

where we have used the fact that two stabilizer states associated with the same stabilizer group are either identical or orthogonal. Using this unitary change of basis, we obtain

$$\sum_{P\in G}\mathrm{tr}^2(WP) = \sum_\sigma\left(\frac{1}{\sqrt{d}}\sum_{P\in G}\mathrm{tr}(P\sigma)\,\mathrm{tr}(WP)\right)^2 = d\sum_\sigma \mathrm{tr}^2(W\sigma) \leq d^2. \tag{S2.86}$$

Finally, since there exist $d+1$ stabilizer groups $\{G_j\}_{j=1}^{d+1}$ that intersect only at the identity, we can write

$$\sum_P \mathrm{tr}^2(WP) \leq \sum_{j=1}^{d+1}\sum_{P\in G_j}\mathrm{tr}^2(WP) \leq (d+1)d^2. \tag{S2.87}$$

Therefore, $\|W\|_2 \leq \sqrt{d(d+1)}$, which concludes the proof. □

Given the definition of $t$-extended robustness, it is possible to prove the following result.

**Lemma 15** ($\rho$-$\mathsf{WMEM}_\varepsilon(\hat{\mathcal{S}}_{n,t}) \in \mathsf{QP}$ for any $t = O(\log\log d)$)**.** *For any $0 < \varepsilon < 1$, there exists an algorithm running in time $\mathsf{poly}(2^{\log^2 d + 2^{2t}(\log d + \log\varepsilon^{-1})})$ that, by computing the $t$-extended robustness of magic up to additive precision $\nu \leq \frac{\varepsilon^6}{2^{26}\cdot 3^3\cdot d^{24}}$, solves $\rho$-$\mathsf{WMEM}_\varepsilon(\hat{\mathcal{S}}_{n,t})$. In particular, the problem lies in $\mathsf{QP}^2$ for any $t < \log\log d$, and in $\mathsf{QP}$ for any $t = O(\log\log d)$.*

*Proof.* The proof employs a strategy analogous to that presented in Lemma 11, substituting the robustness of magic with its $t$-extended counterpart. The primary challenge arises from the continuity of the set $\mathcal{S}_{n,t}$, which would otherwise result in an infinite runtime for the algorithm. To address this, we construct a $\vartheta$-packing net $\mathcal{S}^\vartheta_{n,t}$ of $\mathcal{S}_{n,t}$ such that $\mathcal{S}^\vartheta_{n,t} \subseteq \mathcal{S}_{n,t}$. A straightforward method for constructing this ensemble involves first creating a $\vartheta$-packing net of the set of $t$-qubit states and subsequently applying rotations using the Clifford group. By Definition 18, this ensures that $\mathcal{S}^\vartheta_{n,t} \subseteq \mathcal{S}_{n,t}$. Let $\hat{\mathcal{S}}^\vartheta_{n,t}$ denote the convex hull of $\mathcal{S}^\vartheta_{n,t}$. For any state $\rho \in \hat{\mathcal{S}}_{n,t}$, there exists a state $\rho^\vartheta \in \hat{\mathcal{S}}^\vartheta_{n,t}$ such that $\|\rho - \rho^\vartheta\|_2 \leq \vartheta$.

With the ensemble of states thus defined, we proceed to prove the lemma by demonstrating that the $t$-extended robustness, when evaluated on the $\vartheta$-packing net, is sufficient to solve the one-sided problem $\rho$-$\mathsf{WMEM}^1_\delta(\hat{\mathcal{S}}_{n,t})$. The $t$-extended robustness admits the dual formulation Lemma 14

$$\mathcal{R}_t(\rho) = \max_{\substack{W \\ \mathrm{tr}(W\psi)\leq 1 \ \forall \psi \in \mathcal{S}_{n,t}}} \mathrm{tr}(W\rho), \tag{S2.88}$$

where the feasible set of witnesses is given by

$$\mathcal{W}_t \coloneqq \left\{ W : \max_{\psi \in \mathcal{S}_{n,t}} \mathrm{tr}(W\psi) \leq 1 \right\}. \tag{S2.89}$$

To adapt this definition to $\mathcal{S}^\vartheta_{n,t}$, we obtain

$$\mathcal{R}^\vartheta_t(\rho) = \max_{\substack{W \\ \mathrm{tr}(W\psi)\leq 1 \ \forall \psi \in \mathcal{S}^\vartheta_{n,t}}} \mathrm{tr}(W\rho), \tag{S2.90}$$

with the corresponding feasible set defined as

$$\mathcal{W}^\vartheta_t \coloneqq \left\{ W : \max_{\psi \in \mathcal{S}^\vartheta_{n,t}} \mathrm{tr}(W\psi) \leq 1 \right\}. \tag{S2.91}$$

We begin by considering an idealised setting in which the entire witness set $\mathcal{W}^\vartheta_t$ is accessible exactly. Subsequently, we introduce robustness with respect to finite precision by defining the $\nu$-robust witness set. Recall from Definition 6 that for a set $K$, the inner core

$$S(K, -\nu) = \{W \in K : S(\{W\}, \nu) \subseteq K\} \tag{S2.92}$$

consists of all points whose $\nu$-ball is contained in $K$ where $S(\{W\}, \nu) = \{W' : \|W' - W\|_2 \leq \nu\}$. For $K = \mathcal{W}^\vartheta_t$, we have

$$S(\mathcal{W}^\vartheta_t, -\nu) = \{W \in \mathcal{W}^\vartheta_t : \forall W' \text{ with } \|W' - W\|_2 \leq \nu,\ W' \in \mathcal{W}^\vartheta_t\}. \tag{S2.93}$$

For the reminder of the proof, we adopt the notation $\mathcal{W}^\vartheta_{t,\nu} \equiv S(\mathcal{W}^\vartheta_t, -\nu)$. The associated optimisation problem

$$R^\vartheta_{t,\nu}(\rho) \coloneqq \max_{W \in \mathcal{W}^\vartheta_{t,\nu}} \mathrm{tr}(W\rho) \tag{S2.94}$$

is a linear program that can be solved up to additive error $\nu$ using the ellipsoid algorithm [65], provided that $\mathcal{W}^\vartheta_{t,\nu}$ is nonempty. We now show that computing $R^\vartheta_{t,\nu}(\rho)$ suffices to solve the one-sided problem $\rho$-$\mathsf{WMEM}^1_\delta(\hat{\mathcal{S}}_{n,t})$, from which the full weak membership problem follows by Lemma 10.

- YES. Let $\rho \in \hat{\mathcal{S}}_{n,t}$. By definition of the $\vartheta$-packing net, there exists $\rho^\vartheta \in \hat{\mathcal{S}}^\vartheta_{n,t}$ such that $\|\rho - \rho^\vartheta\|_2 \leq \vartheta$. For any witness $W$, we then have

$$\begin{aligned} \mathrm{tr}(W\rho) &\leq \mathrm{tr}\big(W\rho^\vartheta\big) + \vartheta\|W\|_2 + \nu \\ &\leq 1 + \vartheta\sqrt{d(d+1)} + \nu, \end{aligned} \tag{S2.95}$$

  where we have used $\|W\|_2 \leq \sqrt{d(d+1)}$ (Lemma 14) and the fact that the optimisation is solved up to $\nu$-additive error. Hence, no YES instance can be misclassified as NO, provided that NO instances yield values strictly larger than $1 + \nu + \vartheta\sqrt{d(d+1)}$.

- NO. Let $\rho$ satisfy $\min_{\sigma \in \hat{\mathcal{S}}_{n,t}} \|\rho - \sigma\|_2 \geq \delta$. Then

$$\min_{\sigma \in \hat{\mathcal{S}}^\vartheta_{n,t}} \|\rho - \sigma\|_2 \geq \min_{\sigma \in \hat{\mathcal{S}}_{n,t}} \|\rho - \sigma\|_2 \geq \delta. \tag{S2.96}$$

  By Lemma 7, since $\hat{\mathcal{S}}^\theta_{n,t}$ is convex, there exists a Hermitian operator $\tilde{W}$ such that

$$\mathrm{tr}\Big(\tilde{W}\rho\Big) > \max_{\sigma \in \hat{\mathcal{S}}^\vartheta_{n,t}} \mathrm{tr}\Big(\tilde{W}\sigma\Big) + \delta^2. \tag{S2.97}$$

We rescale $\tilde{W}$ into a valid element of $\mathcal{W}^{\vartheta}_{t,\nu}$ by defining

$$W \coloneqq \frac{\tilde{W}}{\max_{\sigma\in\hat{\mathcal{S}}^{\vartheta}_{n,t}} \operatorname{tr}\left(\tilde{W}\sigma\right)} - \nu I. \tag{S2.98}$$

Indeed, for any $H$ such that $\|H\|_2 \leq \nu$ and $\sigma \in \hat{\mathcal{S}}^{\vartheta}_{n,t}$, we have $\operatorname{tr}[(W+H)\sigma] \leq \operatorname{tr}(W\sigma) + \nu \leq 1 - \nu + \nu \leq 1$. Hence $W \in \mathcal{W}^{\vartheta}_{t,\nu}$. Evaluating $W$ on $\rho$ yields

$$\operatorname{tr}(W\rho) = \frac{\operatorname{tr}\left(\tilde{W}\rho\right)}{\max_{\sigma\in\hat{\mathcal{S}}^{\vartheta}_{n,t}} \operatorname{tr}\left(\tilde{W}\sigma\right)} - \nu > 1 + \frac{\delta^2}{\max_{\sigma\in\hat{\mathcal{S}}^{\vartheta}_{n,t}} \operatorname{tr}\left(\tilde{W}\sigma\right)} - \nu.$$

To tell apart YES and NO instant, we must have $\operatorname{tr}(W\rho) > 1 + \nu + \vartheta\sqrt{d(d+1)}$. It suffices that

$$\frac{\delta^2}{\max_{\sigma\in\hat{\mathcal{S}}^{\vartheta}_{n,t}} \operatorname{tr}\left(\tilde{W}\sigma\right)} > 3\nu + \vartheta\sqrt{d(d+1)}. \tag{S2.99}$$

Using the bound $\max_{\sigma\in\hat{\mathcal{S}}^{\vartheta}_{n,t}} \operatorname{tr}\left(\tilde{W}\sigma\right) \leq \|\tilde{W}\|_2 \leq 2$ (see Lemma 7), we obtain

$$\frac{\delta^2}{2} > 3\nu + \vartheta\sqrt{d(d+1)}. \tag{S2.100}$$

For simplicity, we choose $\nu \leq \frac{\delta^2}{12}$ and $\vartheta < \frac{\delta^2}{4\sqrt{d(d+1)}}$, which ensures that the condition is satisfied. We now examine the runtime of the algorithm. Since we are solving a linear program, the runtime is $\mathsf{poly}(|\mathcal{S}^{\vartheta}_{n,t}|, \log \nu^{-1})$ and therefore it is governed by an upper bound on the size of the set $\mathcal{S}^{\vartheta}_{n,t}$. As shown in Lemma 2, we have:

$$|\mathcal{S}^{\vartheta}_{n,t}| \leq 2^{2\log^2 d + 4\log d} \left(\frac{4}{\vartheta}\right)^{2^{2t}} \tag{S2.101}$$

Using the bound on $\vartheta$, this becomes

$$\begin{aligned} |\mathcal{S}^{\vartheta}_{n,t}| &\leq 2^{2\log^2 d + 4\log d} \left(\frac{4\sqrt{d(d+1)}}{\delta^2}\right)^{2^{2t}} \leq 2^{2\log^2 d + 4\log d} \left(\frac{8d}{\delta^2}\right)^{2^{2t}} \\ &= 2^{2\log^2 d + 4\log d + 2^{2t}(3 + \log d + 2\log \delta^{-1})}. \end{aligned} \tag{S2.102}$$

The above argument shows that computing $R^{\vartheta}_{t,\nu}(\rho)$ solves $\rho$-$\mathsf{WMEM}^1_{\delta}(\hat{\mathcal{S}}_{n,t})$. By Lemma 10, the one-sided and two-sided weak membership problems are polynomially equivalent. In particular, there exists a polynomial relation

$$\delta = \frac{\varepsilon^3}{2^{12}\cdot 3 \cdot d^{12}} \tag{S2.103}$$

such that solving the one-sided problem with precision $\delta$ yields a solution to $\rho$-$\mathsf{WMEM}_{\varepsilon}(\hat{\mathcal{S}}_{n,t})$. Choosing $\nu \leq \frac{\varepsilon^6}{2^{26}\cdot 3^3 \cdot d^{24}}$ completes the reduction.

□

### Appendix S.3: Additional results and applications

This section first introduces the notion of completely $t$-doped stabilizer-preserving channels and explores in Appendix S.3 1, which generalises completely stabilizer-preserving channels (Lemma 4). We then present applications of these results to the computational complexity of channel classification in Appendix S.3 2, and to the limitations of magic-state distillation in Appendix S.3 3. In particular, we show that our hardness results imply both the intractability of certain channel classification tasks and the existence of non-stabilizer states that admit no efficient stabilizer-based distillation protocol.

### 1. Completely $t$-doped stabilizer preserving channels and their characterisation

In this section, we introduce a generalisation of completely stabilizer preserving channels. For stabilizer states, we have seen that there exist channels, called *completely stabilizer-preserving channels*, that map states belonging to the stabilizer polytope into states that remain within the polytope. This notion can be generalised to *$t$-doped stabilizer-preserving channels*, namely channels that map $\hat{\mathcal{S}}_n$ into $\hat{\mathcal{S}}_{n,t}$. Formally, we define the set of completely $t$-doped stabilizer-preserving channels as follows.

**Definition 23** (Completely $t$-doped stabilizer-preserving channels)**.** *Let $\mathcal{E}$ be a quantum channel with an $n$-qubit input. We say that $\mathcal{E}$ is* completely $t$-doped stabilizer-preserving *if*

$$(\mathcal{E} \otimes I)(\sigma) \in \hat{\mathcal{S}}_{n+m,t} \tag{S3.1}$$

*for all $\sigma \in \hat{\mathcal{S}}_{n+m}$ and for all $m \in \mathbb{N}$.*

In words, the action of the channel $\mathcal{E}$ on a stabilizer input of arbitrary size yields a convex combination of $t$-doped states which, for $t = O(\log n)$, can be efficiently simulated classically and, therefore, do not provide any computational advantage.

We now provide a useful characterisation of completely $t$-doped stabilizer-preserving channels. To this end, we first introduce the following lemma.

**Lemma 16** ([82])**.** *Let $|\psi\rangle \in \mathcal{S}_{n+m,t}$ be a $t$-doped stabilizer state vector on $n+m$ qubits. Then:*

1. *For any stabilizer state vector $|\phi\rangle \in \mathcal{S}_m$, the (unnormalised) vector $(I \otimes \langle\phi|)\,|\psi\rangle$ is a $t$-doped stabilizer state on $n$ qubits.*
2. *The reduced state $\mathrm{tr}_m(|\psi\rangle\langle\psi|)$ lies in the convex hull of $t$-doped stabilizer states of $n$ qubits, i.e., $\mathrm{tr}_m(|\psi\rangle\langle\psi|) \in \hat{\mathcal{S}}_{n,t}$.*

*Proof.* We work entirely at the level of the Pauli expansion. Let $d = 2^{n+m}$ and write

$$|\psi\rangle\langle\psi| = \frac{1}{d} \sum_{P \in \mathcal{P}_{n+m}} \alpha_P \, P, \tag{S3.2}$$

where $\mathcal{P}_{n+m}$ is the $n+m$-qubit Pauli group modulo phases and $\alpha_P \coloneqq \mathrm{tr}(P\,|\psi\rangle\langle\psi|)$. By definition of a $t$-doped stabilizer state, $|\psi\rangle\langle\psi|$ can be written as

$$|\psi\rangle\langle\psi| = \left(\frac{1}{2^t} \sum_{Q \in \bar{G}} \alpha_Q Q\right) \left(\frac{1}{2^{n+m-t}} \sum_{R \in G} R\right), \tag{S3.3}$$

where $G$ is the set of Pauli operators that stabilises $|\psi\rangle$, while $\bar{G}$ is a set of Pauli operators that do not stabilise $|\psi\rangle$, but have non-zero component.

Let $|\phi\rangle \in \mathcal{S}_m$ be a stabilizer state vector. Up to a local Clifford unitary on the last $m$ qubits, we may assume that $|\phi\rangle = |0\rangle^{\otimes m}$. Such a Clifford acts by permuting Pauli operators and therefore preserves the cardinality of the set of nonzero coefficients $\{\alpha_P\}$, so $t$ remains unchanged. Now consider the unnormalised post-measurement operator

$$\rho' \coloneqq (I \otimes \langle 0|^{\otimes m})\,|\psi\rangle\langle\psi|\,(I \otimes |0\rangle^{\otimes m}). \tag{S3.4}$$

Using the Pauli expansion, we have

$$\rho' = \frac{1}{d} \sum_{P_n \otimes P_m} \alpha_{P_n \otimes P_m} \, P_n \, \langle 0|^{\otimes m} P_m |0\rangle^{\otimes m}, \tag{S3.5}$$

where the sum runs over $P_n \in \mathcal{P}_n$ and $P_m \in \mathcal{P}_m$. The scalar $\langle 0|^{\otimes m} |P_m| 0\rangle^{\otimes m}$ vanishes unless $P_m$ is a tensor product of $I$ and $Z$ operators. Hence, only those Pauli components of $|\psi\rangle\langle\psi|$ of the form $P_n \otimes Z_m$ contribute, where $Z_m \in \{I, Z\}^{\otimes m}$.

Let $G$ be the stabilizer subgroup associated with the decomposition given in (S3.3). The projection onto $|0\rangle^{\otimes m}$ restricts the Pauli expansion to a subset of the original $G$, and $\bar{G}$ but cannot generate new nonzero Pauli coefficients. Therefore, the number of distinct Pauli operators on the first $n$ qubits that appear with nonzero weight is at most $4^t$ [53]. This shows that $(I \otimes \langle\phi|)\,|\psi\rangle$ is an unnormalised $t$-doped stabilizer state on $n$ qubits.

Let $\{|\phi_k\rangle\}_k$ be any orthonormal stabilizer basis of the $m$-qubit Hilbert space. Then

$$\mathrm{tr}_m(|\psi\rangle\langle\psi|) = \sum_k (I \otimes \langle\phi_k|)\,|\psi\rangle\langle\psi|\,(I \otimes |\phi_k\rangle). \tag{S3.6}$$

Define $v_k \coloneqq (I \otimes \langle\phi_k|)\,|\psi\rangle$ and $p_k \coloneqq \|v_k\|_2^2$. If $p_k > 0$, let $|\psi_k\rangle \coloneqq p_k^{-1/2} v_k$. By Claim 1, each nonzero $v_k$ is an non-normalised $t$-doped stabilizer state, and therefore each $|\psi_k\rangle \in \mathcal{S}_{n,t}$. This yields the convex decomposition

$$\mathrm{tr}_m(|\psi\rangle\langle\psi|) = \sum_k p_k\ |\psi_k\rangle\langle\psi_k|\,, \tag{S3.7}$$

which shows that $\mathrm{tr}_m(|\psi\rangle\langle\psi|) \in \hat{\mathcal{S}}_{n,t}$. □

The next lemma gives a characterisation of completely $t$-doping stabilizer preserving channels in terms of their Choi states.

**Lemma 17** (Completely $t$-doping stabilizer preserving channels in terms of their Choi states)**.** *Let $\mathcal{E}$ be a quantum channel with $n$-qubit input. Then $\mathcal{E}$ is completely $t$-doping stabilizer preserving if and only if its Choi state $\rho_{\mathcal{E}} \in \hat{\mathcal{S}}_{2n,t}$.*

*Proof.* Following the proof in Ref. [23], we divide the proof into two parts. First, we show that in order to verify whether $\mathcal{E}$ is completely $t$-doping stabilizer preserving, it is sufficient to check the condition for $m = n$. Second, we show that this is equivalent to requiring that the Choi state lies in $\hat{\mathcal{S}}_{2n,t}$.

By linearity, it is sufficient to verify that $[\mathcal{E}\otimes I]$ maps pure stabilizer states in $\mathcal{S}_{n+m}$ to states in $\hat{\mathcal{S}}_{n+m,t}$. Let $m = n+m'$ and let $|\sigma\rangle$ be a stabilizer state vector on $n+m$ qubits. It is well known that any stabilizer state on a bipartite system $A|B$ is Clifford-unitarily equivalent to a tensor product of Bell pairs across the cut $A|B$ and local stabilizer states. Let $A$ denote the first $n$ qubits and $B$ the remaining $m$ qubits. Since there can be at most $n$ Bell pairs across the cut, there exists a local Clifford operation $C_B$ such that

$$|\sigma\rangle_{A,B} = (I_A \otimes C_B)\big(|\sigma'\rangle_{A,A'} \otimes |\sigma''\rangle_{B'}\big), \tag{S3.8}$$

where $A'$ denotes the first $n$ qubits of $B$ and $B'$ the remaining $m'$ qubits.

Suppose there exists a state vector $|\sigma\rangle_{A,B}$ with $m' > 0$ such that $[\mathcal{E} \otimes I](\sigma_{A,B}) \notin \hat{\mathcal{S}}_{n+m,t}$. We show that there then exists a corresponding state vector $|\sigma'\rangle_{A,A'}$ supported on $2n$ qubits such that $[\mathcal{E} \otimes I](\sigma_{A,A'}) \notin \hat{\mathcal{S}}_{2n,t}$. Indeed, by denoting $\mathcal{C}_B(\cdot) \coloneqq C_B(\cdot)C_B^\dagger$, we have that

$$\begin{aligned}[\mathcal{E} \otimes I](\sigma_{A,B}) &= [\mathcal{E} \otimes I] \circ [I_A \otimes \mathcal{C}_B]\big(\sigma'_{A,A'} \otimes \sigma''_{B'}\big) \\ &= [I_A \otimes \mathcal{C}_B]\Big([\mathcal{E} \otimes I](\sigma'_{A,A'}) \otimes \sigma''_{B'}\Big),\end{aligned} \tag{S3.9}$$

which implies that $[\mathcal{E} \otimes I](\sigma'_{A,A'}) \notin \hat{\mathcal{S}}_{2n,t}$, since tensoring with auxiliary stabilizer states and applying Clifford operations preserve membership in $\hat{\mathcal{S}}_{n+m,t}$.

We now show that the image of all pure stabilizer states under $[\mathcal{E} \otimes I]$ lies in $\hat{\mathcal{S}}_{2n,t}$ if and only if $\rho_{\mathcal{E}} \in \hat{\mathcal{S}}_{2n,t}$. The forward implication is immediate. For the reverse implication, let $W$ be a witness for $\hat{\mathcal{S}}_{n,t}$, that is, $\mathrm{tr}(W\psi) \leq 0$ for all $\psi \in \mathcal{S}_{n,t}$ (see Definition 21 with $\gamma = 0$). We claim that $W \otimes |\phi\rangle\langle\phi|$ is a valid witness for $\hat{\mathcal{S}}_{n+m,t}$ for any $m$-qubit stabilizer state vector $|\phi\rangle$. For any $\sigma \in \mathcal{S}_{n+m,t}$, we have

$$\mathrm{tr}\big((W \otimes |\phi\rangle\langle\phi|)\sigma\big) = \mathrm{tr}\big((W \otimes I)\tilde{\sigma}\big), \tag{S3.10}$$

where $\tilde{\sigma} = (I \otimes |\phi\rangle\langle\phi|)\,\sigma\,(I \otimes |\phi\rangle\langle\phi|)$. Since $|\phi\rangle$ is a stabilizer state, it follows from Lemma 16 that $\tilde{\sigma}$ is a non-normalized $t$-doped state. We further obtain

$$\mathrm{tr}\big((W \otimes I)\tilde{\sigma}\big) = \mathrm{tr}\big(W\ \mathrm{tr}_m(\tilde{\sigma})\big) \leq 0, \tag{S3.11}$$

where we have used that the partial trace of a $t$-doped state lies in $\hat{\mathcal{S}}_{n,t}$ by Lemma 16, together with the assumption that $W$ is a witness for $\hat{\mathcal{S}}_{n,t}$.

For the final step, suppose for contradiction that $\rho_{\mathcal{E}} \in \hat{\mathcal{S}}_{2n,t}$, but there exists $\phi \in \mathcal{S}_{2n}$ such that $\rho = [\mathcal{E}\otimes I](\phi) \notin \hat{\mathcal{S}}_{2n,t}$. By Lemma 7, there exists a witness $W$ for $\hat{\mathcal{S}}_{2n,t}$ such that $\mathrm{tr}(W\rho) > 0$. Then

$$0 < \mathrm{tr}(W\rho) = 2^n\,\mathrm{tr}\big(\rho_{\mathcal{E}\otimes I}\,(W \otimes \phi^T)\big). \tag{S3.12}$$

Since $W$ is a witness for $\hat{\mathcal{S}}_{2n,t}$ and the transpose of a stabilizer state is again a stabilizer state, the operator $2^n(W \otimes \phi^T)$ is a witness for $\hat{\mathcal{S}}_{4n,t}$. This implies that $\rho_{\mathcal{E}\otimes I} \notin \hat{\mathcal{S}}_{4n,t}$. Noting that $\rho_{\mathcal{E}\otimes I} = \rho_{\mathcal{E}} \otimes \phi^+_{2n}$ completes the proof. □

The set of $t$-doped stabilizer states thus provides an extension of the set of stabilizer states by also incorporating the use of non-Clifford operations, while still remaining classically simulable for $t = O(\log n)$.

### 2. Channel classification computational problem

In this section, we develop the first application of the results established in Appendix S.2 1. In particular, we present two corollaries of Theorem 6 and Lemma 15 in the context of channel classification. Starting from the first corollary, we consider the following computational task: given a description of a quantum channel $\mathcal{E}$, determine whether it admits an efficient description within the stabilizer formalism. We formalise this task as a decision problem, asking whether the channel $\mathcal{E}$ is completely stabilizer-preserving.

**Problem 8** (Completely stabilizer-preserving classification (CSPC))**.** *Let $\mathcal{E}$ be a quantum channel and let $\varepsilon > 0$. Decide whether $\mathcal{E}$ satisfies:*

- *For all $\sigma \in \mathcal{S}_{2n}$, $[\mathcal{E} \otimes \mathbb{1}_n](\sigma) \in S(\hat{\mathcal{S}}_{2n}, -\varepsilon)$; output* YES.
- *There exists $\sigma \in \mathcal{S}_{2n}$ such that $[\mathcal{E} \otimes \mathbb{1}_n](\sigma) \notin S(\hat{\mathcal{S}}_{2n}, \varepsilon)$; output* NO.

*An instance of this problem is denoted by* $\mathsf{CSPC}_\varepsilon$.

**Corollary 4** (Superpolynomial complexity of CSPC)**.** *Let $\hat{\mathcal{S}}_{2n}$ denote the stabilizer polytope on $2n$ qubits. Under the* exponential time hypothesis *(ETH) (see Definition 5), for any $\varepsilon \leq \frac{1}{2^{30}3^3 d^{52}\log^{18} d}$, the decision problem* $\mathsf{CSPC}_\varepsilon$ *is not contained in the complexity class* $\mathsf{QP}^{2-\eta}$ *for any $\eta > 0$.*

*Proof.* The proof proceeds by reduction to a weak membership problem for the stabilizer polytope. By Lemma 4, deciding whether a quantum channel $\mathcal{E}$ is completely stabilizer-preserving can be reduced to a membership test for its Choi state $\rho_\mathcal{E}$. More precisely, $\mathcal{E}$ is completely stabilizer-preserving if and only if its Choi state satisfies $\rho_\mathcal{E} \in \hat{\mathcal{S}}_{2n}$, together with the constraint $\mathrm{tr}_B\, \rho_\mathcal{E} = I/d$, where $B$ denotes the second $n$-qubit subsystem.

This motivates the following decision problem. Given a state $\rho_\mathcal{E}$ satisfying $\mathrm{tr}_B\, \rho_\mathcal{E} = I/d$ and a parameter $\delta > 0$, decide whether

- for all states $\sigma$ such that $\|\rho_\mathcal{E} - \sigma\|_2 \leq \delta$, we have $\sigma \in \hat{\mathcal{S}}_{2n}$; output YES.
- There exists a state $\sigma \in \hat{\mathcal{S}}_{2n}$ such that $\|\rho_\mathcal{E} - \sigma\|_2 > \delta$; output NO.

We denote this problem by $\rho_\mathcal{E}$-$\mathsf{WMEM}_\delta(\hat{\mathcal{S}}_{2n})$.

We first show that $\rho$-$\mathsf{WMEM}_\delta(\hat{\mathcal{S}}_n)$ for an arbitrary $n$-qubit state $\rho$ reduces to $\rho_\mathcal{E}$-$\mathsf{WMEM}_\delta(\hat{\mathcal{S}}_{2n})$. Assume oracle access to $\rho_\mathcal{E}$-$\mathsf{WMEM}_\delta(\hat{\mathcal{S}}_{2n})$. Given $\rho$, define $\omega \equiv I/d$ and consider the state $\rho \otimes \omega$, which is a valid Choi state since $\mathrm{tr}_B(\rho \otimes \omega) = I/d$. Querying the oracle on $\rho \otimes \omega$, a YES instance implies that every state $\sigma$ satisfying $\|\rho \otimes \omega - \sigma\|_2 \leq \delta$ belongs to $\hat{\mathcal{S}}_{2n}$. For any $\tilde{\sigma}$ such that $\|\rho - \tilde{\sigma}\|_2 \leq \delta$, we have

$$\|\rho \otimes \omega - \tilde{\sigma} \otimes \omega\|_2 \leq \|\rho - \tilde{\sigma}\|_2 \|\omega\|_2 = \delta d^{-1/2} \leq \delta, \tag{S3.13}$$

and hence $\tilde{\sigma} \otimes \omega \in \hat{\mathcal{S}}_{2n}$. Since the partial trace preserves stabilizer membership, this implies $\tilde{\sigma} \in \hat{\mathcal{S}}_n$.

Conversely, for a NO instance we have $\min_{\sigma \in \hat{\mathcal{S}}_{2n}} \|\rho \otimes \omega - \sigma\|_2 > \delta$. If there exists $\tilde{\sigma} \in \hat{\mathcal{S}}_n$ with $\|\rho - \tilde{\sigma}\|_2 \leq \delta$, then it would follow $\tilde{\sigma} \otimes \omega \in \hat{\mathcal{S}}_{2n}$, contradicting the inequality. Therefore, $\min_{\sigma \in \hat{\mathcal{S}}_n} \|\rho - \sigma\|_2 > \delta$, which completes the reduction.

We now reduce $\mathsf{CSPC}_\varepsilon$ to $\rho_\mathcal{E}$-$\mathsf{WMEM}_\delta(\hat{\mathcal{S}}_{2n})$. Assume oracle access to $\mathsf{CSPC}_\varepsilon$. Then:

- YES. For all $\sigma \in \mathcal{S}_{2n}$ and all states $\tau$ such that $\|[\mathcal{E} \otimes I](\sigma) - \tau\|_2 \leq \varepsilon$, we have $\tau \in \hat{\mathcal{S}}_{2n}$. In particular, choosing $\sigma = \phi^+_{2n}$ shows that the Choi state $\rho_\mathcal{E} = [\mathcal{E} \otimes I](\phi^+_{2n})$ is a YES instance of $\rho_\mathcal{E}$-$\mathsf{WMEM}_\delta(\hat{\mathcal{S}}_{2n})$, provided that $\delta \leq \varepsilon$.
- NO. There exists $\sigma \in \mathcal{S}_{2n}$ such that $\min_{\tau \in \hat{\mathcal{S}}_{2n}} \|[\mathcal{E} \otimes I](\sigma) - \tau\|_2 > \varepsilon$. By Lemma 7, there exists a witness $W'$ such that $\mathrm{tr}(W'[\mathcal{E} \otimes I](\sigma)) > \max_{\tau \in \mathcal{S}_{2n}} \mathrm{tr}(W'\tau) + \varepsilon^2$. We redefine $W := W' - \max_{\tau \in \mathcal{S}_{2n}} \mathrm{tr}(W'\tau) I$ so that $\max_{\tau \in \mathcal{S}_{2n}} \mathrm{tr}(W\tau) \leq 0$. Using the Choi representation, this can be written as

$$\mathrm{tr}(W[\mathcal{E} \otimes I](\sigma)) = d\,\mathrm{tr}\big(\rho_{\mathcal{E}\otimes I}(W \otimes \sigma^T)\big) > \varepsilon^2. \tag{S3.14}$$

  As shown in Ref. [23], whenever $\sigma$ is a pure stabilizer state, the operator $W \otimes \sigma$ defines a valid witness for $\hat{\mathcal{S}}_{4n}$. Applying Lemma 7 again yields $\|\rho_{\mathcal{E}\otimes I} - \sigma'\|_2 > \varepsilon^2/d/\|W\|_2$ for all $\sigma' \in \hat{\mathcal{S}}_{4n}$. Choosing $\sigma' = \tilde{\sigma} \otimes \phi^+_{2n}$ and using $\rho_{\mathcal{E}\otimes I} = \rho_\mathcal{E} \otimes \phi^+_{2n}$, we obtain

$$\|\rho_\mathcal{E} - \tilde{\sigma}\|_2 > \frac{\varepsilon^2}{d\|W\|_2} \geq \frac{\varepsilon^2}{d(d+2)}, \tag{S3.15}$$

  where we have used the fact that $\|W\|_2 \leq \|W'\|_2 + d \leq d + 2$ because $\max_\tau(W'\tau) \leq 1$ (see Lemma 7).

This establishes the relation $\varepsilon = \sqrt{d(d+2)}\,\delta$. Combining the above reductions with Theorem 6 concludes the proof. □

A similar problem can be defined when asking the question of whether a channel $\mathcal{E}$ is completely $t$-doping stabilizer preserving Definition 23 or not. An analogous decision problem can be defined by asking whether a quantum channel $\mathcal{E}$ is completely $t$-doping stabilizer preserving (see Definition 23).

**Problem 9** (Completely $t$-doped stabilizer-preserving classification (CTDSPC))**.** *Let $\mathcal{E}$ be a quantum channel and let $\varepsilon > 0$. Decide whether $\mathcal{E}$ satisfies:*

- *For all $\sigma \in \mathcal{S}_{2n}$, the state $[\mathcal{E} \otimes \mathbb{1}_n](\sigma)$ belongs to $S(\hat{\mathcal{S}}_{2n,t}, -\varepsilon)$; output* YES.
- *There exists $\sigma \in \mathcal{S}_{2n}$ such that $[\mathcal{E} \otimes \mathbb{1}_n](\sigma) \notin S(\hat{\mathcal{S}}_{2n,t}, \varepsilon)$; output* NO.

*An instance of this problem is denoted by* $\mathsf{CTDSPC}_\varepsilon$.

**Corollary 5** (Superpolynomial complexity of completely $t$-doped stabilizer-preserving classification)**.** *Let $\hat{\mathcal{S}}_{2n,t}$ denote the $t$-doped stabilizer polytope on $2n$ qubits. Under the* exponential time hypothesis *(ETH) (see Definition 5), for any*

$$\varepsilon \leq (2^{115}3^3(\log d)^{90}d^{118})^{-1}, \tag{S3.16}$$

*the decision problem* $\mathsf{CTDSPC}_\varepsilon$ *is not contained in the complexity class* $\mathsf{QP}^{2-\eta}$ *for any $\eta > 0$.*

*Proof.* The argument follows the same reduction strategy used for the completely stabilizer-preserving case. The only modification is that, instead of characterising completely stabilizer-preserving channels via their Choi states using Lemma 4, we invoke the Choi-state characterisation of completely $t$-doping-preserving channels given in Lemma 17. In particular, a YES instance of $\mathsf{CTDSPC}_\varepsilon$ implies that the Choi state of the channel lies within distance $\varepsilon$ of the $t$-doped stabilizer polytope $\hat{\mathcal{S}}_{2n,t}$, while a NO instance yields, via the same witness-based argument as before, a separation of order $\varepsilon/d$ from $\hat{\mathcal{S}}_{2n,t}$. This establishes a polynomial-time reduction from $\mathsf{CTDSPC}_\varepsilon$ to $\rho$-$\mathsf{WMEM}_{\varepsilon/d}(\hat{\mathcal{S}}_{2n,t})$. □

### 3. Existence of super-polynomial magic-state distillation

In this section, we apply our main hardness result to the task of magic-state distillation. Magic-state distillation protocols based on stabilizer operations constitute the standard route to achieving universal quantum computation. A natural question is whether the mere presence of non-stabilizer resources guarantees the existence of an efficient (in the Hilbert space dimension $d$) distillation procedure.

We show that this is not the case. More precisely, we prove that there exist non-stabilizer states for which no stabilizer-based distillation protocol that can be found and executed within polynomial time in $d$ is capable of producing a non-negligible amount of magic, even when allowed polynomially many copies of the input state. This result establishes a fundamental computational obstruction to magic-state distillation beyond information-theoretic considerations.

**Corollary 6** (Superpolynomial-time magic-state distillation)**.** *There exists a (possibly mixed) non-stabilizer state $\tilde{\rho}$ on $n$ qubits such that, for every stabilizer-based distillation protocol $\mathcal{E}$ satisfying the following conditions:*

1. *$\mathcal{E}$ can be specified by an algorithm running in time* $\mathsf{poly}(d)$*,*
2. *$\mathcal{E}$ can be executed in time* $\mathsf{poly}(d)$*,*

*the probability that $\mathcal{E}$ outputs a magic state is at most $\frac{1}{\omega(\mathsf{poly}(d))}$.*

*Proof.* Assume, for the sake of contradiction, that the statement is false. By Theorem 6, there exists a non-stabilizer state $\tilde{\rho}$ such that $\tilde{\rho}$ cannot be distinguished from the stabilizer polytope $\hat{\mathcal{S}}_n$ by any algorithm running in time $\mathsf{poly}(2^{\log^{2-\eta} d})$ for any $\eta > 0$. Equivalently, for any such algorithm, the distinguishing advantage between $\tilde{\rho}$ and any state in $\hat{\mathcal{S}}_n$ is negligible.

Under the above assumption, given $\tilde{\rho}$ one can proceed as follows. First, in time $\mathsf{poly}(d)$, compute the description of the corresponding distillation protocol $\mathcal{E}_{\tilde{\rho}}$. Next, prepare $\mathsf{poly}(d)$ copies of $\tilde{\rho}$ and apply $\mathcal{E}_{\tilde{\rho}}$ to obtain an output state. By assumption, this procedure produces a magic state with a probability of at least $1/\mathsf{poly}(d)$. Repeating this entire experiment $\mathsf{poly}(d)$ times yields an overall algorithm running in time $\mathsf{poly}(d)$ that outputs a magic state with a non-negligible probability whenever the input state is $\tilde{\rho}$. In contrast, if the input state is any $\sigma \in \hat{\mathcal{S}}_n$, then every stabilizer-based operation preserves $\hat{\mathcal{S}}_n$, thereby distinguishing $\tilde{\rho}$ from the stabilizer polytope in time $\mathsf{poly}(d)$. This contradicts the computational indistinguishability guaranteed by Theorem 6, implying our initial assumption to be false. □