

Polynomial-time sampling from peaked shallow quantum circuits with arbitrary connectivity

Yuxuan Zhang

Institute of Physics, École Polytechnique Fédérale de Lausanne (EPFL),
CH-1015 Lausanne, Switzerland

Department of Physics, Princeton University,
Princeton, New Jersey 08544, USA

22 October 2026 (version 1)

Abstract

For fixed positive integers d and k , we give a uniform randomized classical algorithm that samples the complete output distribution of a depth- d , n -qubit circuit whose largest output probability is at least n^{-k} . The circuit may have arbitrary connectivity, and a heavy output string need not be supplied. With bounded-degree algebraic gate entries of total description length b , the algorithm has worst-case runtime $\text{poly}_{d,k}(n, b)$ and total variation error at most any prescribed $\varepsilon \geq n^{-k}$. The proof combines a global L^1 estimate for a rooted inclusion–exclusion expansion with a positive proposal distribution that dominates the target probabilities. Its normalization is polynomial because the total connected-moment weight is $O(\log n)$. A small high-marginal core handles defects that are not uniformly rare. Positive dyadic upper weights and a fixed rejection cap make the bit-complexity and worst-case guarantees explicit. The exponents are large functions of d, k ; the result is an asymptotic sampling theorem rather than a practical simulation bound.

1 The sampling question

Shallow quantum circuits have small backwards light cones, but their complete output distributions can still contain global constraints. For peaked circuits, Bravyi, Gosset and Liu gave a quasipolynomial sampler without geometric locality assumptions; their polynomial sampling result uses two-dimensional locality [1, Theorems 1 and 2]. Coudron, Gullans, Nelson, Rajakumar and Tan subsequently obtained polynomial additive-error estimates of individual output probabilities [2, Theorem 1.1]. Such estimates do not by themselves yield a normalized distribution with small total variation error.

We use the rooted expansion and good-zero methodology of Ref. [2], with a direct derivation below. The additional steps are a bound on the *sum* of all output errors and a positive distribution from which proposals can be drawn efficiently. The proposal factors unconditional joint moments on separated sets. It does not assert a factorization of output probabilities, which would lose zero-event correlations such as parity constraints.

Theorem 1.1. *Fix positive integers d, k . Let U be an n -qubit circuit of depth at most d , consisting in each layer of disjoint one- or two-qubit gates, with arbitrary connectivity. The input is $|0^n\rangle$, and every output qubit is measured in the computational basis. Suppose that gate entries are exact*

algebraic numbers of fixed bounded degree, with total encoding length b , and that entries can be evaluated to any requested precision in polynomial time. Write

$$p_U(x) = |\langle x|U|0^n \rangle|^2, \quad \max_x p_U(x) \geq n^{-k}.$$

There is a uniform randomized classical sampler whose output law $q_{U,\varepsilon}$ satisfies

$$d_{\text{TV}}(q_{U,\varepsilon}, p_U) = \frac{1}{2} \sum_x |q_{U,\varepsilon}(x) - p_U(x)| \leq \varepsilon$$

for every rational $n^{-k} \leq \varepsilon \leq 1/4$. Its runtime and random-bit use are bounded in the worst case by $\text{poly}_{d,k}(n, b)$. The algorithm does not require a heavy output string.

We prove the stronger error target $e = n^{-k}$, independently of a larger requested ε . In particular, an unnecessarily long rational description of that larger tolerance need not be processed. Throughout, $\log$ is the natural logarithm. Constants may depend on d, k , but not on n, b . For a graph and vertex set A , $N[A]$ denotes the closed neighborhood of A . Small bounded n can be sampled by direct simulation, so we take $n \geq 2$.

2 A computable center and rare defects

Join two output vertices when their backwards input light cones overlap, and call the resulting graph G . Each backwards cone contains at most 2^d input qubits, and each such input can influence at most 2^d outputs. Thus G has maximum degree at most $4^d - 1$. Fix $K = 4^d$, increasing it to 2 if needed, and compute a proper K -coloring.

The graph has a strong dependency property. If two disjoint vertex sets have no edge between them, their complete measurement vectors are independent: the backwards-evolved projectors have disjoint supports on the product input. The same assertion holds for any collection of mutually separated sets. This is stronger than pairwise independence.

We will repeatedly use local marginal computation. For a specified event on s outputs, contraction of its backwards cone costs

$$\exp(O_d(s)) \text{ poly}(b, P) \tag{1}$$

bit operations at absolute precision 2^{-P} . The cone contains $O_d(s)$ qubits and gates. Evaluating each gate to $P + O_d(s + \log(P + 1))$ bits suffices, by telescoping the gate errors or bounding the finite contraction directly. There is no need to construct one common exact algebraic number field.

Let x_* be a heavy string, used only in this analysis. Within any color class I ,

$$\prod_{i \in I} \mathbb{P}(X_i = x_{*,i}) = \mathbb{P}(X_I = x_{*,I}) \geq \mathbb{P}(X = x_*) \geq n^{-k}.$$

Since $-\log(1 - u) \geq u$, summing over colors gives

$$\sum_i \mathbb{P}(X_i \neq x_{*,i}) \leq K k \log n.$$

Coordinatewise majority bits minimize individual mismatch probabilities. They are computable approximately without finding x_* .

Choose a fixed dyadic $\delta > 0$ with

$$2K^2 2^K \delta^{1/K} \leq \frac{1}{8}. \tag{2}$$

For example, $\delta = 2^{-K \lceil \log_2(16K^2 2^K) \rceil}$ works. Compute each one-bit marginal to error $\gamma = \min\{\delta/16, 1/(16n^2)\}$, and let b_i be its approximate majority bit. Put

$$Y_i = \mathbf{1}_{\{X_i \neq b_i\}}, \quad p_i = \mathbb{P}(Y_i = 1), \quad C = Kk + 4.$$

An approximate choice increases each mismatch probability by at most 2γ , whence

$$\sum_i p_i \leq C \log n. \quad (3)$$

Certified upper bounds of width at most 2γ for the p_i define the good set G_0 : an index is good when its upper bound is at most δ . Let $H = [n] \setminus G_0$. Then

$$p_i \leq \delta \ (i \in G_0), \quad p_i > \delta/2 \ (i \in H), \quad |H| \leq \frac{2C}{\delta} \log n. \quad (4)$$

No exact comparison of algebraic numbers is required.

Lemma 2.1. *Let $W = \sum_{i \in G_0} Y_i$. For each fixed $a \geq 1$,*

$$\mathbb{E} a^W \leq n^{D_a}, \quad D_a = \frac{C(a^K - 1)}{K}. \quad (5)$$

The same bound holds after restricting the sum to any subset of G_0 .

Proof. Write W_j for the sum in color class j . Hölder's inequality and independence within each class give

$$\mathbb{E} a^W \leq \prod_{j=1}^K (\mathbb{E} a^{KW_j})^{1/K} = \prod_{j=1}^K \prod_{i \in I_j \cap G_0} (1 + (a^K - 1)p_i)^{1/K}.$$

Use $1 + u \leq e^u$ and (3). Removing vertices only decreases the right-hand side. $\square$

Fix

$$h = \lceil (D_2 + k + 12) \log_2 n \rceil. \quad (6)$$

Markov's inequality yields

$$\mathbb{P}(W > h) \leq 2^{-h} \mathbb{E} 2^W \leq e/4096. \quad (7)$$

Ceilings of the fixed constants and $\ell = \lceil \log_2 n \rceil$ can be used throughout to obtain integer cutoffs without a real-number oracle.

Define the unconditional moment

$$\mu(A) = \mathbb{P}(Y_i = 1 \text{ for every } i \in A), \quad \mu(\emptyset) = 1.$$

Strong dependency implies

$$\mu(A) = \prod_{D \in \text{comp}(G[A])} \mu(D). \quad (8)$$

We use this identity for two different purposes: a signed expansion of zero events and a positive proposal envelope.

3 A good-zero evaluator with fixed analytic truncation

For $A \subseteq G_0$, set

$$Z_A(z) = \sum_{T \subseteq A} (-z)^{|T|} \mu(T), \quad Z_A = Z_A(1) = \mathbb{P}(Y_A = 0).$$

Equation (8) expresses this polynomial as a gas of nonempty connected sets, with activity $(-z)^{|T|} \mu(T)$; sets are compatible when disjoint and nonadjacent. This connected-set description and its use for good-zero events follow the approach of Ref. [2, Section 3.3]. We supply the bounds needed here.

Lemma 3.1. *For every $A \subseteq G_0$, $Z_A(z) \neq 0$ on $|z| \leq 2$. Its analytic logarithm f_A , normalized by $f_A(0) = 0$, satisfies*

$$|f_A(z)| \leq n \log 2 \quad (|z| \leq 2).$$

Proof. A largest color class in T has at least $|T|/K$ vertices, so $\mu(T) \leq \delta^{|T|/K}$. There are at most K^{2s} connected s -vertex sets through a specified vertex: choose a canonical spanning-tree traversal of length at most $2s$, pad by self-steps, and encode each step with at most K choices. Also $|N[T]| \leq K|T|$.

Induct on $|A|$, using, for $v \in A$, the exact deletion recurrence

$$Z_A(z) = Z_{A \setminus \{v\}}(z) + \sum_{\substack{T \subseteq A \text{ connected} \\ v \in T}} (-z)^{|T|} \mu(T) Z_{A \setminus N[T]}(z). \quad (9)$$

Assume the deletion ratios already established for smaller sets have modulus at least $1/2$. Deleting the other vertices of $N[T]$ bounds the quotient of the two smaller partition functions by $2^{|N[T]|-1}$. For $|z| \leq 2$,

$$\left| \frac{Z_A(z)}{Z_{A \setminus \{v\}}(z)} - 1 \right| \leq \sum_{s \geq 1} (2K^2 2^K \delta^{1/K})^s \leq \frac{1}{7} < \frac{1}{2}.$$

This proves the induction and nonvanishing. Each ratio lies in a disk centered at 1 of radius $1/7$, and its principal logarithm has modulus less than $\log 2$. Summing at most n such logarithms gives the stated analytic logarithm and bound. $\square$

Cauchy's estimate on the radius-two disk gives

$$\left| f_A(1) - \sum_{j=1}^r [z^j] f_A(z) \right| \leq n \log 2 \, 2^{-r}.$$

For $0 < \eta \leq 1/4$, choose

$$r = \lceil \log_2(8n/\eta) \rceil, \quad \widehat{Z}_A = \exp \left(\sum_{j=1}^r [z^j] f_A(z) \right). \quad (10)$$

The exponent is real, and exponentiating the error shows

$$|\widehat{Z}_A - Z_A| \leq \eta Z_A. \quad (11)$$

Lemma 3.2. *For inverse-polynomial η , the expression (10) has polynomial size and can be evaluated to additive precision 2^{-P} in $\text{poly}_{d,k}(n, b, P)$ time. Increasing P does not change r .*

Proof. For a finite set F , define the connected-support Möbius coefficient

$$c_j(F) = \sum_{D \subseteq F} (-1)^{|F|-|D|} [z^j] \log Z_D(z).$$

Möbius inversion gives

$$[z^j] f_A(z) = \sum_{F \subseteq A} c_j(F).$$

Only $|F| \leq j$ can contribute: a degree- j product of moments in the formal logarithm involves a union of at most j vertices. If F is disconnected, factorization of Z_D across its separated parts makes $\log Z_D$ additive, and the Möbius coefficient vanishes. It is therefore enough to enumerate connected sets of size at most r .

There are at most nK^{2r} traversal encodings. Deduplicate their vertex sets; for each resulting set, compute all local moments and the formal logarithms through degree r . The number of subsets and arithmetic operations per set is $\exp(O_d(r))$ times a polynomial. Equation (1) applies because every marginal uses at most r outputs. As $r = O_{d,k}(\log n)$, the expression has polynomial size.

For explicit precision control, write $\log(1 + Q) = \sum_{j=1}^r (-1)^{j+1} Q^j / j$ through degree r . The coefficient ℓ^1 -norm of Z_D is at most 2^r . Products of at most r such coefficient arrays have magnitudes and error amplification at most $\exp(O(r^2))$. Combining the polynomially many connected sets adds $O(\log n)$ guard bits. Thus $P + O(r^2 + \log n)$ local-moment precision suffices. Evaluate the real exponential by range reduction and rational Taylor arithmetic. Its argument is bounded in modulus by $n \log 2 + 1$, and at the evaluation point it differs from $\log Z_A \leq 0$ by less than one. These bounds give polynomial bit complexity.

The analytic cutoff r is chosen once from η . Evaluating this same finite expression more accurately does not invoke a new approximation algorithm at a smaller analytic tolerance. $\square$

4 Rooted probabilities and a global L^1 estimate

Represent an output by its good defect set $S \subseteq G_0$ and its complete bit pattern $u \in \{0, 1\}^H$. Let $p(S, u)$ be its probability. Only $|S| \leq h$ will be retained. Put $B = H \cup S$. A set $T \subseteq G_0 \setminus S$ is *rooted at B* if every component of $G[T]$ has an edge to B . Define

$$A_T = [n] \setminus N[B \cup T] \subseteq G_0.$$

Lemma 4.1 (Rooted identity). *For every S, u ,*

$$p(S, u) = \sum_{\substack{T \subseteq G_0 \setminus S \\ T \text{ rooted at } B}} (-1)^{|T|} \mathbb{P}(Y_H = u, Y_{S \cup T} = 1) Z_{A_T}. \quad (12)$$

Proof. Apply inclusion–exclusion to the required zero bits in $G_0 \setminus S$. For each selected set, distinguish components adjacent to B from the remaining components. Let T be the union of the first group. Every remaining component lies outside $N[B \cup T]$. Its event is independent of the joint event on $B \cup T$, by strong dependency. Summing all the remaining signed selections gives Z_{A_T} . Conversely every pair consisting of a rooted T and a selection in A_T occurs once, proving the identity. This is the rooted decomposition of Ref. [2, Eqs. (7)–(10)], written for the present center. No independence conditional on observed bits is used. $\square$

Truncate (12) to $|T| \leq R$, calling the result p_R . Replace each Z_{A_T} by $\hat{Z}_{A_T}$, with the same η, r , and call the result $\hat{p}(S, u)$. For $|B|, R = O(\log n)$, this can be evaluated in polynomial time. Indeed,

assign each rooted component to its first adjacent vertex of B , distribute the total sizes among roots, and enumerate a connected animal at each root. The number of encodings is $\exp(O_K(|B| + R))$: size compositions contribute at most $2^{|B|+R}$, and rooted walks contribute $K^{O(|B|+R)}$. If $B = \emptyset$, only $T = \emptyset$ is admissible. All marginal events use $O(|H| + h + R)$ outputs.

Proposition 4.2. *The complete sums of errors over retained outputs obey*

$$\sum_{\substack{|S| \leq h \\ u}} |p(S, u) - p_R(S, u)| \leq 2^{-R} n^{D_5}, \quad (13)$$

$$\sum_{\substack{|S| \leq h \\ u}} |p_R(S, u) - \hat{p}(S, u)| \leq \eta n^{D_3}. \quad (14)$$

Proof. Take absolute values only after expanding each discarded contribution. Since $0 \leq Z_{A_T} \leq 1$, all coefficients to be bounded are nonnegative. Regroup the disjoint pair S, T by $D = S \cup T$. There are at most $2^{|D|}$ possible choices of S ; summing over the complete H -pattern u gives $\mu(D)$. Dropping rootedness and $|S| \leq h$ can only enlarge the sum. A discarded T has $|D| > R$. Consequently

$$\begin{aligned} \sum_{\substack{|S| \leq h \\ u}} |p(S, u) - p_R(S, u)| &\leq \sum_{\substack{D \subseteq G_0 \\ |D| > R}} 2^{|D|} \mu(D) \\ &\leq 2^{-R} \sum_{D \subseteq G_0} 4^{|D|} \mu(D) = 2^{-R} \mathbb{E} 5^W. \end{aligned}$$

The last equality expands $\prod_{i \in G_0} (1 + 4Y_i)$. This proves (13). For the zero-evaluator contribution, use (11) and $Z_{A_T} \leq 1$; the same regrouping gives

$$\sum_{\substack{|S| \leq h \\ u}} |p_R - \hat{p}| \leq \eta \sum_{D \subseteq G_0} 2^{|D|} \mu(D) = \eta \mathbb{E} 3^W.$$

Apply Lemma 2.1. □

Choose

$$R = \lceil (D_5 + k + 12) \log_2 n \rceil, \quad \eta = n^{-D_3 - k - 12}. \quad (15)$$

Larger integer exponents or conservative dyadic tolerances may be used. Each bound in Proposition 4.2 is at most $e/4096$. Extending $\hat{p}$ by zero outside $|S| \leq h$ and using (7), we obtain

$$\|\hat{p} - p\|_1 < e/1024. \quad (16)$$

The estimate is global. It has no multiplicative factor equal to the number of strings in a logarithmic-radius Hamming ball. At this stage $\hat{p}$ may have negative entries, so a sampling construction is still required.

5 A positive envelope with polynomial normalization

List all nonempty connected $C' \subseteq G_0$ with $|C'| \leq h$. Call these sets polymers. Their number M_{poly} is polynomial, by the connected-set count.

Lemma 5.1. *The sum of all connected good-set moments satisfies*

$$\Lambda = \sum_{\substack{C' \subseteq G_0 \\ C' \neq \emptyset \text{ connected}}} \mu(C') \leq \delta^{-1} C \log n. \quad (17)$$

Proof. Fix $v \in C'$, with $|C'| = s$. Remove $N[v]$, which has at most K vertices, and take a largest color class in the remainder. These vertices are jointly independent and their joint event is independent of Y_v . Hence

$$\mu(C') \leq p_v \delta^{\max\{0, \lceil (s-K)/K \rceil\}} \leq p_v \delta^{-1} \delta^{s/K}.$$

Overcount each C' at every possible root v , and use the traversal bound to get

$$\Lambda \leq \delta^{-1} \sum_v p_v \sum_{s \geq 1} (K^2 \delta^{1/K})^s \leq \delta^{-1} C \log n,$$

where (2) bounds the geometric series by one. $\square$

The high-marginal core consists of pairs (J, D) , with $J \subseteq H$, $D \subseteq G_0$, $|D| \leq h$, such that every component of $G[D]$ has an edge to H . The empty D is allowed and is the only choice when $H = \emptyset$. Give this pair the moment weight $\mu(J \cup D)$. The rooted enumeration above gives

$$M_{\text{core}} \leq 2^{|H|} \exp(O_K(|H| + h)) = \text{poly}_{d,k}(n).$$

The sets $J \cup D$ are distinct, and therefore

$$A_{\text{exact}} = \sum_{\text{cores}} \mu(J \cup D) \leq \sum_{F \subseteq [n]} \mu(F) = \mathbb{E} 2^{\sum_i Y_i} \leq 2^{|H|} n^{D_2}. \quad (18)$$

Tiny or zero moments must not create an expensive division. Choose a dyadic ν satisfying

$$\frac{1}{32(M_{\text{core}} + M_{\text{poly}} + 1)} \leq \nu \leq \frac{1}{16(M_{\text{core}} + M_{\text{poly}} + 1)}.$$

Bracket each needed moment to width at most $\nu/4$, round its upper endpoint upwards to a multiple of ν , and add ν . For every resulting rational upper weight a ,

$$\nu \leq a, \quad \mu \leq a \leq \mu + 3\nu. \quad (19)$$

Let $a_{J,D}$ and $a_{C'}$ be the core and polymer weights, and put

$$A = \sum_{\text{cores}} a_{J,D}, \quad K_{\text{prop}} = A \prod_{C'} (1 + a_{C'}). \quad (20)$$

The aggregate additive slack is less than one. By (17) and (18),

$$A \leq 2^{|H|} n^{D_2} + 1, \quad \prod_{C'} (1 + a_{C'}) \leq e n^{C/\delta}.$$

With

$$J_* = \lceil 4C/\delta + D_2 + 10 \rceil, \quad K_{\text{prop}} \leq n^{J_*} \quad (n \geq 2). \quad (21)$$

For example, (4) implies $2^{|H|} \leq n^{2C/\delta}$, which gives the displayed conservative exponent. All weights have $O_{d,k}(\log n)$ -bit descriptions; the polynomial number of factors in (20) gives polynomial total bit length.

A proposal attempt draws a core with law $a_{J,D}/A$, and independently activates every polymer C' with probability $a_{C'}/(1 + a_{C'})$. The attempt is invalid if active polymers overlap or share an edge, if an active polymer meets $N[H \cup D]$, or if their union with D has more than h vertices. Otherwise, set

$$S = D \cup \bigcup_{\text{active } C'} C',$$

and output pattern J on H , before the further acceptance step.

Proposition 5.2 (Unique representation and domination). *Every output with $|S| \leq h$ has exactly one valid core/polymer representation. Its unconditional proposal probability equals*

$$\frac{w(S, J)}{K_{\text{prop}}}, \quad w(S, J) = a_{J,D} \prod_{\text{active } C'} a_{C'} \geq p(S, J). \quad (22)$$

Proof. The components of $G[S]$ that have an edge to H must belong to D ; all others must be the active polymers. This choice is valid, and the validity rules force it, proving uniqueness.

Every active component is separated from $H \cup D$ and from the other active components. Thus (8) gives

$$\mu(J \cup S) = \mu(J \cup D) \prod_{\text{active } C'} \mu(C').$$

The event for the exact output is contained in the event defining the left side, so (19) proves $w(S, J) \geq p(S, J)$. Finally, the core probability and all activation/nonactivation probabilities have common denominator $A \prod_{C'} (1 + a_{C'})$; their numerator is precisely $w(S, J)$. $\square$

This construction uses positive unconditional moments as an envelope. It does not sample the signed activities in the zero-event expansion. Invalid proposals simply count as failed attempts.

6 Acceptance, finite precision and a worst-case cap

Given a valid proposal, accept it with ideal probability

$$\text{clip}\left(\frac{\hat{p}(S, J)}{w(S, J)}, 0, 1\right).$$

The accepted, unnormalized mass multiplied by K_{prop} is

$$r(S, J) = \text{clip}(\hat{p}(S, J), 0, w(S, J)).$$

Since the true probability lies in $[0, w(S, J)]$, clipping cannot increase its pointwise error. Set $r = 0$ outside the support cutoff. By (16),

$$\|r - p\|_1 < e/1024, \quad Z_r := \sum_{S, J} r(S, J) \geq 1 - e/1024.$$

Moreover,

$$d_{\text{TV}}(r/Z_r, p) \leq \frac{1}{2}(\|r/Z_r - r\|_1 + \|r - p\|_1) \leq \|r - p\|_1.$$

Thus the conditional accepted law is close to the target, and its success probability per attempt is $Z_r/K_{\text{prop}} \geq 1/(2n^{J^*})$.

Use at most

$$T_{\text{max}} = \left\lceil 2n^{J^*}(k+5)\lceil \log_2 n \rceil \right\rceil \quad (23)$$

attempts. If they all fail, output the center b . Since the exponent in the failure bound is at least $\log(16/e)$, the fallback probability is at most $e/16$.

We now replace each ideal random choice by a bounded number of fair bits. Approximate the rational core cumulative distribution with one dyadic uniform number, to total variation error at most $e/(32T_{\text{max}})$ per attempt. Round each activation and acceptance probability to a fixed dyadic grid. Choose enough bits that the sum of all these per-choice errors over at most $T_{\text{max}}(M_{\text{poly}} + 2)$

choices is at most $e/16$. A stepwise coupling, including adaptive acceptance decisions, bounds the resulting output-law error by that sum. Hence no unbounded exact Bernoulli routine is used.

The numerical acceptance calculation also has bounded cost. A valid proposal has at most h active polymers, so

$$w(S, J) \geq \nu^{h+1} = \exp(-O_{d,k}(\log^2 n)). \quad (24)$$

It follows that $O_{d,k}(\log^2 n)$ bits of absolute precision in the numerator suffice for the inverse-polynomial acceptance tolerance. Evaluate the already fixed expression $\hat{p}$ to this accuracy. Lemma 3.2, the polynomial number of rooted terms, and (1) give polynomial bit complexity. Crucially, this extra arithmetic accuracy changes neither r nor R in the analytic truncations.

The ideal accepted-law error, fallback probability and finite-bit implementation error sum to less than e . Every proposal list, contraction, arithmetic expression and loop cap has size polynomial in n, b for fixed d, k . This proves Theorem 1.1.

7 Scope and verification

The sampler is an existence result with explicit finite computation. The constants in δ , the truncation orders and the cap (23) lead to very large polynomial degrees. No efficient dependence on growing depth or on a growing peakedness exponent is claimed. The theorem also does not address exponentially small target errors.

Exact rational finite controls were used to check rooted identities, connected-moment factorizations, global error regrouping, unique proposal representations and domination. They include a noncommuting shallow circuit and a parity-constrained law. A separate six-cycle dependency law was checked with a genuine support truncation and nonzero invalid proposal probability. These calculations support the algebra; the analytic arguments above establish the asymptotic and bit-complexity claims.

AI assistance and verification status. This work was developed and written in an author-directed research workflow using OpenAI Codex. AI assistance included mathematical derivation, literature comparison, computational checks and manuscript drafting. Separate agent reviews within the same system examined the complete proof and targeted controls. These internal reviews are fallible and are not external expert review or formal verification. The manuscript reports the proof and its dependencies directly; no external acceptance is claimed.

References

- [1] S. Bravyi, D. Gosset and Y. Liu, *Classical simulation of peaked shallow quantum circuits*, arXiv:2309.08405v1 (2023). [arXiv:2309.08405v1](#).
- [2] M. Coudron, M. J. Gullans, J. Nelson, J. Rajakumar and S. J. S. Tan, *Polynomial-time additive-error estimation of output probabilities for shallow quantum circuits*, arXiv:2610.02146v1 (2026). [arXiv:2610.02146v1](#).