

Provable quantum–classical separation for continuous Gibbs sampling

Enrico Olivucci^{*,1}, Mariia Sobchuk^{*,2,3}, Sehmimul Hoque^{*,2,3,4}, Jeffrey Hnybida¹,
William Kim^{2,3}, Ala Shayeghi^{2,5}, Pooya Ronagh^{2,3,4,6}

¹Irréversible Inc., Sherbrooke, QC, Canada ²Institute for Quantum Computing, Univ. of Waterloo, Canada

³Dept. of Physics and Astronomy, Univ. of Waterloo, Canada ⁴Perimeter Institute for Theoretical Physics

⁵National Research Council Canada ⁶Microsoft, Redmond, WA, USA *Equal contribution.

Motivation. Establishing rigorous separations between quantum and classical computation is a central objective of quantum algorithms research, yet proven separations are rare. Previous separations — Grover’s unstructured search [1], the hidden-structure problems of Bernstein–Vazirani and Simon [2, 3], promise problems such as Deutsch–Jozsa and the collision problem [4–6], and quantum-walk speedups for welded trees and element distinctness [7, 8] — all concern discrete (and often contrived) computational tasks. For Gibbs sampling, a fundamental task in statistical physics [9], computational science, and machine learning [10–12], no analogous separation over continuous domains has, to our knowledge, been established. For non-convex, multimodal potentials V , classical samplers such as the unadjusted Langevin algorithm and other Markov chain Monte Carlo methods [13] suffer mixing times exponential in the barrier heights [14, 15]. Quantum Gibbs samplers for continuous potentials have recently been proposed [16–19] (see [20–22] for the discrete setting), but comparing against specific classical algorithms cannot rule out faster ones. We establish a quantum–classical separation for Gibbs sampling on the torus $\mathbb{T}^d$ for densities in a Gevrey regularity class: we prove an information-theoretic lower bound on the query complexity of *every* classical sampling algorithm interacting with *any local oracle* for the target distribution, and exhibit a quantum Gibbs sampler whose query complexity falls below this lower bound by a factor growing exponentially in the dimension d .

Setting and main result. Fix $s > 0$ and let $t = 1 + 1/s$. A t -Gevrey probability distribution on $\mathbb{T}^d$ [23, 24] is a measure $\pi \ll \mu$ (μ the uniform measure) whose density $\rho = d\pi/d\mu$ is smooth, positive, and satisfies $|\partial^a \rho| \leq (a!)^t C r^{-\|a\|_1}$ for all multi-indices a and some amplitude $C > 0$ and radius $r > 0$. Every such distribution is a Gibbs state $\rho = e^{-\beta V}/Z$ for a smooth periodic potential V , and the parameter $\alpha := \max \rho / \min \rho = e^{\beta \text{osc} V}$ measures its non-log-concavity: it upper bounds the Poincaré constant, $C_{\text{PI}} \leq e^{\beta \text{osc} V}$ (a Holley–Stroock-type comparison [25, 26]), which governs the mixing of Langevin dynamics. A classical algorithm accesses π through a *local oracle* — any deterministic oracle whose answer at a query point x depends only on the germ of ρ at x ; this class contains every finite-jet oracle returning $\rho(x), \nabla \rho(x), \dots, \nabla^{\otimes k} \rho(x)$, hence zeroth- and first-order queries to V . The quantum algorithm makes queries to ∇V ; query complexity is minimax over the instance class, to accuracy ε in total variation. Our main result:

Theorem 1 (informal). *For every $s > 0$ there is a family of t -Gevrey Gibbs distributions on $\mathbb{T}^d$, $t = 1 + 1/s$, with $\log \alpha = \beta \text{osc} V = \Omega(d)$, such that:*

- (i) (Classical lower bound) *every classical randomized algorithm with access to any local oracle requires $N_{\text{cl}} = \Omega(\alpha) = \Omega(e^{\beta \text{osc} V})$ queries to sample to constant total-variation accuracy;*

(ii) (Quantum upper bound) a QSVT-based quantum algorithm with temperature annealing samples to accuracy ε using

$$N_q = \tilde{O}\left(e^{\beta E_M} \beta^{2.5} E_M^2 \left(N\sqrt{d} E_M + R\right)\right) = \alpha^{\frac{1}{2} + \frac{p}{d}} \text{poly}(d^s, \beta, \log^s(\alpha/\varepsilon))$$

queries to ∇V , where $E_M = \frac{1}{2} \text{osc} V$, $R = \sup_x \|\nabla V(x)\|_2$, N is the grid size (polylogarithmic in $1/\varepsilon$), and p is the constant power of the inverse Gevrey radius in the quantum bound.

Consequently $N_{\text{cl}}/N_q = \Omega(\alpha^{c(d)}/(\text{poly}(d) \text{polylog}(\alpha, \frac{1}{\varepsilon})))$ with $c(d) = \frac{1}{2} - \frac{p}{d}$ for $d > 2p$; in the regime $\log \alpha = \Omega(d)$ and $d = \Omega(\log \log(1/\varepsilon))$ this yields

$$\frac{N_{\text{cl}}}{N_q} = \Omega(e^{d/2}).$$

The lower bound (i) is *regularity-agnostic*: it holds for every class containing the hard family, and the Gevrey hypothesis enters only to certify that the hard instances are as smooth as the quantum algorithm requires; both sides thus concern the same instances under comparable oracles.

Classical lower bounds from local oracles. Our lower-bound technique is an adversarial coupling argument in the information-based complexity framework [27–29], in the tradition of the local-oracle lower bounds of convex optimization [30], formalized measure-theoretically for the sampling task (oracles as sectionwise-measurable response maps, algorithms as probability spaces with non-anticipating query maps, output laws as pushforwards). The hard family consists of *bump distributions*: perturbations of the uniform measure that are constant off a small patch $\Omega_i \subset \mathbb{T}^d$ and concentrate mass $\delta = \pi_i(\Omega_i)$ on it. Locality of the oracle implies that all bumps in the family answer identically — with a well-defined *null response* — at any point lying outside their patch, so a coupling of each run against the null run shows that an algorithm learns nothing about the planted instance until a query lands in its patch. Averaging the per-instance inequalities against a probability measure ν on the index set whose *patch density* obeys $\sup_x \nu(\{i \mid x \in \Omega_i\}) \leq 1/M$ yields the algorithm-independent bound

$$N \geq M(\bar{\delta} - \varepsilon) - 1, \quad \bar{\delta} = \mathbb{E}_{i \sim \nu}[\delta_i], \quad \text{for every } \varepsilon < \delta_i \text{ } \nu\text{-a.e.}$$

The measure ν is an optimization device: for equal-measure patches the uniform coverage is optimal and attains $M = \mu(\Omega)/\mu(\Omega_0)$. On a compact abelian group the Haar-indexed family of *all* translates of a single cell realizes this optimum, and the resulting covering bound provably dominates every bound obtainable from a packing, strictly unless the packing already tiles the group (cf. [31]). The argument prices only oracle calls and quantifies over any oracle constrained only by locality; moreover it identifies exactly where regularity matters: for the ∞ -jet oracle the null response must consist of *flat* jets coexisting with non-trivial bumps, which forces non-quasi-analyticity. This is precisely where the Gevrey hypothesis $t > 1$ enters, and it shows the construction cannot survive in the analytic class ($t = 1$), consistently with the identity theorem. Query lower bounds for sampling itself are scarce: prior results concern log-concave targets [32, 33] or, in the non-log-concave regime, the weaker Fisher-information accuracy [34] (cf. [35]); to our knowledge, ours are the first in total variation with exponential-in- d hardness, and the first against arbitrary local oracles.

Hard instances: Gevrey bumps with explicit constants. To place the bump family inside the t -Gevrey Gibbs class we construct it explicitly: starting from the function $\psi_s(x) = \exp(-1/(1-x^2)^s)$, whose membership in $G^t(1, r)$ with an explicit radius $r = r(s)$ is proven via a contour-optimized Cauchy estimate, we build a mollified indicator $\chi_{\nu, \ell}$ of the cube $[0, \ell]^d$ with transition width ν , and set $\rho \propto 1 + (\alpha - 1)\chi_{\nu, \ell}$. All Gevrey constants of ρ are explicit in (α, d, ℓ, ν) , as is the

	Model	Oracle access	Query complexity
This work (Thm. 1(i))	classical, any algorithm	any local oracle	$\Omega(e^{\beta \text{osc} V})$
This work (Thm. 1(ii))	quantum, QSVT + annealing	∇V	$\tilde{O}(e^{\beta \text{osc} V/2} \text{poly}(d^s, \beta))$
Langevin / MRW (via C_{PI})	classical, specific	∇V	$O(e^{\beta \text{osc} V} \text{poly}(d, \frac{1}{\varepsilon}))$

Table 1: Query complexities at constant accuracy on the hard family of Theorem 1 ($\beta \text{osc} V = \log \alpha$). The lower bound applies to *every* classical algorithm and every local oracle; the quantum algorithm beats it by $e^{\beta \text{osc} V/2} = \Omega(e^{d/2})$.

gradient bound for the associated Gibbs potential, $R \leq \frac{4s\sqrt{d}}{\beta\nu} \log^t(\frac{2(\alpha-1)}{Z_s\nu})$ with $Z_s = \int_{-1}^1 \psi_s(y) dy$. The lower bound then reads $N_{\text{cl}} \geq M(\delta - \varepsilon) - 1$ with $M = \ell^{-d}$, and the separation is governed by a two-variable trade-off between the packing number M and the mass budget $S := (\alpha - 1)(1 - 2\kappa_1 \varrho)^d$, where $\varrho = \nu/\ell$ and $\kappa_1 \in (0, 1)$ is an explicit constant of the mollifier. Optimizing the ratio of the lower to the upper bound is a clean two-step problem: at fixed ϱ the optimal cell count satisfies $M_* \asymp S \min\{\varepsilon^{-1/2}, d/p\}$ with concentration $\delta_* \asymp \max\{\sqrt{\varepsilon}, p/d\}$; in turn, the logarithmic profile in ϱ is strictly concave with unique maximizer $\varrho_* = p/(2\kappa_1 d)$. At the optimum, $\ell_* \asymp \alpha^{-1/d}$ and $\nu_* \asymp \alpha^{-1/d}/d$, giving the exponent $c(d) = \frac{1}{2} - \frac{p}{d}$ of Theorem 1; the same $e^{d/2}$ asymptotics is already achieved, with slightly worse constants, by the sub-optimal choice $\delta = \Theta(1)$, $M \asymp \alpha$.

The quantum upper bound. The quantum algorithm combines the QSVT-based [36] Gibbs sampler of Leng, Ding, Chen, and Lin [17] with a temperature-annealing schedule. That algorithm filters the singular values of the discretized Witten Laplacian H , whose kernel encodes the target $\sqrt{\sigma} \propto e^{-\beta V/2}$, but it requires a warm start (non-negligible initial overlap with the target) and rests on an unproven spatial-discretization assumption [17, Assumption 11]. We resolve both. First, we prove the discretization assumption for the class at hand: comparing H with the discretized Fokker–Planck generator of [16] via exponentially convergent Fourier pseudo-spectral error bounds for Gevrey functions, we show the two operators are $\tilde{\varepsilon}$ -close in operator norm at grid size $N = \tilde{\Omega}((d \log \frac{dRDC}{\beta\tilde{\varepsilon}})^s)$, with $D = \max\{e^{-\beta V_{\min}}, e^{\beta V_{\max}}\}$ and C the Gevrey amplitude; since the latter has one-dimensional kernel spanned by the discretized Gibbs state and spectral gap at least $e^{-\beta \text{osc} V}$, Weyl and Davis–Kahan perturbation bounds [37] transfer both properties to H , yielding an assumption-free QSVT threshold, implemented via an erf-based polynomial approximation of the sign function [38]. Second, to remove the warm start we run an annealing schedule [18, 39] of $\beta^2 E_M^2$ steps with increments $\delta\beta = 1/(\beta E_M^2)$: consecutive Gibbs states overlap by at least $1/\cosh(\delta\beta E_M/2)$, so each state warm-starts the next; the overlap survives the per-step filtering error, and the algorithm converges from a uniform superposition to the target at accuracy ε with constant success probability. The end-to-end gradient-query complexity is $\tilde{O}(e^{\beta E_M} \beta^{2.5} E_M^2 (N\sqrt{d} E_M + R))$.

Impact. To our knowledge, this establishes the first provable quantum–classical separation for a sampling problem over a continuous domain — of intrinsic interest for quantum computing, where unconditional separations are notoriously difficult to obtain. It also concerns a regime of practical relevance: Gibbs sampling from non-log-concave distributions is an active research area in machine learning [10, 11], underlying training and inference in energy-based models, Boltzmann machines, and related models, with scalable samplers as the computational bottleneck. Our result shows that quantum algorithms accelerate sampling precisely where the task is provably hard for *every* classical algorithm, not merely harder than known methods. The result cuts both ways: for quantum algorithms, a natural continuous task with an unconditional query advantage; for machine learning, quantum acceleration certified where classical sampling is intractable.

References

- [1] L. K. Grover. A fast quantum mechanical algorithm for database search. In *Proceedings of the 28th Annual ACM Symposium on Theory of Computing (STOC)*, pages 212–219, 1996.
- [2] E. Bernstein and U. Vazirani. Quantum complexity theory. *SIAM Journal On Computing*, **26**, 1411–1473 (1997), <https://doi.org/10.1137/S0097539796300921>
- [3] D. R. Simon. On the power of quantum computation. In *Proceedings of the 35th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 116–123, 1994.
- [4] D. Deutsch and R. Jozsa. Rapid solution of problems by quantum computation. *Proceedings of the Royal Society of London A*, 439(1907):553–558, 1992.
- [5] G. Brassard, P. Høyer, and A. Tapp. Quantum cryptanalysis of hash and claw-free functions. In *LATIN’98: Theoretical Informatics*, Lecture Notes in Computer Science 1380, pages 163–169. Springer, 1998.
- [6] Y. Shi. Quantum lower bounds for the collision and the element distinctness problems. In *Proceedings of the 43rd Annual Symposium on Foundations of Computer Science (FOCS)*, pages 513–519, 2002.
- [7] A. M. Childs, R. Cleve, E. Deotto, E. Farhi, S. Gutmann, and D. A. Spielman. Exponential algorithmic speedup by a quantum walk. In *Proceedings of the 35th Annual ACM Symposium on Theory of Computing (STOC)*, pages 59–68, 2003.
- [8] A. Ambainis. Quantum walk algorithm for element distinctness. *SIAM Journal on Computing*, 37(1):210–239, 2007.
- [9] D. Ruelle. *Statistical Mechanics: Rigorous Results*. World Scientific, 1999.
- [10] Y. Du and I. Mordatch. Implicit generation and generalization in energy-based models. arXiv:1903.08689, 2019.
- [11] Y. Song and D. P. Kingma. How to train your energy-based models. arXiv:2101.03288, 2021.
- [12] J. Ho, A. Jain, and P. Abbeel. Denoising diffusion probabilistic models. *Advances in Neural Information Processing Systems*, 33:6840–6851, 2020.
- [13] R. Dwivedi, Y. Chen, M. J. Wainwright, and B. Yu. Log-concave sampling: Metropolis–Hastings algorithms are fast! In *Conference on Learning Theory*, pages 793–797. PMLR, 2018.
- [14] A. Bovier, M. Eckhoff, V. Gayrard, and M. Klein. Metastability in reversible diffusion processes I: Sharp asymptotics for capacities and exit times. *Journal of the European Mathematical Society*, 6:399–424, 2004.
- [15] N. Berglund. Kramers’ law: validity, derivations and generalisations. arXiv:1106.5799, 2011.
- [16] A. Motamedi and P. Ronagh. Gibbs sampling of continuous potentials on a quantum computer. *Proceedings of Machine Learning Research*, 235:36322–36371, 2024.
- [17] J. Leng, Z. Ding, Z. Chen, and L. Lin. Operator-level quantum acceleration of non-logconcave sampling. arXiv:2505.05301, 2025.
- [18] G. Ozgul, X. Li, M. Mahdavi, and C. Wang. Stochastic quantum sampling for non-logconcave distributions and estimating partition functions. *Proceedings of Machine Learning Research*, 235:38953–38982, 2024.
- [19] A. M. Childs, T. Li, J.-P. Liu, C. Wang, and R. Zhang. Quantum algorithms for sampling log-concave distributions and estimating normalizing constants. *Advances in Neural Information Processing Systems*, 35, 2022.
- [20] B. M. Terhal and D. P. DiVincenzo. Problem of equilibration and the computation of correlation functions on a quantum computer. *Physical Review A*, 61:022301, 2000.
- [21] D. Poulin and P. Wocjan. Sampling from the thermal quantum Gibbs state and evaluating partition functions with a quantum computer. *Physical Review Letters*, 103:220502, 2009.
- [22] A. N. Chowdhury and R. D. Somma. Quantum algorithms for Gibbs sampling and hitting-time estimation. arXiv:1603.02940, 2016.
- [23] M. Gevrey. Sur la nature analytique des solutions des équations aux dérivées partielles. Premier mémoire. *Annales scientifiques de l’École Normale Supérieure*, 35:129–190, 1918.
- [24] L. Rodino. *Linear Partial Differential Operators in Gevrey Spaces*. World Scientific, 1993.
- [25] R. Holley and D. Stroock. Logarithmic Sobolev inequalities and stochastic Ising models. *Journal of Statistical Physics*, 46:1159–1194, 1987.
- [26] D. Bakry, I. Gentil, and M. Ledoux. *Analysis and Geometry of Markov Diffusion Operators*. Springer,

- 2014.
- [27] J. F. Traub, G. W. Wasilkowski, and H. Woźniakowski. *Information-Based Complexity*. Academic Press, 1988.
 - [28] E. Novak. *Deterministic and Stochastic Error Bounds in Numerical Analysis*. Lecture Notes in Mathematics 1349, Springer, 1988.
 - [29] E. Novak and H. Woźniakowski. *Tractability of Multivariate Problems. Volume I: Linear Information*. European Mathematical Society, 2008.
 - [30] A. S. Nemirovsky and D. B. Yudin. *Problem Complexity and Method Efficiency in Optimization*. Wiley, 1983.
 - [31] C. A. Rogers. *Packing and Covering*. Cambridge University Press, 1964.
 - [32] S. Chewi, P. R. Gerber, C. Lu, T. Le Gouic, and P. Rigollet. The query complexity of sampling from strongly log-concave distributions in one dimension. In *Conference on Learning Theory*, PMLR 178:2041–2059, 2022.
 - [33] S. Chewi, J. de Dios Pont, J. Li, C. Lu, and S. Narayanan. Query lower bounds for log-concave sampling. In *64th IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 2139–2148, 2023. *Journal of the ACM*, 71(4):29, 2024.
 - [34] S. Chewi, P. Gerber, H. Lee, and C. Lu. Fisher information lower bounds for sampling. In *Algorithmic Learning Theory*, PMLR 201:375–410, 2023.
 - [35] K. Talwar. Computational separations between sampling and optimization. *Advances in Neural Information Processing Systems*, 32, 2019.
 - [36] A. Gilyén, Y. Su, G. H. Low, and N. Wiebe. Quantum singular value transformation and beyond: exponential improvements for quantum matrix arithmetics. In *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing (STOC)*, pages 193–204, 2019.
 - [37] Y. Yu, T. Wang, and R. J. Samworth. A useful variant of the Davis–Kahan theorem for statisticians. *Biometrika*, 102:315–323, 2015.
 - [38] G. H. Low and I. L. Chuang. Hamiltonian simulation by uniform spectral amplification. arXiv:1707.05391, 2017.
 - [39] R. Somma, S. Boixo, and H. Barnum. Quantum simulated annealing. arXiv:0712.1008, 2007.