

On Quantum Advantage in Sampling Using Instantaneous Quantum Polynomial Circuits

Afrad Basheer¹, Mario Ponce¹, Martin Leib¹, and Thomas Cope¹

¹*IQM Germany GmbH, Georg-Brauchle-Ring 23-25, 80992 München, Germany*

{Afrad.Basheer, Mario.Ponce, Martin.Leib, Thomas.Cope}@iqm.tech

Abstract

The field of quantum generative models based on Instantaneous Quantum Polynomial (IQP) circuits has seen significant interest in recent years, driven by results proving classical difficulty of sampling from sparse anti-concentrated IQP circuits and efficient trainability. Practical demonstrations of the former are crucial in establishing the superiority of these circuits in generating samples over classical methods. In this work, we show that Linear Cross Entropy benchmarking, the most widely considered test to demonstrate quantum supremacy, using IQP circuits can be spoofed efficiently and argue that beating the spoofer using quantum computers will require near fault-tolerant levels of fidelities, which puts such experiments firmly beyond the capabilities of current hardware. Furthermore, we introduce an algorithm that can compute amplitudes of large sparse IQP circuits consisting of 30-50 qubits in under 1 second on GPUs, which can be leveraged to generate samples in under a minute. As a final contribution, we show how one can generate peaked IQP circuits defined on hundreds of qubits efficiently, using which one can set up verifiable supremacy tasks, thus providing a better path towards the experimental demonstration of quantum supremacy using IQP circuits. Finally, we argue that such peaked circuits can be the cornerstone of classical difficulty arguments for IQP-based generative models, since we believe it is unlikely that flat anti-concentrated distributions, which are challenging to distinguish from uniform, will appear in practically relevant datasets.

1 Overview

Recent results linking trainability with classical simulation have led to growing concerns that quantum machine learning could be limited to physics-based use cases or will have to wait for fault-tolerant devices [CLGM⁺25]. One of the leading approaches that goes against this belief is generative machine learning using Instantaneous Quantum Polynomial (IQP) circuits [RAB25, SPDW26]. These circuits start and end with a layer of Hadamards each, and contain only diagonal gates in between. Since all diagonal gates commute, there is no time ordering for the gates and hence the term "instantaneous". Their suitability for quantum advantage in generative machine learning is predominantly driven by research in two broad directions: complexity theoretical results showing that classically sampling from the output distributions of certain families of random IQP circuits is provably hard [BMS17, BJS10], and the ability to train them classically efficiently to solve generative learning problems [RAB25, RAAB25, SPDW26]. The focus of our work is in the former direction, especially its practical consequences.

The first step to physically achieving this advantage for quantum computers is to implement a protocol that demonstrates this classical sampling difficulty on currently available quantum hardware. Since the complexity advantage of IQP sampling relies on a property known as *anti-concentration* (very "flat" output distributions), which is also shared by the Google *supremacy* experiments [AAB⁺19, MVM⁺24], a natural experimental demonstration of hardness would be the Linear Cross Entropy (XEB) benchmarking tests that they pioneered.

An apt IQP candidate is the sparse IQP circuit family introduced in [BMS17], whose output distributions are anti-concentrated and provably hard for classical computers to sample from, and can be transpiled onto quantum systems with square grid topologies in sublinear depth. They are also considered good candidates for generative learning tasks, as it has been proven that they avoid the notorious barren plateau issue [LTW⁺25] that has plagued many quantum machine learning models [SPDW26]. Hence, XEB-based experiments using sparse IQP circuits have gained traction recently [HKB⁺24, PK25, BEG⁺23].

In this work, we show how the XEB-based supremacy experiments using IQP circuits can be efficiently *spoofed*, where one aims to not sample from a distribution close to the target but from one that has high XEB with the target. This is possible since XEB is not a true metric on the set of probability distributions. Though the technique is quite

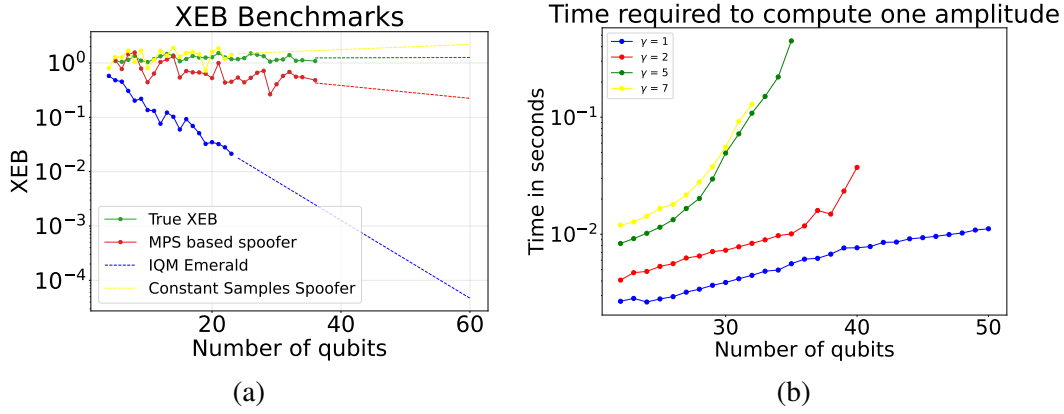


Figure 1: (a) XEB scores computed on a real quantum computer vs spoofers discussed in this work. We can see that extremely high gate fidelities will be required to beat the MPS-based spoofer, especially on devices with restricted topologies. (b) Time required to compute one output probability using the tensor network-based technique introduced in this work. All amplitudes are computed under 1 second, raising the bar significantly for demonstrating quantum advantage through sampling, even for fault tolerant IQP circuits.

different, it follows the spirit of [TGB26, GKC⁺24], who showed similar results for Google’s family of circuits. This research direction is crucial for exploring the boundaries of classical computation, as well as for understanding the right way to design tasks that demonstrate the superiority of quantum computers in generating samples in the pre-fault-tolerant era. The existence of this spoofer implies that one would require near fault-tolerant levels of gate fidelities to produce samples that can beat the spoofer in an XEB test, and demonstrate something genuinely quantum.

Furthermore, we introduce an algorithm inspired by the work of [CvdW23] which is able to calculate sparse IQP circuit amplitudes in under 1 second on GPUs for circuits of size 30-50 qubits. Whilst this enables us to calculate XEB values for much higher circuit sizes, it also allows us to sample exactly from the distribution classically with a number of amplitudes scaling linearly in the number of qubits by leveraging the methods developed in [BGL22], further raising the bar for a genuine show of quantum advantage. Also, we discuss a new method to spoof the XEB-based experiments by computing a constant number of amplitudes, thus increasing the difficulty even higher for XEB-based experiments.

One approach proposed to avoid these pitfalls of XEB experiments has been to focus on *peaked* circuits, where the aim is to construct quantum circuits that have output distributions with only a few bitstrings having large probabilities. Given this capability, one can frame a supremacy problem where the task is to find these bitstrings - a quantum computer can easily find them by sampling from the circuit, and hopefully it is difficult for any classical algorithm [AZ24, Uda25, GMS⁺25]. The hardness of finding these peaks implies the hardness of sampling from such distributions. This is relevant to quantum generative models, since we believe it is unlikely that flat anti-concentrated distributions, which are challenging to distinguish from uniform, will appear in practically relevant datasets.

Our third and final contribution is a novel method of generating peaked IQP circuits, which could be used for designing efficiently verifiable supremacy circuits. This is also achieved by starting with sparse random difficult IQP circuits, followed by IQP circuits that are trained to peak them. The commutation property of gates, along with the presence of the difficult random IQP circuit, is leveraged to make the circuit difficult for classical algorithms. We demonstrate that one can generate descriptions of peaked IQP circuits defined on hundreds of qubits easily on a laptop, and empirical results reveal that Matrix Product State (MPS) based classical simulation algorithms struggle to find these bitstrings in a reasonable time.

2 Results

2.1 Spoofing XEB-Based Supremacy Experiments of IQP Circuits

The XEB achieved by our efficient spoofer is presented using the red curve in Figure 1 (a). One can see that the XEB scores achieved by the spoofer are many orders of magnitude higher than the XEB on a real quantum computer (IQM Emerald). Further analysis shows that such an experiment will require near fault-tolerant levels of fidelities on fixed topologies to beat this spoofer. The spoofer is built using the local expectations based MPS approximation algorithm introduced in [CPF⁺10], as such expectations can be computed efficiently classically for IQP circuits [vdN10]. We also design another inefficient spoofer, which, although it requires computation of amplitudes, only requires very

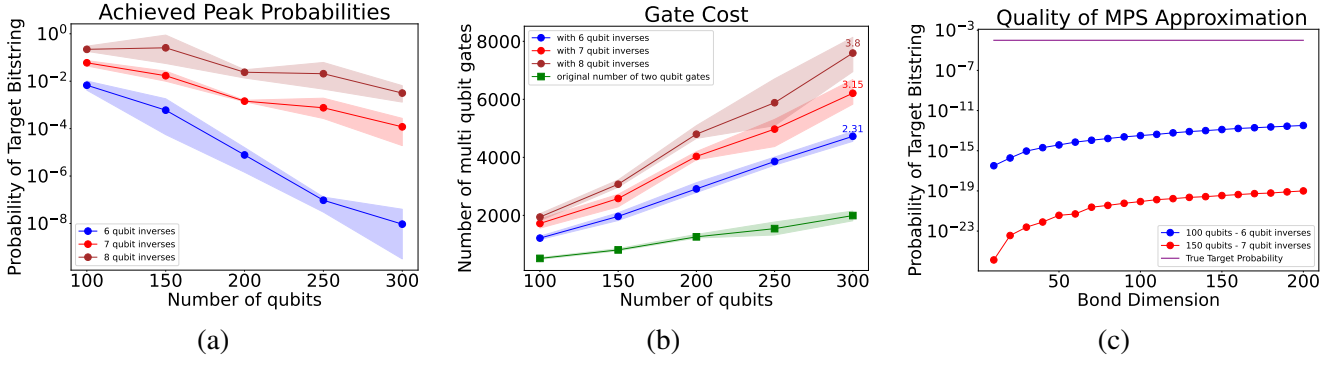


Figure 2: (a) Peak probability values of IQP circuits generated using our method. (b) The number of multi-local gates required. (c) Probability of the target bitstring of the MPS approximation. Here, as an example, we have created a 100 qubit IQP circuit with around 1200 multiqubit gates that has peak probabilities around 10^{-4} , which will take an MPS approximation algorithm with bond dimension 200 around 10^{15} samples to find these bitstrings.

few of them. Its performance is plotted with the yellow curve in Figure 1 (a), where for each circuit, we use only 5 amplitudes for the spoofer. As we can see, the XEBs are very high, and for the chosen sparsity levels, higher than the XEBs that the true distributions would produce. This shows that for sparse IQP circuits, since there are methods like the ones introduced in this work that can compute amplitudes much faster than in the case of generic circuits, and since only very few amplitude computations suffice, even a fault-tolerant sparse IQP implementation should at least have a number of logical qubits that is beyond this capability to demonstrate quantum advantage in sampling using XEB experiments.

2.2 IQP Probability Computation Using Tensor Networks

The computation times required for single amplitudes for different sparse IQP circuits are given in Figure 1 (b), where we show that amplitudes of large circuits can be computed under 1 second. Here, γ is a parameter that controls the density of the circuit; the higher the γ , the more two-qubit gates are involved. The algorithm is inspired by the graph-based IQP amplitude computation algorithm introduced in [CvdW23]. In that work, ZX calculus is used to reduce an amplitude computation into a series of "lighter" amplitude computations over maximally independent sets of qubits of the circuit. In our version, we show that similar computational reductions into graphs that are "almost" trees can be carried out without ZX calculus by using techniques developed in [RWL24], making the final evaluation easy for tensor networks. The results shown in Figure 1 b) have not been optimised - with appropriate hyperparameter choices, we expect to improve this plot to scale linearly for each fixed γ .

2.3 Peaked IQP Circuits

In Figure 2, we provide the results of different circuits that we have generated, which are trained to peak over 2 bitstrings, with their gate costs, as well as the quality of solutions produced by MPS approximation algorithms. Here, as an example, we have created 100 and 150 qubit IQP circuits with around 1200 and 2000 multiqubit gates, respectively, that have peak probabilities around 10^{-3} , which will take an MPS approximation algorithm with bond dimension 200 around 10^{15} and 10^{20} samples to find these bitstrings. These circuits are created by starting with a sparse random IQP circuit, and then inverting different subcircuits within this circuit as much as possible with newer subcircuits, defined on more qubits, that do not contain the connections featured in the original subcircuits. Furthermore, by using commutation of the gates, we mix and combine all the gates of the final circuit in random, different ways, which makes it difficult for classical algorithms to identify the inversions and peaks. An interesting feature here is that peaked circuits like these can be prepared for other commuting classes of circuits, which consist of diagonal gates sandwiched by layers of some arbitrary single-qubit gates (and their inverses) instead of Hadamards. This allows us to generate circuits that are robust towards marginal-based classical simulation techniques proposed for certain classes of peaked IQP circuits, such as [PBG20]. The results presented above involve multi-local gates, making them difficult to run on fixed topologies. Our current research is focused on developing new methods to generate similar circuits that are easier to apply on square grid topologies, implementing them on IQM hardware, and also testing our circuits against other classical algorithms such as tensor networks equipped with belief propagation [AA21]. We hope to additionally present these results at QTML '26.

References

- [AA21] R. Alkabetz and I. Arad. Tensor networks contraction and the belief propagation algorithm. *Physical Review Research*, 3(2), April 2021.
- [AAB⁺19] Frank Arute, Kunal Arya, Ryan Babbush, Dave Bacon, Joseph C. Bardin, Rami Barends, Rupak Biswas, Sergio Boixo, Fernando G. S. L. Brandao, David A. Buell, Brian Burkett, Yu Chen, Zijun Chen, Ben Chiaro, Roberto Collins, William Courtney, Andrew Dunsworth, Edward Farhi, Brooks Foxen, Austin Fowler, Craig Gidney, Marissa Giustina, Rob Graff, Keith Guerin, Steve Habegger, Matthew P. Harrigan, Michael J. Hartmann, Alan Ho, Markus Hoffmann, Trent Huang, Travis S. Humble, Sergei V. Isakov, Evan Jeffrey, Zhang Jiang, Dvir Kafri, Kostyantyn Kechedzhi, Julian Kelly, Paul V. Klimov, Sergey Knysh, Alexander Korotkov, Fedor Kostritsa, David Landhuis, Mike Lindmark, Erik Lucero, Dmitry Lyakh, Salvatore Mandr, Jarrod R. McClean, Matthew McEwen, Anthony Megrant, Xiao Mi, Kristel Michielsen, Masoud Mohseni, Josh Mutus, Ofer Naaman, Matthew Neeley, Charles Neill, Murphy Yuezhen Niu, Eric Ostby, Andre Petukhov, John C. Platt, Chris Quintana, Eleanor G. Rieffel, Pedram Roushan, Nicholas C. Rubin, Daniel Sank, Kevin J. Satzinger, Vadim Smelyanskiy, Kevin J. Sung, Matthew D. Trevithick, Amit Vainsencher, Benjamin Villalonga, Theodore White, Z. Jamie Yao, Ping Yeh, Adam Zalcman, Hartmut Neven, and John M. Martinis. Quantum supremacy using a programmable superconducting processor. *Nature*, 574:505–510, oct 2019.
- [AZ24] Scott Aaronson and Yuxuan Zhang. On verifiable quantum advantage with peaked circuit sampling. *arXiv:2404.14493*, 2024.
- [BEG⁺23] Dolev Bluvstein, Simon J. Evered, Alexandra A. Geim, Sophie H. Li, Hengyun Zhou, Tom Manovitz, Sepehr Ebadi, Madelyn Cain, Marcin Kalinowski, Dominik Hangleiter, J. Pablo Bonilla Ataides, Nishad Maskara, Iris Cong, Xun Gao, Pedro Sales Rodriguez, Thomas Karolyshyn, Giulia Semeghini, Michael J. Gullans, Markus Greiner, Vladan Vuletić, and Mikhail D. Lukin. Logical quantum processor based on reconfigurable atom arrays. *Nature*, 626(7997):58–65, December 2023.
- [BGL22] Sergey Bravyi, David Gosset, and Yincheng Liu. How to simulate quantum measurement without computing marginals. *Physical Review Letters*, 128(22), June 2022.
- [BJS10] Michael J. Bremner, Richard Jozsa, and Dan J. Shepherd. Classical simulation of commuting quantum computations implies collapse of the polynomial hierarchy. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 467(2126):459–472, August 2010.
- [BMS17] Michael J Bremner, Ashley Montanaro, and Dan J Shepherd. Achieving quantum supremacy with sparse and noisy commuting quantum computations. *Quantum*, 1:8, 2017.
- [CLGM⁺25] Marco Cerezo, Martin Larocca, Diego García-Martín, N. L. Diaz, Paolo Braccia, Enrico Fontana, Manuel S. Rudolph, Pablo Bermejo, Aroosa Ijaz, Supanut Thanasilp, Eric R. Anschuetz, and Zoë Holmes. Does provable absence of barren plateaus imply classical simulability? or, why we need to rethink variational quantum computing. *Nature Communications*, 16, 7907, 2025.
- [CPF⁺10] Marcus Cramer, Martin B. Plenio, Steven T. Flammia, Rolando Somma, David Gross, Stephen D. Bartlett, Olivier Landon-Cardinal, David Poulin, and Yi-Kai Liu. Efficient quantum state tomography. *Nature Communications*, 1(1), December 2010.
- [CvdW23] Julien Coudi and John van de Wetering. Classically simulating quantum supremacy IQP circuits through a random graph approach. *arXiv:2212.08609*, 2023.
- [GKC⁺24] Xun Gao, Marcin Kalinowski, Chi-Ning Chou, Mikhail D. Lukin, Boaz Barak, and Soonwon Choi. Limitations of linear cross-entropy as a measure for quantum advantage. *PRX Quantum*, 5(1), February 2024.
- [GMS⁺25] Hrant Gharibyan, Mohammed Zuhair Mullath, Nicholas E. Sherman, Vincent P. Su, Hayk Tepanyan, and Yuxuan Zhang. Heuristic quantum advantage with peaked circuits. *arXiv:2510.25838*, 2025.

- [HKB⁺24] Dominik Hangleiter, Marcin Kalinowski, Dolev Bluvstein, Madelyn Cain, Nishad Maskara, Xun Gao, Aleksander Kubica, Mikhail D Lukin, and Michael J Gullans. Fault-tolerant compiling of classically hard IQP circuits on hypercubes. *arXiv:2404.19005*, 2024.
- [LTW⁺25] Martin Larocca, Supanut Thanasilp, Samson Wang, Kunal Sharma, Jacob Biamonte, Patrick J. Coles, Lukasz Cincio, Jarrod R. McClean, Zoë Holmes, and Marco Cerezo. A review of barren plateaus in variational quantum computing. *Nat Rev Phys* 7, 174–189, 2025.
- [MVM⁺24] A. Morvan, B. Villalonga, X. Mi, S. Mandrà, A. Bengtsson, P. V. Klimov, Z. Chen, S. Hong, C. Erickson, I. K. Drozdov, J. Chau, G. Laun, R. Movassagh, A. Asfaw, L. T. A. N. Brandão, R. Peralta, D. Abanin, R. Acharya, R. Allen, T. I. Andersen, K. Anderson, M. Ansmann, F. Arute, K. Arya, J. Atalaya, J. C. Bardin, A. Bilmes, G. Bortoli, A. Bourassa, J. Bovaird, L. Brill, M. Broughton, B. B. Buckley, D. A. Buell, T. Burger, B. Burkett, N. Bushnell, J. Campero, H.-S. Chang, B. Chiaro, D. Chik, C. Chou, J. Cogan, R. Collins, P. Conner, W. Courtney, A. L. Crook, B. Curtin, D. M. Debroy, A. Del Toro Barba, S. Demura, A. Di Paolo, A. Dunsworth, L. Faoro, E. Farhi, R. Fatemi, V. S. Ferreira, L. Flores Burgos, E. Forati, A. G. Fowler, B. Foxen, G. Garcia, É. Genois, W. Giang, C. Gidney, D. Gilboa, M. Giustina, R. Gosula, A. Grajales Dau, J. A. Gross, S. Habegger, M. C. Hamilton, M. Hansen, M. P. Harrigan, S. D. Harrington, P. Heu, M. R. Hoffmann, T. Huang, A. Huff, W. J. Huggins, L. B. Ioffe, S. V. Isakov, J. Iveland, E. Jeffrey, Z. Jiang, C. Jones, P. Juhas, D. Kafri, T. Khatyar, M. Khezri, M. Kieferová, S. Kim, A. Kitaev, A. R. Klots, A. N. Korotkov, F. Kostritsa, J. M. Kreikebaum, D. Landhuis, P. Laptev, K.-M. Lau, L. Laws, J. Lee, K. W. Lee, Y. D. Lensky, B. J. Lester, A. T. Lill, W. Liu, W. P. Livingston, A. Locharla, F. D. Malone, O. Martin, S. Martin, J. R. McClean, M. McEwen, K. C. Miao, A. Mieszala, S. Montazeri, W. Mruczkiewicz, O. Naaman, M. Neeley, C. Neill, A. Nersisyan, M. Newman, J. H. Ng, A. Nguyen, M. Nguyen, M. Yuezhen Niu, T. E. O’Brien, S. Omonije, A. Opremcak, A. Petukhov, R. Potter, L. P. Pryadko, C. Quintana, D. M. Rhodes, C. Rocque, E. Rosenberg, N. C. Rubin, N. Saei, D. Sank, K. Sankaragomathi, K. J. Satzinger, H. F. Schurkus, C. Schuster, M. J. Shearn, A. Shorter, N. Shutty, V. Shvarts, V. Sivak, J. Skrzynny, W. C. Smith, R. D. Somma, G. Sterling, D. Strain, M. Szalay, D. Thor, A. Torres, G. Vidal, C. Vollgraff Heidegger, T. White, B. W. K. Woo, C. Xing, Z. J. Yao, P. Yeh, J. Yoo, G. Young, A. Zalcman, Y. Zhang, N. Zhu, N. Zobrist, E. G. Rieffel, R. Biswas, R. Babbush, D. Bacon, J. Hilton, E. Lucero, H. Neven, A. Megrant, J. Kelly, P. Roushan, I. Aleiner, V. Smelyanskiy, K. Kechedzhi, Y. Chen, and S. Boixo. Phase transitions in random circuit sampling. *Nature*, 634(8033):328–333, October 2024.
- [PBG20] Hakop Pashayan, Stephen D. Bartlett, and David Gross. From estimation of quantum probabilities to simulation of quantum circuits. *Quantum*, 4:223, January 2020.
- [PK25] Chae-Yeun Park and Michael J. Kastoryano. Complexity phase transitions in instantaneous quantum polynomial-time circuits. *Phys. Rev. Res.*, 7:013001, Jan 2025.
- [RAAB25] Erik Recio-Armengol, Shah Nawaz Ahmed, and Joseph Bowles. Train on classical, deploy on quantum: scaling generative quantum machine learning to a thousand qubits. *arXiv:2503.02934*, 2025.
- [RAB25] Erik Recio-Armengol and Joseph Bowles. IQPopt: Fast optimization of instantaneous quantum polynomial circuits in JAX. *arXiv:2501.04776*, 2025.
- [RWL24] Joel Rajakumar, James D Watson, and Yi-Kai Liu. Polynomial-time classical simulation of noisy IQP circuits with constant depth. *arXiv:2403.14607*, 2024.
- [SPDW26] Kevin Shen, Susanne Pielawa, Vedran Dunjko, and Hao Wang. Characterizing trainability of instantaneous quantum polynomial circuit born machines. *arXiv:2602.11042*, 2026.
- [TGB26] Andrew Tanggara, Mile Gu, and Kishor Bharti. Classically spoofing system linear cross entropy score benchmarking. *arXiv:2405.00789*, 2026.
- [Uda25] O. G. Udalov. A method for constructing quasi-random peaked quantum circuits. *arXiv:2508.07491*, 2025.
- [vdN10] M. van den Nest. Quantum information and computation. *arXiv:0911.1624*, 2010.