

Generative AI Beyond Tokens: Quantum Resource Consumption of IQP Circuits

Tom Krüger

Technical University of Applied Sciences Regensburg
FI CODE, Universität der Bundeswehr München
Regensburg/Munich, Germany
tom.krueger@othr.de

Wolfgang Maurer

Technical University of Applied Science Regensburg
Siemens AG, Foundational Technology
Regensburg/Munich, Germany
wolfgang.maurer@othr.de

Abstract—Quantum generative modelling casts sampling as a generative task: a parametrised quantum circuit is trained such that sampling reproduces a target probability distribution. Instantaneous Quantum Polynomial-time (IQP) circuits combine structural simplicity with complexity-theoretic evidence for quantum advantage. Yet their practical value depends not only on expressivity, but on how efficiently they consume genuinely quantum resources.

We study this question through the lens of magic, or non-stabiliserness, as a resource for quantum generative modelling. We show that established fidelity- and geodesic-based notions of computational progress in a projective Hilbert space are ill-suited to generative models, since operational performance is determined by output probability distributions rather than quantum states themselves. We evaluate magic-consumption directly on the probability simplex, using changes in Jensen-Shannon divergence to quantify progress. Applying this framework to trained random γ -sparse IQP circuits shows signatures of efficient magic use, with the dominant contribution arising from two-qubit gates. As IQP circuits produce remarkably low intermediate magic relative to phase-randomised states with the same sampling distributions, this renders IQP-based quantum generative models as promising candidates for resource-efficient demonstrations of quantum advantage on early fault-tolerant architectures.

Index Terms—Quantum Resource Theory, Quantum Generative Models, IQP Circuits, Quantum Magic

I. INTRODUCTION

Quantum generative modelling (QGM) turns sampling quantum circuits into a generative task: a parameterised circuit is trained such that measurements of its output state reproduce a target probability distribution. This perspective is particularly natural in view of complexity-theoretic supremacy arguments that show certain quantum sampling protocols cannot be implemented efficiently classically [1]. Instantaneous Quantum Polynomial-time (IQP) circuits are especially attractive: They provide restricted Born-machine ansätze for learning and generating probability distributions, while retaining the complexity-theoretic hardness of sampling from their output distributions under standard assumptions [2]–[4].

Apart from the question if they can represent useful distributions, it is important to understand which genuinely quantum resources are required, particularly for fault-tolerant quantum computing (FTQC), where logical Clifford operations are comparatively benign, while non-Clifford resources remain costly and must typically be supplied through magic-state

preparation, injection, or distillation. Consumption of non-stabiliserness provides a means of determining non-classical contributions to quantum generative modelling [5]–[7]. Evaluating magic consumption allows for (a) physical cost estimation and (b) benchmarking the efficiency of converting quantumness into computational utility.

Existing approaches to magic-consumption for variational and structured circuits use geometric notions of progress in a projective Hilbert space [7]. For quantum generative models, however, distinct quantum states may induce the same measurement distribution, and are equivalent for generative performance; fidelity-based or geodesic measures on a Hilbert space can misrepresent progress in the learning task. We argue that magic-consumption evaluation must instead be formulated on the manifold of probability distributions, where the circuit output is compared to the target distribution. We realise this by replacing projective-state progress with a Jensen-Shannon divergence based measure on the probability simplex.

We analyse the efficiency of magic consumption in random γ -sparse IQP circuits trained as quantum generative models, and find that they exhibit signatures of efficient magic use; relevant contributions predominantly arise from two-qubit gates. As trained IQP circuits also produce intermediate states with remarkably low magic compared with phase-randomised states representing the same output distributions, we suggest that IQP-based quantum generative models offer a compelling architecture for benchmarking (and providing) resource-efficient quantum advantage on early FTQC devices.

II. PRELIMINARIES

We fix notation and define basics in the following remarks.

A. Quantum Generative Modeling

Given a probability distribution $Q(x)$, quantum generative modelling [8], [9] seeks to train a parametrised quantum circuit $C(\phi)$ such that the output probability distribution $P(x) = |\langle x|C(\phi)|0\rangle|^2$ matches $Q(x)$. If $C(\phi)$ can be trained efficiently and $Q(x)$ is classically complex to sample from, QGMs can achieve quantum advantage.

B. Sparse IQP

Definition 1. A n qubit IQP circuit is given by (a) an input state $|0\rangle$; (b) an implemented a unitary $U_{\text{IQP}} = H^{\otimes n} D H^{\otimes n}$,

where $H^{\otimes n}$ is the n qubit Hadamard gate and D is a diagonal operator in the computational basis; and (c) a final state measurement in the computational basis.

D is commonly constructed from Pauli-Z rotations $RZ_I(\theta)$ and controlled Pauli-Z rotations $C_j RZ_I(\theta)$. Here $RZ_I(\theta) = \exp(i\theta Z_I)$, where Z_I is the Pauli string with Z operators acting on all qubits $i \in I$; we also use $Z_{I \cap [a:b]}$ for Z_I restricted to qubits $a \leq i \leq b$. Further, $C_j RZ_I(\theta)$ is a $RZ_I(\theta)$ gate controlled by $j \notin I$.

We rewrite U_{IQP} by applying the Hadamard transformation to D : Uncontrolled $RZ_I(\theta)$ gates are transformed to $RX_I(\theta) = \exp(i\theta X_I)$ gates, where X_I is a tensor product of Pauli-X operators acting on the qubits in $I \subset \mathbb{N}$. For controlled $C_j RZ_I(\theta)$ gates, we assume qubits $j \cup I = \{k_1, k_2, \dots, k_m\}$ are all adjacent and ordered such that $k_i < k_{i+1}$ and $k_1 = j$ (if not, we permute wlog). Given such an order, $C_j RZ_I(\theta) = \text{diag}(\mathbb{1}, RZ_I(\theta))$. Consider a controlled Z rotation $\text{diag}(\mathbb{1}, RZ(\theta) = \exp(i\theta \text{diag}(\mathbf{0}, Z)))$ with $\text{diag}(\mathbf{0}, Z) = (\mathbb{1} Z - Z Z)/2$. The Hadamard transform of the exponential of a diagonal operator D is obtained by transforming D , thus $H \exp(i\theta \text{diag}(\cdot)) H = \exp(i\theta H \text{diag}(\cdot) H)$. In our case $H(\mathbb{1} Z - Z Z)/2 H = (\mathbb{1} X - X X)/2$ and $H C_j RZ(\theta) H = \exp(i\theta (\mathbb{1} X - X X)/2)$.

Generalising to multi-qubits as $C_j RZ_I(\theta) = \exp(i\theta (Z_{I \cap [1:j-1]} \otimes \mathbb{1} \otimes Z_{I \cap [j+1:n]} - Z_{I \cup \{j\}})/2)$. As gates in $\{RZ_I(\theta), C_j RZ_I(\theta)\}$ commute pairwise, a Hadamard transformed circuit constructed from this set can be expressed with gates out of $\{RX_I(\theta), \exp(i\theta (X_{I \cap [1:j-1]} \otimes \mathbb{1} \otimes X_{I \cap [j+1:n]} - X_{I \cup \{j\}})/2)\}$.

Random γ -sparse IQP Circuits: To sample a γ -sparse IQP circuit, for each pair of qubits one adds a 2-qubit gate with probability $\gamma \log(n)/n$, for $0 < \gamma \leq n/\log(n)$ [10]. Additionally, a 1-qubit gate is added per qubit. On average, a random γ -sparse IQP circuit has $O(n \log(n))$ 2-qubit gates. In our case for a qubit pair $\{i, j\}$ we sampled the 2-qubit gates uniformly at random from $\{RZ_{\{i,j\}}(\theta), C_k RZ_l(\theta)\}$, where $(k, l) = (i, j)$ or (j, i) with equal probability. Lastly, all gates were randomly shuffled to evenly distribute one-qubit and two-qubit gates.

C. Magic

We call the set of states reachable by Clifford operations the Clifford orbit or stabiliser states; they can be simulated efficiently classically. States beyond the Clifford orbit pose inherent non-stabiliser properties (non-stabiliserness) commonly referred to as *magic*. Given the generally accepted hypothesis that non-stabiliser cannot be simulated efficiently classically, magic serves as a resource theoretic measure [5] of the amount of actual quantum resources needed to prepare a state. Although, there are many proposed measures of magic, the α -Stabiliser-Rényi-Entropy (SRE_α) gained widespread adoption in recent years, for a detailed definition we refer to [6]. Unless specified, we use $\alpha = 2$.

D. Geometry

The expected SRE of a randomly Haar sampled n -qubit state $\mathbb{E}_{|\psi\rangle \sim \text{Haar}}(\text{SRE}_\alpha(|\psi\rangle)) \in O(n)$ for all $\alpha \geq 2$, which means that magic is a common feature of the average quantum state. Consequently, magic consumption of a quantum circuit is to be expected and does not necessarily contribute to the computational progress in a meaningful way. Therefore, the efficiency of magic consumption needs to be evaluated. In [7] it was shown how based on the idea of geometric computational efficiency [11] the efficiency of magic consumption can be evaluated. Let $s_0(|\psi\rangle, |\phi\rangle)$ be the geodesic distance between $|\psi\rangle$ and $|\phi\rangle$ on the projective Hilbert space. Using s_0 as proposed in [7] does not work in the case of QGM due to how quantum states are interpreted differently. To see this compare the two states $|+\rangle$ and $|-\rangle$, we have $s_0(|+\rangle, |-\rangle) = \pi$. In QGMs states are interpreted as probability distributions where $P_{|\psi\rangle}(x) = |\langle x|\psi\rangle|^2$, clearly we have $P_{|+\rangle} = P_{|-\rangle}$. In this work we therefore propose to move the efficiency tracing of magic consumption to the space of probability distributions *i.e.* the probability simplex.

In information geometry, closeness between distributions is commonly evaluated with divergencies. A divergence does not fulfill all criteria of a metric, most importantly symmetry. Nevertheless, locally they coincide with metrics on the statistical manifold [12]. A divergence widely adopted is the Kullback-Leibler divergence $D_{\text{KL}}(P : Q)$. Unfortunately D_{KL} is not suited for our analysis as it is not well defined for x s.t. $Q(x) = 0$ and $P(x) \neq 0$. Given two states $|\psi\rangle, |\phi\rangle$, mapping them to the probability simplex, we can not guarantee this requirement. We propose using the Jensen-Shannon divergence

$$D_{\text{JS}}(P : Q) = \frac{1}{2}(D_{\text{KL}}(P : M) + D_{\text{KL}}(Q : M)) \quad (1)$$

with $M = (P + Q)/2$, as D_{JS} is well defined for all pairs of probability distributions mapped from quantum states, and is symmetric and bounded.

E. Efficiency of Magic Consumption

First we note that Magic is a state property, while resource consumption is a trait of state dynamics *e.g.* circuits, gate operations, state propagation. A quantum gate U only manipulates the magic of a state if U is a non-Clifford operation. Vice versa, if $\text{SRE}(U|\psi\rangle) \neq \text{SRE}(|\psi\rangle)$, we know that U performed a non-classical operation. Based on this observation, the *magic consumption* of U can be defined by $|\Delta \text{SRE}| = |\text{SRE}(U|\psi\rangle) - \text{SRE}(|\psi\rangle)|$ [7]. The efficiency of said consumption can be evaluated by comparing it with the geometric approach towards the desired target. As we proposed above, in the case of QGMs a fitting measure is the delta in D_{JS} . In summary, the efficiency of magic consumption can be expressed by $|\Delta \text{SRE}|$ vs $-\Delta D_{\text{JS}}$.

III. EXPERIMENTS / SIMULATIONS

To consider the efficiency of magic consumption in random γ -sparse IQP circuits, we generate random γ -sparse circuits

and train them on random binomial mixture distributions (simulations are performed with QuantumToolbox [13]):

$$P(x; p_1, \dots, p_4) = \frac{1}{4} \sum_{i=1}^4 P_{\text{Bin}}(x; p_i), \quad (2)$$

where $P_{\text{Bin}}(x; p_i)$ is the probability of the of $\sum_{k=1}^n x_k 2^k \in \mathbb{N}$ for $x \in \mathbb{F}_2^n$ in a binomial distribution of $2^n - 1$ draws with probability $p = p_i$. We performed two batches of simulations: 1) to analyse the efficiency of magic consumption throughout the circuit and 2) to analyse the magic of intermediate states in the context of the degree of freedom allowed by relative qubit phases.

Batch 1: For the first batch we randomly generated 500 7-qubit circuit-distribution pairs (C, P) for each $\gamma \in \{1, 1.4, 1.8, 2.2, 2.6, 3, 3.4\}$, here C is a random γ -sparse IQP circuit as described above and P is generated according to Eq. (2) by sampling $0 < p_1, \dots, p_4 < 1$ uniformly at random. For each pair C got optimised to produce a sampling distribution close to P . Concretely, the circuit was trained to minimise the D_{JS} to P .

Batch 2: A notable degree of freedom arises when mapping quantum states to probability distributions: Given a state $|\psi\rangle$ we obtain a distribution $P_{|\psi\rangle}(x) = |\langle x|\psi\rangle|^2$. Now consider a second state

$$|\psi(\theta)\rangle = \sum_x e^{i\theta_x} \sqrt{P_{|\psi\rangle}(x)} |x\rangle, \quad (3)$$

then $P_{|\psi(\theta)\rangle} = P_{|\psi\rangle}$. To investigate this degree of freedom, we generated and trained 50 random circuit-distribution pairs (C, P) similar to batch 1 with a fixed $\gamma = 1.5$. Let C_k be the k -th gate of C , then $|C_k\rangle = C_k \cdots C_1 |0\rangle$. For each $|C_k\rangle$ of the trained circuit we generated 100 phase-randomised states $|C_k(\theta)\rangle$ according to Eq. (3) by choosing the $0 \leq \theta_k \leq 2\pi$ uniform at random.

IV. RESULTS

We again discuss results in two parts, as both batches capture different aspects of magic consumption efficiency.

A. Batch 1

In batch 1, we analyse magic consumption efficiency of each gate in C . Let again $|C_k\rangle = C_k \cdots C_1 |0\rangle$ be the state after the k -th gate in C . We then calculate $|\Delta_{k,k+1} \text{SRE}| = |\text{SRE}(|C_{k+1}\rangle) - \text{SRE}(|C_k\rangle)|$ and $\Delta_{k,k+1} D_{\text{JS}} = D_{\text{JS}}(P_{|C_{k+1}\rangle}, P) - D_{\text{JS}}(P_{|C_k\rangle}, P)$.

Figure 1 shows that magic consumption $|\Delta \text{SRE}|$ correlates with computational progress $-\Delta D_{\text{JS}}$. Further we note two jumps in magic efficiency at circuit densities $1 < \gamma < 1.4$ and $3 < \gamma < 3.4$. Figure 2 shows a stark difference: The clear correlation between magic consumption and computational progress for 2-qubit gates is not present for 1-qubit gates. A detailed split of correlation values over gate types and circuit densities is summarised in Table I. Additionally, the comparison to correlation values between magic consumption and geodesic distance changes Δs_0 on the projective Hilbert space further motivates our proposed evaluation metric ΔD_{JS} .

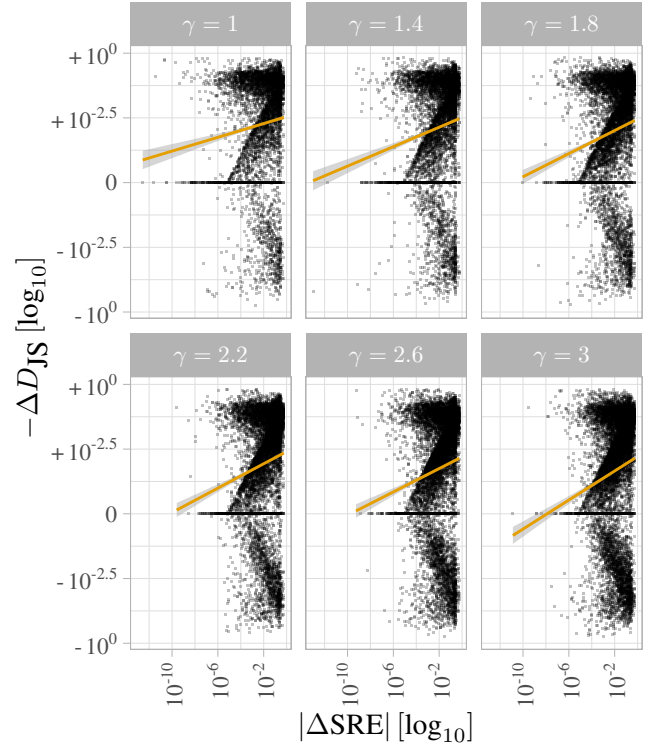


Fig. 1. Magic consumption $|\Delta \text{SRE}|$ is correlated with computational progress measured by $-\Delta D_{\text{JS}}$. It seems like a minimal density $1 < \gamma < 1.4$ is needed but for higher γ values the correlation does not get stronger.

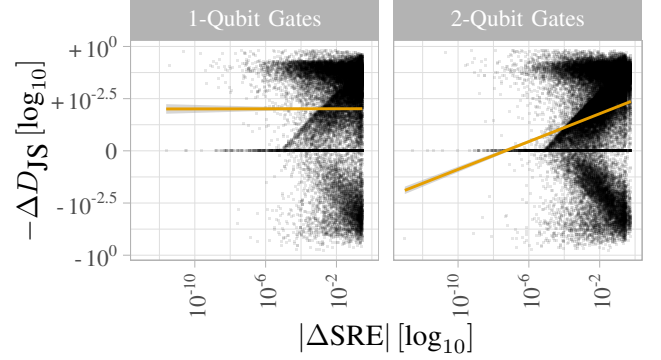


Fig. 2. The majority of 1-qubit gates contributes to computational progress measured in $-\Delta D_{\text{JS}}$, but without clear correlation to magic consumption $|\Delta \text{SRE}|$. For 2-qubit gates, a clear correlation pattern emerges.

TABLE I
CORRELATION: MAGIC CONSUMPTION AND EVALUATION METRICS.

γ	1	1.4	1.8	2.2	2.6	3	3.4
All Gates							
$\rho_{ \Delta \text{SRE} , -\Delta D_{\text{JS}}}$	0.03	0.07	0.07	0.07	0.07	0.06	0.09
$\rho_{ \Delta \text{SRE} , -\Delta s_0}$	0.01	0.02	0.02	0.01	0.01	0.00	0.02
1-Qubit Gates							
$\rho_{ \Delta \text{SRE} , -\Delta D_{\text{JS}}}$	-0.06	-0.01	-0.03	-0.04	-0.03	-0.03	0.01
$\rho_{ \Delta \text{SRE} , -\Delta s_0}$	0.02	0.03	0.02	0.01	0.01	-0.02	0.04
2-Qubit Gates							
$\rho_{ \Delta \text{SRE} , -\Delta D_{\text{JS}}}$	0.11	0.12	0.13	0.13	0.11	0.10	0.12
$\rho_{ \Delta \text{SRE} , -\Delta s_0}$	0.00	0.02	0.02	0.00	0.01	0.00	0.01

B. Batch 2

Following up on the observation made above regarding the degree of freedom presented by alternative choices of relative phases, we will now investigate a second aspect of efficiency, when it comes to magic consumption. Figure 3 shows the SRE distribution of phase-randomised states $|C_k(\theta)\rangle$ for four selected circuits C out of batch 2. States $|C_k\rangle$ produced by the trained circuit exhibit significantly less magic than the average state $\mathbb{E}_\theta(|C_k(\theta)\rangle)$. To further investigate this, we look at the z -score of $\text{SRE}(|C_k\rangle)$ within the distribution of $\{\text{SRE}(|C_k(\theta)\rangle)\}_\theta$. The z -score is defined as $z(x) = (x - \mu)/\sigma$, with μ being the sample mean and σ the standard deviation. As shown in Fig. 4, all but two states $|C_k\rangle$ have a z -score < 0 : Their magic is below average. The significant majority of states, especially at the later stages of the circuit, have a z -score < -3 and can be considered statistical outliers.

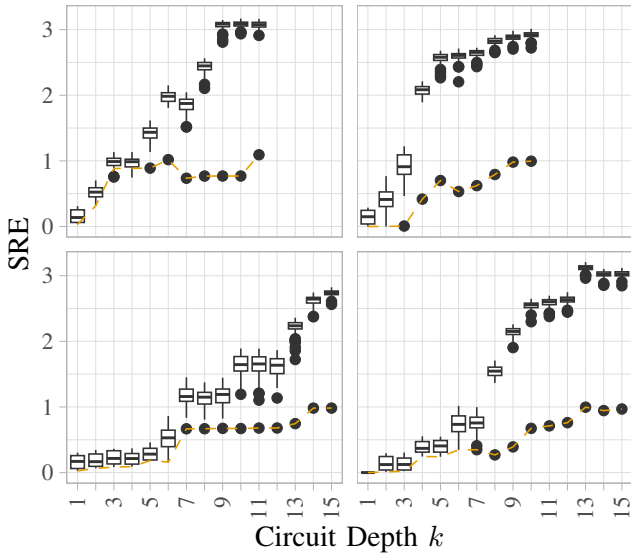


Fig. 3. SRE distribution of intermediate circuit states with random relative phases $\{|C_k(\theta)\rangle\}_\theta$. Dashed orange: SRE of original in circuit states $|C_k\rangle$.

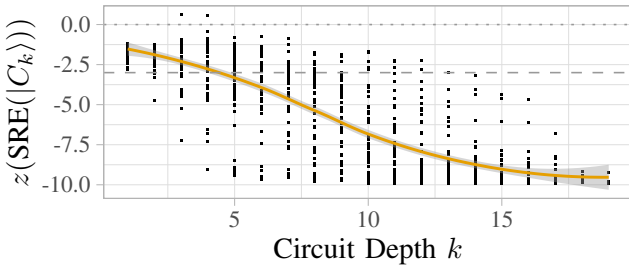


Fig. 4. z -scores for the SRE of each $|C_k\rangle$ in its batch of random $\{|C_k(\theta)\rangle\}_\theta$. All states $|C_k\rangle$ have a below average magic (below dotted line) and a majority of states can be categorised as statistical outliers (below dashed $z = -3$ line).

V. CONCLUSION

We showed why geometric distance measures like s_0 on the projective Hilbert space are inapt to evaluate computational progress quantum generative models. Moreover, we propose

to treat the quantum states as representations of probability distributions and to use D_{JS} to replace s_0 . Using this approach, we were able to evaluate the efficiency of magic consumption in random γ -sparse IQP circuits. Our analysis showed a correlation between magic consumption and computational progress in these circuits, with 2-qubit gates seemingly being the driving factor of magic efficiency. Worth further investigation is the influence of the circuit density γ . The jumps in magic efficiency in the low and high γ regimes could be further investigated, to clarify whether they are mere experimental artifacts of low γ resolution or if they alternatively hint at phase transitions in the magic efficiency of random γ -sparse IQP circuits. In our second batch of simulations we showed that random γ -sparse IQP circuits produce exceptionally low magic states compared to other states representing the same magic distribution. This further suggests IQP based quantum generative models to be a promising architecture for early fault tolerant systems.

Reproducibility A reproduction package [14] can be found at github.com/lfd/qce26_IQP_magic_consumption additionally a DOI safe version is available at DOI:10.5281/zenodo.21630858.

Acknowledgment This work was supported by the German Federal Ministry of Research, Technology and Space (BMFTR), funding program *quantum technologies—from basic research to market*, grant number 13N17387.

REFERENCES

- [1] A. Bouland et al., Quantum supremacy and the complexity of random circuit sampling, *arXiv preprint arXiv:1803.04402*, 2018.
- [2] M. J. Bremner, B. Cheng, and Z. Ji, Instantaneous quantum polynomial-time sampling and verifiable quantum advantage: Stabilizer scheme and classical security, *PRX Quantum*, 2025.
- [3] M. J. Bremner, A. Montanaro, and D. J. Shepherd, Achieving quantum supremacy with sparse and noisy commuting quantum computations, *Quantum*, 2017.
- [4] H. Pashayan, S. D. Bartlett, and D. Gross, From estimation of quantum probabilities to simulation of quantum circuits, *Quantum*, 2020.
- [5] V. Veitch et al., The resource theory of stabilizer quantum computation, *New Journal of Physics*, 2014.
- [6] L. Leone, S. F. E. Oliviero, and A. Hamma, Stabilizer rényi entropy, *Physical Review Letters*, 2022.
- [7] T. Krüger and W. Maurer, Geometric and resource-theoretic characterisation of non-stabiliserness in quantum algorithms, *Physical Review A*, 2026.
- [8] M. S. Rudolph et al., Trainability barriers and opportunities in quantum generative modeling, *npj Quantum Information*, 2024.
- [9] A. Barthe et al., Parameterized quantum circuits as universal generative models for continuous multivariate distributions, *npj Quantum Information*, 2025.
- [10] L. Paletta et al., Robust sparse iqp sampling in constant depth, *Quantum*, 2024.
- [11] J. Anandan and Y. Aharonov, Geometry of quantum evolution, *Physical Review Letters*, 1990.
- [12] S.-i. Amari, *Information Geometry and Its Applications*. Springer Japan, 2016.
- [13] A. Mercurio et al., QuantumToolbox.jl: An efficient Julia framework for simulating open quantum systems, *Quantum*, 2025.
- [14] W. Maurer and S. Scherzinger, 1-2-3 reproducibility for quantum software experiments, *SANER*, 2022.