

## Overdispersion in Gate Tomography

### Experiments and Continuous, Two-Scale Random Walk Model on the Bloch Sphere

Nowak, Wolfgang; Bruennette, Tim; Schalkers, Merel Annelise; Möller, Matthias

**DOI**

[10.1145/3688857](https://doi.org/10.1145/3688857)

**Publication date**

2024

**Document Version**

Final published version

**Published in**

ACM Transactions on Quantum Computing

**Citation (APA)**

Nowak, W., Bruennette, T., Schalkers, M. A., & Möller, M. (2024). Overdispersion in Gate Tomography: Experiments and Continuous, Two-Scale Random Walk Model on the Bloch Sphere. *ACM Transactions on Quantum Computing*, 5(4), Article 24. <https://doi.org/10.1145/3688857>

**Important note**

To cite this publication, please use the final published version (if applicable).  
Please check the document version above.

**Copyright**

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

**Takedown policy**

Please contact us and provide details if you believe this document breaches copyrights.  
We will remove access to the work immediately and investigate your claim.

***Green Open Access added to TU Delft Institutional Repository***

***'You share, we take care!' - Taverne project***

***<https://www.openaccess.nl/en/you-share-we-take-care>***

Otherwise as indicated in the copyright section: the publisher is the copyright holder of this work and the author uses the Dutch legislation to make this work public.



# Overdispersion in Gate Tomography: Experiments and Continuous, Two-Scale Random Walk Model on the Bloch Sphere

**WOLFGANG NOWAK**, Institute for Modelling Hydraulic and Environmental Systems, University of Stuttgart, Stuttgart, Germany

**TIM BRUENNETTE**, Institute for Modelling Hydraulic and Environmental Systems, University of Stuttgart, Stuttgart, Germany

**MEREL ANNEISE SCHALKERS**, Applied Mathematics, Delft University of Technology, Delft, Netherlands

**MATTHIAS MÖLLER**, Applied Mathematics, Delft University of Technology, Delft Netherlands

Noisy intermediate-scale quantum computers (NISQ) are computing hardware in their childhood, but they are showing high promise and growing quickly. They are based on so-called qubits, which are the quantum equivalents of bits. Any given qubit state results in a given probability of observing a value of zero or one in the readout process. One of the main concerns for NISQ machines is the inherent noisiness of qubits, i.e., the observable frequencies of zeros and ones do not correspond to the theoretically expected probability, as the qubit states are subject to random disturbances over time and with each additional algorithmic operation applied to them. Models to describe the influence of this noise exist.

In this study, we conduct extensive experiments on quantum noise. Based on our data, we show that existing noise models lack important aspects. Specifically, they fail to properly capture the aggregation of noise effects over time (or over an algorithm's runtime), and they are underdispersed. With underdispersion, we refer to the fact that observable frequencies scatter much more between repeated experiments than what the standard assumptions of the binomial distribution would allow for. Based on these shortcomings, we develop an extended noise model for the probability distribution of observable frequencies as a function of the number of gate operations. The model roots back to a known continuous random walk on the (Bloch) sphere, where the angular diffusion coefficient can be used to characterize the standard noisiness of gate operations. Here, we superimpose a second random walk at the scale of multiple readouts to account for overdispersion. Further, our model has known, explicit components for noise during state preparation and measurement.

We thank the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) dfg for supporting this work by funding EXC2075–390740016 dfg under Germany's Excellence Strategy. We acknowledge the support by the Stuttgart Center for Simulation Science (SimTech). We furthermore thank the Delft Institute of Applied Mathematics (DIAM) diam for the financial support of Merel Schalkers and acknowledge the use of IBM Quantum services for this work. Moreover, we thank the Collaborative Research Centre SFB 1313, Project Number 327154368 sfb for the support of Tim Brünnette.

Authors' Contact Information: Wolfgang Nowak, Institute for Modelling Hydraulic and Environmental Systems, University of Stuttgart, Stuttgart, Baden-Württemberg, Germany; e-mail: wolfgang.nowak@iws.uni-stuttgart.de; Tim Bruennette, Institute for Modelling Hydraulic and Environmental Systems, University of Stuttgart, Stuttgart, Baden-Württemberg, Germany; e-mail: tim.bruennette@iws.uni-stuttgart.de; Merel Annelise Schalkers, Applied Mathematics, Delft University of Technology, Delft, Netherlands; e-mail: m.a.schalkers@tudelft.nl; Matthias Möller, Applied Mathematics, Delft University of Technology, Delft, Netherlands; e-mail: M.Moller@tudelft.nl.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](https://permissions.acm.org).

© 2024 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ACM 2643-6817/2024/10-ART24

<https://doi.org/10.1145/3688857>

The interaction of these two random walks predicts theoretical, runtime-dependent bounds for probabilities. Overall, it is a three-parameter distributional model that fits the data much better than the corresponding one-scale model (without overdispersion), and we demonstrate the better fit and the plausibility of the predicted bounds via Bayesian data-model analysis.

CCS Concepts: • **Hardware** → **Quantum error correction and fault tolerance**; • **Mathematics of computing** → *Markov processes*;

Additional Key Words and Phrases: Qubit noise model, angular diffusion, Bayesian parameter inference

#### ACM Reference Format:

Wolfgang Nowak, Tim Bruennette, Merel Annelise Schalkers, and Matthias Möller. 2024. Overdispersion in Gate Tomography: Experiments and Continuous, Two-scale Random-walk Model on the Bloch Sphere. *ACM Trans. Quantum Comput.* 5, 4, Article 24 (October 2024), 17 pages. <https://doi.org/10.1145/3688857>

## 1 Introduction

Quantum computing is an emerging computing technology that has the potential to revolutionize the way we solve computational problems in the future. The potential power of quantum computing stems from the fact that information is stored in so-called quantum bits (qubits), which, unlike their binary counterparts, can hold a superposition of the two orthonormal basis states  $|0\rangle$  and  $|1\rangle$ . However, qubit states as such are not observable. Instead, when a qubit is measured, it collapses to either the zero or one state with a *probability* that relates to the superimposed state (see Section 2).

The above-described concept generalizes to quantum registers. An  $n$ -qubit register can hold a superposition of all  $2^n$  bitstrings  $|00 \dots 00\rangle, |00 \dots 01\rangle, \dots, |11 \dots 11\rangle$ , which makes it possible to perform operations on an exponential amount of input data simultaneously. Unlike classical computing, this quantum parallelism does neither require an exponential overhead in time nor the replication of hardware in space to enable parallel execution. It is a result of the qubits' capability to hold multiple states simultaneously.

Quantum computing cannot surpass the no-free-lunch theorem, which manifests itself in the retrieval of information from the output state. As states cannot be observed directly, quantum algorithms need to be executed multiple times to obtain an approximation of the probabilities from which the most probable 'outcome(s)' of the quantum algorithm can be deduced.

Today's **noisy intermediate-scale quantum (NISQ)** computers are not yet ready for solving meaningful real-world problems due to the immaturity of the quantum hardware. State-of-the-art quantum computers from IBM provide up to 433 qubits, which enables the superposition of  $2.2 \times 10^{130}$  states and should bring the solution of real-world problems into reach [12]. However, qubits are very susceptible to noise and feature extremely short decoherence times. This noise limits the complexity of quantum algorithms to perform a few basic single- and two-qubit operations (to be explained below) before the output state degenerates to meaningless noise.

The noise that occurs in practical quantum computation is manifold and can be more or less pronounced for different qubit technologies such as superconducting qubits, trapped ion qubits, photonic qubits, neutral-atom qubits, and semiconductor spin qubits [14, 21, 22, 25, 26]. Without going too much into the technical details one can differentiate at least between two types of errors:

- *systematic errors* such as the mistuned microwave pulse leading to a systematic under-rotation, over-rotation, or rotation around an imprecise axis
- *spurious cross-talk* between neighboring qubits when applying an operation to one or more of them or measuring one or more qubit

However, the standard noise models implemented in today's quantum computer simulators such as depolarizing channel models [17] are typically very simplistic and not capable of modeling the complex noise behavior.

Thus, useful methods for quantum characterization, verification, and validation are highly relevant.

A commonly used and scalable method to partially characterize the noise associated with quantum gates is **randomized benchmarking (RB)** [11]. RB was originally proposed in 2005 [8] with different versions and flavors coming out ever since [7]. In RB, the average gate fidelities of a given NISQ machine (on average over many types and repetitions of gates) are found by performing a randomized series of mutually canceling operations of varying lengths. Upon measurement, one theoretically expects to always find the initial (zero) state but does not in practice. The exponential decay rate explaining these imperfect measurements is the core error metric. One known downside of this technique is that it makes the assumption that noise is time independent [23].

Tomographic methods aim to characterize the quantum operations including noise in full detail. While some of the methods focus on specific parts of the quantum computations (such as state tomography, process tomography, or measurement tomography), the widely used gate set tomography aims to reconstruct the entire pipeline, including **state preparation and measurement errors (SPAM)** [15]. While the comprehensiveness of this approach has its appeal, the level of detail requires large, sometimes prohibitively large amounts of experimental data.

Our proposed method, data (see Section 2.2), and model (see Section 2.4) are akin to RB in that they use lumped error characteristics, with the twist that these are happening at multiple levels. We model the qubits' movement on the Bloch sphere with errors occurring on an individual as well as a pool level while also considering finite sample effects. The two levels roughly represent random errors that affect individual qubits only and random effects, possibly caused by environmental factors, affecting batches of qubits identically. This combination helps to explain overdispersion, an effect noticed in References [16] and [18], in an easily accessible manner while avoiding a generic catch-all error. Further, the model could easily allow an extension to incorporate non-Markovian error types [2]. So far, we focus on the characterization of one particular basis gate, however, an application to random gate sequences, as commonly used in RB [9], is possible.

To explore our idea, we pose the following research questions:

- Can the experimental data confirm that we need two levels?
- Our hypothesized model will predict growth and then decay of overdispersion over algorithm runtime, and even provides theoretical bounds for the overdispersion at any given algorithm runtime. Can we observe evidence of this?

## 2 Methods

### 2.1 Bloch Sphere Representation for a Single Qubit

A qubit is the quantum-computing version of a bit. Its state is generally denoted as  $|\Psi\rangle$ , defined by the following:

$$|\Psi\rangle = \gamma_0 |0\rangle + \gamma_1 |1\rangle, \quad \gamma_0, \gamma_1 \in \mathbb{C}, \quad |0\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad |1\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad (1)$$

where the complex numbers  $\gamma_0$  and  $\gamma_1$  are sometimes called probability amplitudes. Furthermore,  $|0\rangle$  and  $|1\rangle$  are orthonormal basis states forming the computational basis. Equation (1) represents a superposition of the two basis states.

Qubit states as such are not observable. When a qubit is measured, it collapses to either  $|0\rangle$  or  $|1\rangle$  (here simply called 0 or 1) with probabilities  $P$  given by

$$P[0] = |\gamma_0|^2, \quad P[1] = |\gamma_1|^2. \quad (2)$$

As the events to measure 0 or 1 are mutually exclusive and collectively exhaustive, the corresponding probabilities sum up to unity,

$$|\gamma_0|^2 + |\gamma_1|^2 = P[0] + P[1] = 1. \quad (3)$$

From the total of four degrees of freedom contained in the two complex numbers  $\gamma_0$  and  $\gamma_1$ , the constraint in Equation (3) removes one. Another degree of freedom has no physically observable consequences, and so a qubit has two relevant degrees of freedom [17]. Due to Equation (3), these form a unit sphere, called the Bloch sphere [17], which can be thought of as a sphere in  $\mathbb{R}^3$ .

The Bloch sphere is like a globe, where the north pole represents the zero-state  $|0\rangle$  and the south pole represents the one-state  $|1\rangle$ . The so-called colatitude is given by the angle  $\theta \in [0, \pi]$  with zero at the north pole (in contrast to regular latitude). For  $\theta = 0$  at the north pole, a measurement returns the zero-state with probability  $P[0] = 1$  and similarly at the south pole with  $\theta = \pi$ . The equator at  $\theta = \pi/2$  corresponds to  $P[1] = P[0] = 1/2$ . The longitude is given by the angle  $\varphi \in [0, 2\pi]$ , which does not influence the probabilities at all, but is still relevant in quantum computing as it changes the readout probabilities after appropriate rotations.

Since the Bloch sphere is a powerful visualization and a good basis for constructing noise models, it is worthwhile to note the transformations between the probability amplitudes  $\gamma_0$  and  $\gamma_1$  and the angles  $\theta$  and  $\varphi$ ,

$$\gamma_0 = \cos\left(\frac{\theta}{2}\right), \quad \gamma_1 = e^{i\varphi} \sin\left(\frac{\theta}{2}\right). \quad (4)$$

As the longitude  $\varphi$  does not influence the probabilities  $P[0]$  and  $P[1]$ , we look specifically at the relation between  $\theta$  and  $P = P[0]$ ,

$$\begin{aligned} P = P[0] = |\gamma_0|^2 &= \cos^2\left(\frac{\theta}{2}\right) = \frac{1}{2} + \frac{1}{2} \cos(\theta) \\ \Rightarrow \theta &= \arccos(2P - 1). \end{aligned} \quad (5)$$

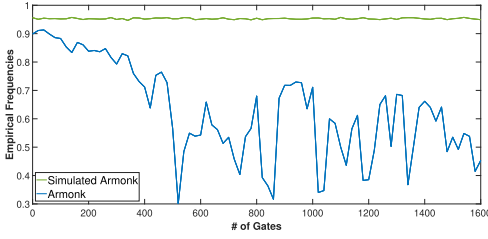
For transforming probability density functions between  $\theta$  and  $P$ , the Jacobian of this relation will be useful,

$$\frac{\partial \theta}{\partial P} = -\frac{1}{\sqrt{P - P^2}}. \quad (6)$$

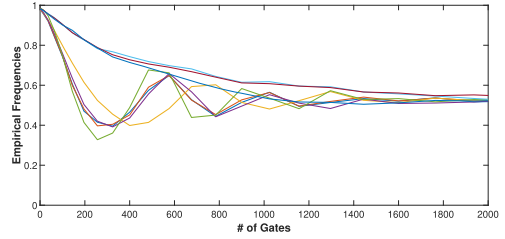
The relation is monotonic, which is an important prerequisite in transformations of random variables (e.g., Reference [4]).

## 2.2 Data

**2.2.1 Data Creation.** To build and test our extended noise model, we first acquire the necessary data by running experiments on open-access quantum computers. The experiments are set up such that, in noise-free theory, we know we should always measure the zero-state with probability 1. To obtain observed frequencies (as the probabilities cannot be directly observed), we repeat each algorithm 8,192 times, which is the upper limit on the freely available machine we use. We then repeat this for multiple circuits with varying lengths (i.e., varying numbers of gate operations), always ensuring that the expected theoretical outcome remains the zero state.



(a) Results from the simulated and real Armonk device. For the real device, each data point was produced at a different wall clock time.



(b) Multiple runs of repeated  $\sqrt{x}$  applications. Each curve corresponds to a single queue, thus effectively a single wallclock time.

Fig. 1. Plots showing the different noise behaviors in three cases. The simulated device, the real machine with mixed wallclock times and the real machine with (multiple) single wallclock times.

We collect the resulting data and plot the observed frequency of measuring the zero-state as a function of the number of gates applied to the qubit. In this article, we consider each such gate applied as a timestep, therefore the axis representing the number of gates applied can be interpreted as a time axis.

In the data for our model, we make use of the  $\sqrt{x}$ -gate, one of IBM's native gates. In each considered circuit, the  $\sqrt{x}$ -gate is applied a number of times that is a multiple of four, ensuring that the total effect on the qubits should vanish.<sup>1</sup> The circuits are compiled using optimization level-0, which guarantees that the gates are actually performed on the physical qubit(s) (and not optimized away as they cancel each other out).

In Figure 1(a), we plot the frequencies of measuring the correct (here: zero) state in the case that the experiments for the different data points are run at different wallclock times. In Figure 1(b), we plot the frequencies of measuring the correct state in the case that the experiments for the different data points are run at the same wallclock time. Here, “different time” means through different jobs submitted to the system's waiting queue that have been processed in intervals of, on average, 7 minutes between the individual jobs; “same time” means submitted as a bulk job without any waiting time in between. Regardless of this difference, one can see how the degree of dispersion (i.e., the scatter of the lines around the hypothetical mean line) increases at first, until somewhere before 500 gate operations in the specific case, and then decreases again.

**2.2.2 Data Availability.** The full settings used for 1-qubit IBM Armonk machine [12] can be found in Appendix B. The data produced are available (*link to GitLab repository will be released after publication*) [online](#). The data file contains the resulting frequency of measuring the zero-state, the amount of  $\sqrt{x}$ -gates applied, and the timestamp of when the experiment was run on the IBM Armonk machine.

### 2.3 Relevant Levels of Randomness and Probability Distributions

In the following derivations, four distinct levels of randomness will appear. To avoid confusion, we declare them upfront:

- Bernoulli level:** Consider the state  $|\Psi\rangle$  (here: specifically  $\theta$ ) of a qubit to be fixed. For an individual measurement ( $n = 1$ ), we know the theoretical probability  $P[0]$  via Equation (5). But when conducted we see either a zero or one, with no meaningful way to define a (frequentist) probability in a single experiment due to  $n = 1$ .

<sup>1</sup>Note that  $\sqrt{x} = \frac{1}{2} \begin{bmatrix} 1+i & i-1 \\ 1-i & 1+i \end{bmatrix}$ , therefore  $\sqrt{x}\sqrt{x}\sqrt{x}\sqrt{x} = I$ .



- **Binomial level:** Again consider a qubit in fixed state with theoretically known  $P$ . After  $n > 1$  repetitions, one can observe a meaningful frequency of the binary outcomes  $\{0, 1\}$ . For  $n \rightarrow \infty$ , the observed frequency converges to the Bernoulli-level probability. For distinguished notation, we use  $\bar{P} = P$  for probabilities at the binomial level and  $\bar{F}$  as observed frequencies. The discrepancy between  $\bar{P}$  and  $\bar{F}$  is described by the well-known binomial distribution [10].
- **Pool level:** Consider a total of  $m$  repetitions of the binomial level (with  $n$  repetitions each). These  $m$  “outer” repetitions form a so-called pool of observed frequencies  $\bar{F}_q; q = 1 \dots m$ . Each of the  $m$  repetitions can have a different qubit state with corresponding theoretical (binomial-level) probability  $\bar{P}_q, q = 1 \dots m$ . When analyzing the pooled data altogether, one observes a pooled frequency  $\bar{F}$ , which converges for  $n, m \rightarrow \infty$  toward the theoretical mean  $\bar{P}$  of the binomial-level probabilities  $\bar{P}_q$ . However, the observed  $\bar{F}_q$  will not only scatter around  $\bar{P}$  due to the classical binomial law but also due to differences in the respective values  $\bar{P}_q$  and due to finite  $m$ ; therefore, the binomial distribution does not apply to such pooled data.
- **Parametric level:** The levels so far are all classical frequentist uncertainties. Probability distribution models fitted to the observed data are typically governed by parameters. Parametric uncertainty refers to the fact that these distributional parameters can only be inferred with finite precision at finite  $m, n$ . The corresponding parametric uncertainty is often quantified via Bayesian inference [3].

Relevant, well-known probability distributions used or referred to in the upcoming derivations are as follows:

- The binomial distribution describes the observable frequency under  $n$  independent repetitions of a binary-outcome (e.g., 0, 1) random experiment with constant theoretical probability  $P$ :  $\bar{F} \sim \text{Bino}(n, P)$ , with parameters  $n, P$ .
- The Uniform distribution  $X \sim \mathcal{U}(0, 1)$  assigns identical probability density to all values of a continuous random variable  $X$  in the interval  $[0, 1]$  [10].
- Gilbert’s sine distribution  $X \sim \text{Sine}() = \sin(2x), \forall x \in [0, \pi/2]$  assigns a sinus-shaped probability density [6]. It can be scaled to any interval  $[L, U]$  with width  $W = U - L$  by  $\text{Sine}(L, U) = W/2\pi \cdot \sin(2\pi/W \cdot 2(x + L))$ .

## 2.4 Two-level Random Walk on the Bloch Sphere

As a possible noise model, we see the noise of qubit states as a random walk on the Bloch sphere: Imagine a qubit initialized at the north pole. At later times, the position on the sphere will not necessarily be identical to the starting position; instead, the position on the sphere is randomly moving over time, e.g., due to noisy gate operations.

The new model to account for overdispersion proposed here is a two-level random walk on the Bloch sphere. The first level represents classical noise ideas as just described, while the second level, the pool level, is responsible for the overdispersion. Graphically, one can imagine the pool-level random walk as an additional, joint component in the random walks of many random walkers, specifically a joint motion of their center of mass on the surface of the sphere.

The average position over many random positions of pure states *on* the Bloch sphere due to the first level of the random walk actually lies *within* the sphere. That means it has a shortened radius smaller than one. We will exploit this fact later to summarize the relevant statistics of the first random walk and combine it with the second random walk.



**2.4.1 Known Results for  $\theta(t)$  in Single-level Random Walks.** In quantum computing, a probability distribution over states is called a mixed or *decoherent* state, while a deterministic state according to Equation (1) is called a *coherent* state [17]. At the binomial level (see Section 2.3), we define a theoretical probability  $\bar{P}$ . This  $\bar{P}$  is the expected value of the noise-affected Bernoulli-level probabilities  $P$  (see Section 2.3) over the probability distribution of qubit states  $p(\theta, \varphi)$ :  $\bar{P} = E_{\theta, \varphi}[|\gamma_0|^2]$ . Such a distribution on the sphere is in fact equivalent to a mixed state, and could as well be represented by a shortened Bloch vector, or by density matrices [17], i.e., as a decoherent state. However, as only  $\theta$  is the probability-relevant angle, we can afford to simplify our derivations.

As a starting point for our noise model, assume a random walk on the Bloch sphere, here written as time-discrete,

$$|\Psi\rangle_{t+1} = |\Psi\rangle_t + \Delta_{|\Psi\rangle, t}. \quad (7)$$

Further, assume that the increments  $\Delta_{|\Psi\rangle, t}$  are independent and identically distributed for any given time lag  $\Delta t = t_i - t_j$ . Due to the independence property, the random series  $|\Psi\rangle_t$  is a Markov process (e.g., Reference [13]). If, additionally, we assume that these increments are Gaussian distributed with zero mean and fixed variance, and we define these increments as angular increments, then we arrive at Fickian diffusion on the Bloch sphere. Written in terms of angles  $\theta, \varphi$ , it is governed by

$$\frac{\partial p(\theta, \varphi; t)}{\partial t} = D_n \nabla p(\theta, \varphi; t), \quad (8)$$

where  $p(\theta, \varphi; t)$  is the joint **probability density function (PDF)** of  $\theta$  and  $\varphi$  as a function of time  $t$ ,  $D_n$  is the rotational diffusion coefficient in  $\text{rad}^2/\text{s}$ .

We restrict ourselves to initial conditions of  $\theta = 0$ . Also, as we do not consider systematic actions on  $\varphi$  that would make  $\varphi$  relevant to probabilities  $P$ , we can afford to neglect  $\varphi$  as of now. A small discussion on the limitation due to this omission can be found in Section 4. So let  $p(\theta; t)$  be the PDF for the colatitude  $\theta$  as a function of time  $t$ . For the case of constant  $D_n$  and the initial condition  $p(\theta; t = 0) = \delta(\theta)$ , expansion in spherical harmonics provides the following analytical solution, as derived in Reference [20]:

$$p(\theta; t) = \sum_{k=0}^{\infty} \frac{2k+1}{2} e^{-D_n k(k+1)t} \cdot L_k(\cos \theta) \sin \theta \quad \forall t > 0, \quad (9)$$

where  $L_k$  is the Legendre polynomial of order  $k$  [1]. In practice, the infinite sum has to be truncated and is known to converge only slowly for small  $t$ . However, with fast and accurate algorithms for evaluating Legendre polynomials, truncation at  $K = 1000$  or higher is unproblematic.

**2.4.2 Mapping onto  $P(t)$  (Bernoulli Level).** While the above analytical solution is a known result, we are interested in transforming it into the PDF for the Bernoulli-level measurement probability  $P = P[0]$ . For clarity, we will now write  $p_\theta(\theta)$  and  $p_P(P)$ , where one must realize that  $p_P(P)$  is a probability density that describes what theoretical probabilities  $P$  one may encounter. Using standard rules for PDFs after change of variables [4], we have the following:

$$p_P(P; t) = p_\theta(\theta(P); t) \cdot \left| \frac{\partial \theta(P)}{\partial P} \right|. \quad (10)$$

Recalling the transformation in Equation (5) and its Jacobian in Equation (6), inserting into Equation (9) yields

$$p_P(P; t) = \sum_{k=0}^{\infty} (2k+1) e^{-D_n k(k+1)t} L_k(2P-1), \quad (11)$$

since

$$\cos \theta = 2P - 1, \quad \sin \theta = 2\sqrt{P - P^2}. \quad (12)$$

**2.4.3 Limit Cases and Information Entropy.** For a better understanding, it is worthwhile to look at certain limit cases of Equation (11):

$$\lim_{t \rightarrow \infty} p_P(P; t) = L_0(2P - 1) = 1, \quad \forall P \in [0, 1], \quad (13)$$

$$p_P(P = 0; t) = \sum_{k=0}^{\infty} (-1)^k (2k + 1) e^{-D_n k(k+1)t}, \quad (14)$$

$$p_P(P = 1; t) = \sum_{k=0}^{\infty} (2k + 1) e^{-D_n k(k+1)t}, \quad (15)$$

which is due to  $L_k(-1) = (-1)^k$  and  $L_k(1) = 1$ , respectively. Looking at the first limit, we see that qubit states approach a *uniform* distribution over their measurement probabilities at late times, the distribution with maximal information entropy, i.e., the least informative one. At the same time, Equation (14) approaches its  $t \rightarrow \infty$  limit of 1 strictly from above, and Equation (15) approaches the same  $t \rightarrow \infty$  limit but strictly from below and starting from zero.

Upon reflection, Equation (13) is intuitive. Due to symmetry, the theoretical average of the uniform distribution is 0.5. Therefore, the observable frequencies over many measurements (each one with a different, random probability from the uniform distribution) converge to the intuitive uninformative 50:50 coin toss between  $|0\rangle$  and  $|1\rangle$ . Such types of averages over 11 will be treated in more detail in the upcoming section. But, as a noteworthy fact, the random walk on the Bloch sphere installs the maximum-entropy property (here: *uniform* distribution) in terms of measurable probabilities, and not in terms of  $\theta$ ,

$$\lim_{t \rightarrow \infty} p_{\theta}(\theta; t) = \frac{1}{2} \sin(\theta), \quad (16)$$

which ends up as the *Sine()* distribution on the  $[0, \pi/2]$  interval. A uniform distribution on the *surface* of the Bloch sphere is not the same as a uniform distribution on the *angles*.

**2.4.4 Transfer to Theoretical Probabilities  $\bar{P}(t)$  (Binomial Level).** Equation (11) is a distribution over the individual probabilities of single ( $n = 1$ ) measurements at the Bernoulli level. Accordingly, each individual measurement will result in an observed zero-state or one-state with individual random probability  $P$  and  $1 - P$ , respectively. Therefore, the probability distribution as in Equation (11) will never be observable. Instead, when repeating the measurement process  $n$  times, with  $n \rightarrow \infty$ , one will empirically observe a probability  $\bar{P}(t) = E[P(t)]$  at the binomial level. The fact that the empirically visible probability is the expected value  $E[P(t)]$  follows from the properties of the Poisson-binomial distribution [24].

Hence, we are now interested in the expected value  $\mu_P = E[P(t)]$  and the general  $\ell$ th moments  $\mu_{P,\ell}(t) = E[P^\ell(t)]$  of Equation (11),

$$\begin{aligned} \mu_{P,\ell}(t) &= E[P^\ell(t)] = \int_0^1 p_P(P; t) P^\ell \, dP \\ &= \int_0^1 \sum_{k=0}^{\infty} (2k + 1) e^{-D_n k(k+1)t} P^\ell L_k(2P - 1) \, dP \\ &= \sum_{k=0}^{\infty} (2k + 1) e^{-D_n k(k+1)t} \int_0^1 P^\ell \mathcal{L}_k(P) \, dP, \end{aligned} \quad (17)$$

with the *shifted* Legendre polynomials  $\mathcal{L}_k = L_k(2x - 1)$ , whose orthogonality property  $\int_0^1 \mathcal{L}_\ell(x) \mathcal{L}_k(x) dx = (2k + 1)^{-1} \delta_{\ell k}$  [1] we exploit in the following. Specifically for  $\ell = 0$  we have  $P^0 = \mathcal{L}_0(P)$ , and hence

$$\mu_{P,0}(t) = \sum_{k=0}^{\infty} (2k + 1) e^{-D_n k(k+1)t} \cdot \underbrace{\int_0^1 \mathcal{L}_0(P) \mathcal{L}_k(P) dP}_{(2k+1)^{-1} \delta_{k0}} = 1 \quad \forall t, \quad (18)$$

which is a requirement for valid PDFs.

The first interesting case is for  $\ell = 1$ , giving the probability  $\bar{P}(t)$ . Here, we use  $P^1 = \frac{1}{2} \mathcal{L}_0(P) + \frac{1}{2} \mathcal{L}_1(P)$ ,

$$\begin{aligned} \bar{P}(t) = \mu_{P,1}(t) &= \sum_{k=0}^{\infty} (2k + 1) e^{-D_n k(k+1)t} \cdot \underbrace{\int_0^1 \left( \frac{1}{2} \mathcal{L}_0(P) + \frac{1}{2} \mathcal{L}_1(P) \right) \mathcal{L}_k(P) dP}_{\frac{1}{2}(2k+1)^{-1} \delta_{k0} + \frac{1}{2}(2k+1)^{-1} \delta_{k1}} \\ &= \underbrace{\frac{1}{2}}_{k=0} + \underbrace{\frac{1}{2} e^{-2D_n t}}_{k=1}. \end{aligned} \quad (19)$$

This is a beautiful result; for qubit states subject to a random walk on the Bloch sphere with isotropic, constant angular diffusion coefficient  $D_n$  after initialization to the north pole, the measurements asymptotically approach the indifferent 50:50 probability with the first-order rate  $2D_n$ , allowing to describe the development by

$$\frac{\partial \bar{P}(t)}{\partial t} = -2D_n \left( \bar{P} - \frac{1}{2} \right). \quad (20)$$

One should remain clear that this does not imply that qubit states as such tend to the equator. Instead, the ensemble average over all possible noise-affected states tends toward the uninformative 50:50 probability, when the qubit states are uniformly distributed over the Bloch surface, i.e., an entirely decoherent state.

The second moment will play a role in overdispersed data at the pool level. Hence, we set  $\ell = 2$ , use  $P^2 = \frac{1}{3} \mathcal{L}_0(P) + \frac{1}{2} \mathcal{L}_1(P) + \frac{1}{6} \mathcal{L}_2(P)$  and repeat the procedure:

$$\begin{aligned} \mu_{P,2}(t) &= \sum_{k=0}^{\infty} (2k + 1) e^{-D_n k(k+1)t} \cdot \int_0^1 \left( \frac{1}{3} \mathcal{L}_0 + \frac{1}{2} \mathcal{L}_1 + \frac{1}{6} \mathcal{L}_2 \right) \mathcal{L}_k dP \\ &= \underbrace{\frac{1}{3}}_{k=0} + \underbrace{\frac{1}{2} e^{-2D_n t}}_{k=1} + \underbrace{\frac{1}{6} e^{-6D_n t}}_{k=2}. \end{aligned} \quad (21)$$

By transformation from raw moments  $\mu_\ell$  to central moments  $\mu_{\ell c}$  (e.g., Reference [19]), we obtain the variance  $\sigma_P^2(t)$  for the Bernoulli-type probabilities  $P$ :

$$\begin{aligned} \sigma_P^2 &= \mu_{P,2c} = \mu_{P,2} - \mu_{P,1}^2 = \frac{1}{3} + \frac{1}{2} e^{-2D_n t} + \frac{1}{6} e^{-6D_n t} - \left( \frac{1}{2} + \frac{1}{2} e^{-2D_n t} \right)^2 \\ &= \frac{1}{12} - \frac{1}{4} e^{-4D_n t} + \frac{1}{6} e^{-6D_n t}. \end{aligned} \quad (22)$$

Special cases are as follows:

$$\sigma_P^2(t=0) = 0, \quad (23)$$

$$\lim_{t \rightarrow \infty} \sigma_P^2(t) = \frac{1}{12}, \quad (24)$$

where the former is due to the deterministic initial condition, and the latter is known to be the variance of the uniform distribution [10], the limit distribution as discussed in Section 2.4.3.

**2.4.5 Pool-level Randomness.** For extending the above results to overdispersed data, we now invoke the pool level with  $m$  repetitions of a binomial-level experiment. Each binomial-level experiment with  $n$  measurements at a given time  $t$  has a different binomial-level probability  $\bar{P}_q(t)$ ,  $q = 1 \dots m$ . We now assume that the random walk on the Bloch sphere has two components: (1) an i.i.d. noise component as discussed above, which works individually on all Bernoulli trials, and (2) a common binomial-level component for all  $n$  random walkers of a size- $n$  binomial experiment that introduces randomness into a size- $m$  pool of binomial experiments, causing the probabilities  $\bar{P}_q(t)$  within the pool to vary.

For all linear relations  $y = f(x)$ , the simplification  $E[y] = f(E[x])$  can be used. We can use this simplification due to the linear character of qubit gates that represent rotations (and hence random moves) on the Bloch sphere. Hence, we can directly work with the *expected* position from the binomial-level random walk on the Bloch sphere to construct our pool-level random walk. The expected position, due to symmetry, stays constant  $\theta(t) = 0$  at the north pole for all times but simply has a reduced *length*  $R$  of the Bloch vector given by

$$R(t) = 1 - 2\bar{P}(t) = e^{-2D_n t}, \quad (25)$$

which follows from Equation (19). We use this reduced length to represent the statistics of the first level of random walk at the binomial level. Then, we use Equation (11) anew for representing the second level, just working on the reduced-length vector. Thus, we construct the overdispersed probability distribution of  $\bar{P}_q(t)$  by using Equation (9) together with the reduced length from Equation (25):

$$\bar{P}_q(\theta_q; t) = \frac{1}{2} + R(t) \cdot \frac{1}{2} \cos(\theta_q), \quad (26)$$

where  $\theta_q$  is the random position of the joint center of mass of all  $n$  random walkers in the size- $n$  binomial experiment number  $q$ . While  $\theta_q$  played no role before and always stayed at the north pole, as of now, it is subject to its own random walk that also follows Equation (7).

Next, we follow all derivation steps from Equations (5) and (6) to Equation (11), but with Equation (26) as the starting point to obtain

$$p_{\bar{P}_q}(\bar{P}_q; t) = p_{\theta_q}(\theta_q; t) \left| \frac{\partial \theta_q}{\partial \bar{P}_q} \right| = \frac{1}{R(t)} \sum_{k=0}^{\infty} (2k+1) e^{-D_q k(k+1)t} L_k \left( \frac{1}{R(t)} (2\bar{P}_q - 1) \right),$$

where  $D_q$  is the pool-level angular diffusion coefficient (subscript  $q$  signifying the pool level), and the previous noise-related diffusion coefficient  $D_n$  is contained in  $R(t)$  (see Equation (25)).

Although the combined distribution from Equation (27) cannot be simplified analytically, we know several properties. To visualize them, Figure 2 shows how the distribution of empirical frequencies  $\bar{F}$  changes over “time,” i.e., over the number of applied gates, from  $10^0$  to  $2 \times 10^4$  on a logarithmic scale. The top plot shows a case of our model with dominant noise ( $D_n \gg D_q$ ); the center plot shows the same for  $D_n \approx D_q$ , and the bottom plot shows the case of dominant pool-level errors ( $D_q \gg D_n$ ). The relevant properties are as follows:

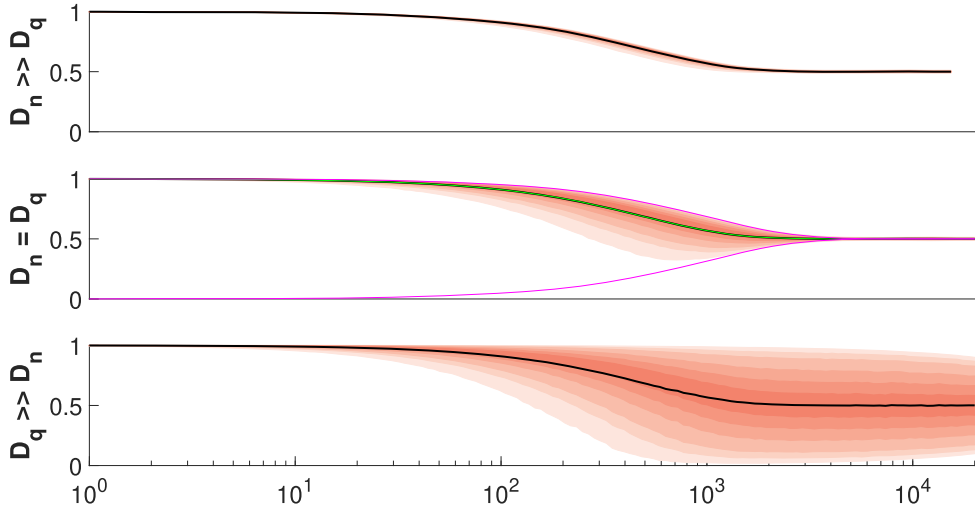


Fig. 2. Simulated distribution of empirical frequencies over the number of applied gates for different relations between  $D_n$  and  $D_q$ .

- **Upper bound:** At any given time  $t$ , the maximal attainable binomial probability  $\bar{P}$  occurs when  $\theta_q = 0$ , and the corresponding probability  $\bar{P}_{q,\max}(t)$  is given by Equation (11), and hence converges to 0.5 with exponential rate of  $D_n$ , see upper magenta line in the centre plot.
- **Lower bound:** At any given time  $t$ , the minimal attainable probability  $\bar{P}$  occurs at  $\theta_q = \pi$ , and the corresponding probability is  $\bar{P}_{q,\min}(t) = 1 - \bar{P}_{q,\max}(t)$ , see lower magenta line in the centre plot.
- **Total expectation:** The increments of the two random walks in angular coordinates are simply additive. Thus, the overall statistical expectation (i.e., the across-pool mean  $\check{P}$ ) experiences a combined random walk with  $D_{\text{tot}} = D_n + D_q$ . Therefore, the across-pool mean  $\check{P}$  has to evolve according to Equation (11), with  $D_{\text{tot}}$  substituted for  $D_n$ , see black line in all three plots. The black line always stays between (but not in the middle) of the two bounds. As we chose  $D_q + D_n = \text{const.}$  the black line is always the same.
- **Overdispersion:** The overdispersion of binomial-level probabilities  $\bar{P}_q$  against the across-pool mean  $\check{P}$  expressed as variance is given by Equation (22) with diffusion coefficient  $D_q$ . The corresponding overdispersion is indicated in the plots via red-shaded percentile intervals. The corresponding overdispersion is hardly visible in the top plot, well visible in the center, and dominant in the bottom plot.

This means, we know time-dependent bounds for  $\bar{P}_q(t)$ , and we know its theoretical mean  $\check{P}(t)$  and its theoretical variance as a function of  $t$ . We also know that the upper and lower bounds both converge to the same limit,

$$\lim_{t \rightarrow \infty} \bar{P}_{q,\max}(t) = \lim_{t \rightarrow \infty} \bar{P}_{q,\min}(t) = \frac{1}{2}, \quad (27)$$

and that the limit distribution within these bounds for  $t \rightarrow \infty$  is the *Sine()* distribution. While the large-time limit (*Sine()*) distribution is apparently symmetric, the early-time distributions are concentrated at the upper bound. Accordingly, in the center plot, we can see how overdispersion initially does not fill the possible range between the limits, then opens up, and is finally suppressed

by actual noise. In the bottom plot, we can see how the distribution of overdispersion asymptotically becomes symmetric, following a *Sine()* distribution, before it would become suppressed by actual noise at timescales beyond the plotted range.

## 2.5 Bayesian Inference of Parameters in the Probabilistic Noise Model

Having stochastically modeled the uncertainties at the Bernoulli, binomial, and pool levels, we set out to fit the model to our experimental data. We lump the diffusion coefficients on the different levels  $D_n, D_q$  plus an initial  $D_{ini}$  representing the SPAM noise into a model parameter vector  $M$ . Also, we arrange our experimental frequencies  $\bar{F}$  after different lengths of quantum algorithms (see Section 2.2) into a data vector  $y$ .

We interpret one read-out frequency, i.e., a batch of 8,192 algorithm executions with a certain number of gates, as a binomial collection that undergoes a joint second-level random walk, meaning we look at all of the data in the same way as it is also presented in Figure 1(a). We interpret every individual gate operation as a timestep in both levels of the random walk and additionally let the binomial-level random walk have a virtual zeroth timestep with  $D_{ini}$  as its diffusion coefficient.

As we are also interested in statements on significance and confidence about our model and its parameters  $M$ , we apply Bayesian parameter inference [3]. That means, we seek the distribution  $p(M|y)$  of model parameters  $M$  conditional on data  $y$ :

$$p(M|y) \propto \int p(y|M) p(M) d\theta_q. \quad (28)$$

Here  $p(M)$  is a prior PDF over the model parameters, set to be improper flat over the positive real numbers;  $p(y|M)$  serves as the so-called likelihood function, expressing for any values in  $M$  the model's quality of fit to the data  $y$ . In our case, it is given by the binomial distribution, but only if we knew the values of  $\theta_q$  within each binomial ( $n = 8192$ ) trial, i.e., per observed frequency  $\bar{F}$ . Therefore, we treat the collection of  $\theta_q$ , one per binomial-level trial, as hidden variables also to be inferred by the data, and use Equation (9) as its PDF for any given value of  $D_q$  substituted for  $D_n$ .

Technically, we use a Metropolis–Hastings algorithm [5] to solve Equation (28), equipped with a so-called Gibbs split, where the proposals for the hidden variables  $\theta_q$  and the primary parameters  $M$  are done separately, and mutual independence among the entries of  $\theta_q$  can be exploited. We generate  $10^6$  parameter samples after burn-in from Equation (28) with our MCMC, out of which we store every 20th, and so obtain 50,000 approximately independent samples of  $M$  to work with.

## 3 Results and Discussion

We apply our two-level model from Section 2.4 to the data from Section 2.2 via Bayesian inference as described in Section 2.5. As mean conditional parameter values, we obtain  $\bar{D}_{ini} = 0.0218$ ,  $\bar{D}_n = 4.9764 \times 10^{-4}$ ,  $\bar{D}_q = 3.2418 \times 10^{-4}$ , i.e.,  $D_n$  and  $D_q$  are on a similar magnitude like in the center plot of Figure 2, and  $D_{ini}$  for state preparation and measurement is about 50 times larger, i.e., it is comparable to about 50 initial  $\sqrt{x}$  gate operations.

Figure 3 shows the conditional distributions of  $D_n$  and  $D_q$  together with their joint distribution. Looking at the small scatter of inferred values relative to the inferred mean values, it becomes clear that both diffusion coefficients, including the one on the pool level, are distinctly different from zero to achieve a good model fit. In  $10^6$  evaluated equiprobable cases, not a single realization shows  $D_q$  anywhere below  $2 \times 10^{-4}$ . This shows a substantial difference from zero for the pool-level  $D_q$  with extreme confidence. Also, looking at the joint distribution, one cannot see any correlation between  $D_n$  and  $D_q$ , i.e., more noise via larger  $D_n$  could not compensate to replace  $D_q$ . The claim that the pool level does indeed improve the model is strengthened further by comparing the maximal found log-likelihood values to the single-level model (obtained by simply forcing  $D_q = 0$ ).

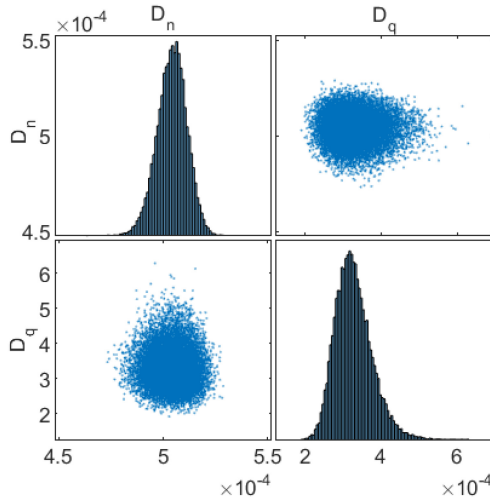
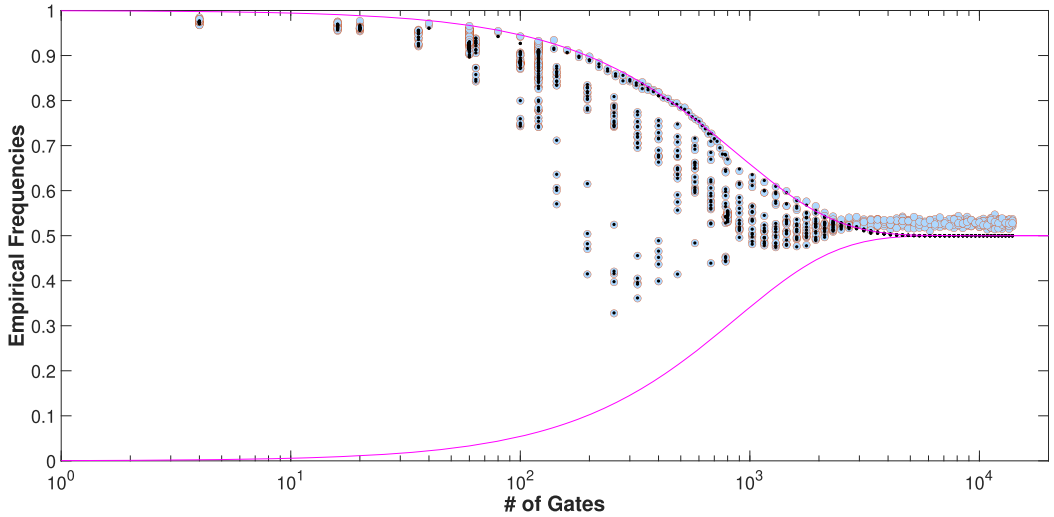
Fig. 3. The (joint) distributions of  $D_n$  and  $D_q$ .

Fig. 4. How well the empirical frequencies are matched.

Here, we have values of  $\mathcal{L}(M_1) = -8974$  and  $\mathcal{L}(M_2) = -1186$ , a log-ratio of about 7000 in favor of our two-level model.

Figure 4 shows the resulting fit of the model to the data. Specifically, the inferred probabilities  $\bar{P}_q$  at the binomial level, based on the hidden variables  $\theta_q$  (black dots), nicely match the observed frequencies  $\bar{F}_q$  (blue circles). Also, the data neatly fall in between the two theoretical bounds predicted by our model and populates especially the space close to the upper bound. Therefore,  $D_n$  (the noise-related diffusion coefficient that dictates the upper bound) could almost be found directly by graphical analysis. As additional visual evidence in favor of our two-level model, one can see that trying to fit the data from Figure 4 into the upper plot of Figure 2 is quickly rejected, i.e., the pool-level random walk with its resulting explanation for overdispersion is indeed required.

There remains a noticeable discrepancy between the data and the model in the late stages. However, this long-time behavior is not of interest when it comes to characterizing the error for



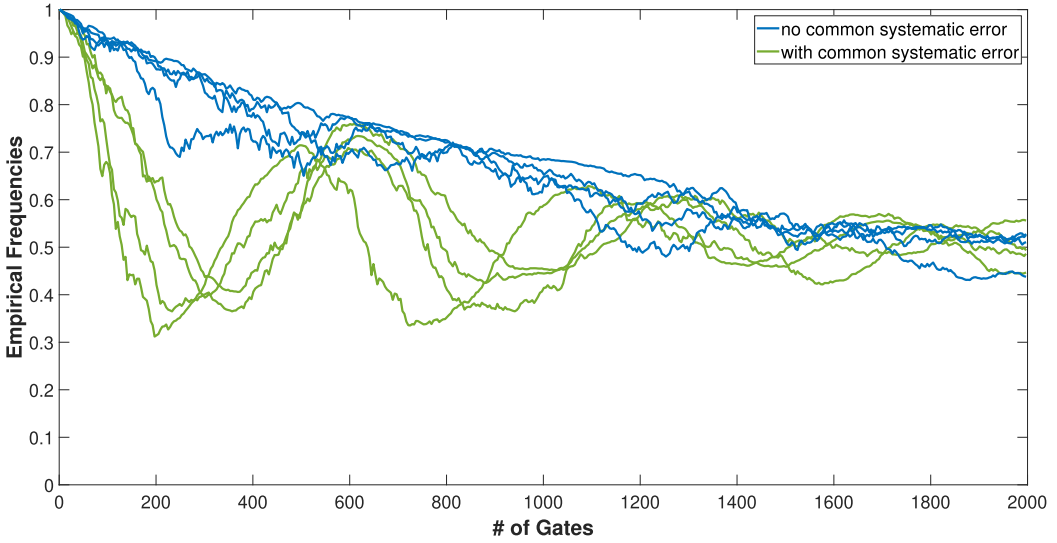


Fig. 5. Forward runs simulated with the MAP estimates. The green curves have an added common deterministic error per run, resulting in oscillations comparable to Figure 1(b).

quantum algorithms, as the distribution is almost completely uninformative by then anyway. In ad hoc modeling variants not shown here, we improved this aspect by introducing a different asymptote, with another parameter to be tuned, into the model, giving even better visual fit and likelihood values.

As a last check, we perform Monte Carlo simulations of the random walk to see how closely we could match the optics of a figure like Figure 1(b). Results are shown in Figure 5. Many of the runs in Figure 1(b) display oscillations, with systematic errors as the most likely explanation. These oscillations occur at some times of the day, and not at others, so some curves show this behavior, while others do not. Our model reproduces the time-sliced data in the distributional sense. The blue curves in Figure 5 represent repeated gate applications simulated with the parameters obtained from the distributional, time-sliced viewpoint. To emulate the behavior of the oscillatory runs, a shared systematic over-rotation was added to a constant fraction of the qubits producing the green curves in Figure 5. The specific quantities stem from manual ad hoc tuning, but it is possible to also include such parameters in the Bayesian inference if one has enough data to use full runs as data points. As we set out to characterize random (diffusive) rather than coherent errors as a function of algorithmic lengths, we omit this aspect in our current analysis.

Nevertheless, a possible extension of our model would be to add to the current pool-level random walk a coherent component, which has increments that are constant or autocorrelated over wallclock time. In the diffusion literature, such non-Markovian behavior is well known. Then, the diffusion coefficient, constant in our model, is defined as half the rate of growth of the variance,

$$D_x(t) = \frac{1}{2} \frac{\partial \sigma_x^2}{\partial t}.$$

Thus, we could modify our above results by re-scaling the time coordinate with an exponent  $0 \leq \kappa \leq 2$  to obtain

$$\bar{P}(t) = \frac{1}{2} + \frac{1}{2} \exp \left( -2D_r \Delta t \left( \frac{t}{\Delta t} \right)^{1+\kappa} \right).$$

This could be done separately at the pool level and the binomial levels, introducing two new parameters into our model to alleviate the assumption of Markovianity behind Equation (7).

#### 4 Conclusion and Outlook

In this study, we investigated the issue of overdispersion in noise that appears on NISQ machines. We proposed a two-level random walk model on the Bloch sphere and tested it against experimental data obtained from the IBM Armonk NISQ machine. Our model assumes that the effect of classical noise can be represented as individual random walkers on the Bloch sphere, while overdispersion is a pool-level random walk that jointly affects the theoretical center of mass of all algorithmic repetitions that are used in quantum computing to obtain observable read-out frequencies. For given parameters (diffusion coefficients for noise, for overdispersion, and for state preparation and measurement), it predicts the probability distribution of observable, noisy frequencies as a function of number of quantum gates applied. It also predicts theoretical bounds for the observable frequencies.

Our model fits very well with the experimental data. We employed Bayesian calibration to highlight the necessity of the newly introduced pool level. It is substantial in strength and statistically significant in its effect, with a log-likelihood ratio of more than 7,000 in favor of our new model, compared to a conventional single-level random walk model. In practice, computationally cheaper optimization algorithms can be used for calibrating the model, e.g., by maximizing Equation (28) with respect to the noise-model parameters, if the parametric uncertainty is not of interest.

Several limitations and aspects for future investigation became apparent:

- The late-time limit for long algorithms in our experimental data is clearly offset against the theoretical 50:50 probabilities. It would be interesting to apply models about the equilibrium (energy-optimal) distribution of qubit states to replace the 50:50 asymptote in our model with an offset asymptote.
- Parts of the experimental data show oscillatory behavior along data acquisition, pointing at coherent error types. Our model could be extended to relax its underlying Markovian assumption in its random walks to account for coherent errors.
- The omission of the non probability relevant angle  $\varphi$  enables the semi-analytic description of the error model presented. While diffusive errors in this angle can be incorporated by the parameters of the model, an explicit modelling of this additional angle would enable representations of many more relevant effects due to systematic errors.
- Our study's scope was focused on data from a single machine (IBM Armonk), which is a single-qubit architecture. This is a limited study based on specific empirical observations. It would definitely be interesting to test the two-level approach on other hardware types. Even though highly speculative, we would anticipate some generalization properties. The two-level structure is partially motivated by two apparent error pathways, inherent per qubit randomness and shared randomness due to environmental effects. This initial motivation still holds for other hardware types.

## Appendices

### A Shifted Legendre Polynomials

For completeness, we provide here the first three shifted Legendre polynomials,

$$\mathcal{L}_0(X) = 1, \tag{29}$$

$$\mathcal{L}_1(X) = 2x - 1, \tag{30}$$

$$\mathcal{L}_2(X) = 6x^2 - 6x + 1. \tag{31}$$

## B Qiskit Specifications

We run our experiments on the open-access one-qubit IBM Armonk machine [12]. The circuits were created and run using the following versions the qiskit packages “qiskit-terra”: “0.19.2,” “qiskit-aer”: “0.10.3,” “qiskit-ignis”: “0.7.0,” “qiskit-ibmq-provider”: “0.18.3.” These packages are all part of the qiskit framework and provide the user with the needed functionality to write and perform quantum circuits on the IBM hardware or simulators. The exact explanation of the functions contained in each package and version can be found on the IBM quantum website [12].

## References

- [1] Milton Abramowitz and Irene A. Stegun. 1968. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. Vol. 55. U.S. Government printing office.
- [2] Robin Blume-Kohout, Marcus P. da Silva, Erik Nielsen, Timothy Proctor, Kenneth Rudinger, Mohan Sarovar, and Kevin Young. 2022. A taxonomy of small markovian errors. *PRX Quantum* 3, 2 (May 2022), 020335. DOI : <https://doi.org/10.1103/PRXQuantum.3.020335>
- [3] George E. P. Box and George C. Tiao. 2011. *Bayesian Inference in Statistical Analysis*. John Wiley & Sons.
- [4] George Casella and Roger L. Berger. 2021. *Statistical Inference*. Cengage Learning.
- [5] Siddhartha Chib. 2001. Markov chain Monte Carlo methods: Computation and inference. *Handbook of Econometrics* 5 (2001), 3569–3649.
- [6] A. W. F. Edwards. 2000. Gilberts sine distribution. *Teach. Stat.* 22, 3 (2000), 70–71. DOI : <https://doi.org/10.1111/1467-9639.00026>
- [7] J. Eisert, D. Hangleiter, N. Walk, I. Roth, D. Markham, R. Parekh, U. Chabaud, and E. Kashefi. 2020. Quantum certification and benchmarking. *Nature Rev. Phys.* 2 (2020), 382–390.
- [8] J. Emerson, R. Alici, and K. Zyczkowski. 2005. Scalable noise estimation with random unitary operators. *J. Optics B: Quantum Classic. Opt.* 3 (2005), S347–S352.
- [9] J. Emerson, M. Silva, O. Moussa, C. Ryan, M. Laforest, J. Baugh, D. Cory, and R. Laflamme. 2022. Symmetrized characterization of noisy quantum processes. *Science* 137, 5846 (2022), 1893–1896.
- [10] Catherine Forbes, Merran Evans, Nicholas Hastings, and Brian Peacock. 2011. *Statistical Distributions*. John Wiley & Sons.
- [11] J. Helsen, I. Roth, E. Onorati, A. H. Werner, and J. Eisert. 2022. General framework for randomized benchmarking. *PRX Quantum* 3 (2022).
- [12] IBM Quantum Experience. Retrieved June 2022 <https://quantum-computing.ibm.com/>
- [13] Samuel Karlin. 2014. *A First Course in Stochastic Processes*. Academic Press.
- [14] Thomas Monz, Philipp Schindler, Julio T. Barreiro, Michael Chwalla, Daniel Nigg, William A. Coish, Maximilian Harlander, Wolfgang Hänsel, Markus Hennrich, and Rainer Blatt. 2011. 14-qubit entanglement: Creation and coherence. *Phys. Rev. Lett.* 130506, 106 (2011). DOI : <https://doi.org/10.1103/PhysRevLett.106.130506>
- [15] Erik Nielsen, John King Gamble, Kenneth Rudinger, Travis Scholten, Kevin Young, and Robin Blume-Kohout. 2021. Gate set tomography. *Quantum* 5, 557 (Oct. 2021). DOI : <https://doi.org/10.22331/q-2021-10-05-557>
- [16] Erik Nielsen, Kenneth Rudinger, Timothy Proctor, Kevin Young, and Robin Blume-Kohout. 2021. Efficient flexible characterization of quantum processors with nested error models. *New J. Phys.* 23, 9 (Sep. 2021), 093020. DOI : <https://doi.org/10.1088/1367-2630/ac20b9>
- [17] Michael A. Nielsen and Isaac L. Chuang. 2010. *Quantum Computation and Quantum Information*. Cambridge University Press.
- [18] Corey Ostrove, Erik Nielsen, Kevin Young, and Robin Blume-Kohout. 2022. Out-of-model effects and overdispersion in gate set tomography. In *APS March Meeting Abstracts (APS Meeting Abstracts)*, Vol. 2022. Article M38.007, M38.007 pages.
- [19] Athanasios Papoulis and S. Unnikrishna Pillai. 2002. *Probability, Random Variables and Stochastic Processes*. McGraw–Hill.
- [20] Jonathan Saragosti, Pascal Silberzan, and Axel Buguin. 2012. Modeling E. coli tumbles by rotational diffusion. implications for chemotaxis. *PLoS One* 7, 4 (Apr. 2012). DOI : <https://doi.org/10.1371/journal.pone.0035412>
- [21] A. K. Sharma and R. Sharma. 2021. Effect of noise on the performance of deterministic CNOT gate for photonic Qubits. *Opt. Quant. Electron.* 607 (2021). DOI : <https://doi.org/10.1007/s11082-021-03249-4>
- [22] M. Mohamed El Kordy Shehata, George Simion, Ruoyu Li, Fahd A. Mohiyaddin, Danny Wan, Massimo Mongillo, Bogdan Govoreanu, Iuliana Radu, Kristiaan De Greve, and Pol Van Dorpe. 2023. Modeling semiconductor spin qubits and their charge noise environment for quantum gate fidelity estimation. *Phys. Rev. B* 108, 4 (Jul. 2023), 045305. DOI : <https://doi.org/10.1103/PhysRevB.108.045305>

- [23] Joel J. Wallman and Steven T. Flammia. 2014. Randomized benchmarking with confidence. *New J. Phys.* 16, 10 (Oct. 2014), 103032. DOI : <https://doi.org/10.1088/1367-2630/16/10/103032>
- [24] Y. H. Wang. 1993. On the number of successes in independent trials. *Stat. Sinica* 3, 2 (1993), 295–312.
- [25] C. D. Wilen, S. Abdullah, N. A. Kurinsky, C. Stanford, L. Cardani, G. D’Imperio, C. Tomei, L. Faoro, L. B. Ioffe, C. H. Liu, B. G. Christensen, J. L. DuBois, and R. McDermott. 2021. Correlated charge noise and relaxation errors in superconducting qubits. *Nature* 594 (2021), 369–373. DOI : <https://doi.org/10.1038/s41586-021-03557-5>
- [26] Karen Wintersperger, Florian Dommert, Thomas Ehmer, Andrey Hoursanov, Johannes Klepsch, Wolfgang Maurer, Georg Reuber, Thomas Strohm, Ming Yin, and Sebastian Luber. 2023. Neutral atom quantum computing hardware: Performance and end-user perspective. *Arxiv Open Access* 607 (2023).

Received 6 February 2024; revised 8 July 2024; accepted 28 July 2024