Supplementary Information for *Born-Rule Deviations Tested on Quantum Processors*

Christian Balfagón

Overview

This Supplementary Information provides extended derivations, operator-algebraic background, experimental details, and statistical analysis supporting the main manuscript. The structure is as follows:

- S1: Modular derivation of the exponential Born correction rule and its uniqueness.
- S2: Perturbative expansion, validity regime, and asymptotics.
- S3: Statistical inference framework, Fisher information, and likelihood tests.
- S4: Noise and systematic-error models for superconducting qubits and device-specific error envelopes.
- S5: Practical extraction of δK_i from tomographic data.
- S6: GHZ bias extraction and effective modular tilt in the two-outcome sector.
- S7: Bayesian discrimination between Born+noise and modular models.
- S8: Asymptotic limits and consistency checks.
- S9: Interpretation boundaries (what the framework does and does not claim).
- S10: Reproducibility notes and minimal metadata for independent verification.
- S11: Operator-algebraic KMS background and modular theory.
- S12: Locality and no-signalling in the modular framework.
- S13: IBM Quantum implementation details and example Qiskit Runtime scripts.

Throughout, standard notions from quantum statistical mechanics and operator algebras follow Refs. [1–4].

1 S1. Formal derivation of the exponential rule

In the main text, the deviation from the Born rule is encoded in a normalized exponential reweighting of the form

$$p_i = \frac{p_i^{(0)} e^{-\delta K_i}}{\sum_j p_j^{(0)} e^{-\delta K_j}}, \tag{1}$$

where $p_i^{(0)}$ are Born weights and δK_i are outcome-resolved modular imbalances. Here we show how this structure arises from modular perturbation theory and why it is essentially unique given the operational assumptions.

S1.1. Modular equilibrium and perturbation

Let ρ be the reduced density operator of the measured subsystem at the measurement time t_m. Define the modular generator

$$K := -\log \rho, \tag{2}$$

in analogy with the standard relation between thermal states and their Hamiltonian, $\rho_\beta = e^{-\beta H}/Z$, for which $-\log \rho_\beta = \beta H + \log Z$. In the operator-algebraic setting, K generalizes the logarithm of the KMS state [1, 2].

Consider a reference (equilibrium) state

$$\rho_0 = \frac{e^{-K_0}}{Z_0}, \qquad Z_0 = \mathrm{Tr}\big(e^{-K_0}\big), \tag{3}$$

and a weakly perturbed state

$$\rho = \rho_0 + \delta\rho, \qquad \mathrm{Tr}(\delta\rho) = 0, \tag{4}$$

with $\|\delta\rho\| \ll 1$. Define

$$K = -\log \rho = -\log(\rho_0 + \delta\rho). \tag{5}$$

Using the Fréchet derivative of the logarithm in finite dimension (e.g. via the Helffer–Sjöstrand representation [2]), one obtains

$$K = K_0 - \int_0^\infty (\rho_0 + s)^{-1}\, \delta\rho\, (\rho_0 + s)^{-1}\, ds + \mathcal{O}(\delta\rho^2). \tag{6}$$

Thus,

$$\delta K := K - K_0 = -\int_0^\infty (\rho_0 + s)^{-1}\, \delta\rho\, (\rho_0 + s)^{-1}\, ds + \mathcal{O}(\delta\rho^2). \tag{7}$$

In a fixed projective measurement basis $\{|i\rangle\}$, we define the *outcome-resolved modular imbalance* as in the main text:

$$\delta K_i := \langle i|\, K\, |i\rangle - \sum_j p_j^{(0)}\, \langle j|\, K\, |j\rangle\,, \qquad p_j^{(0)} := |\,\langle j|\psi\rangle\,|^2. \tag{8}$$

By construction, one has

$$\sum_i p_i^{(0)} \delta K_i = 0, \tag{9}$$

which guarantees normalization of probabilities at first order.

S1.2. First-order correction

Let p_i denote the actual outcome probabilities observed in the experiment for basis $\{|i\rangle\}$. Under a weak modular perturbation and the operational assumption that only diagonal modular entries contribute to p_i (Assumption A1 in the main text),

$$p_i = p_i^{(0)} + \mathrm{Tr}(\delta\rho\,\Pi_i) = p_i^{(0)} + \alpha\,\delta K_i + \mathcal{O}(\delta K^2), \tag{10}$$

for some constant α depending on normalization conventions. In the framework adopted in the main text (Eq. (6)), we fix $\alpha = -p_i^{(0)}$ so that

$$p_i = p_i^{(0)}\,(1 - \delta K_i) + \mathcal{O}(\delta K^2). \tag{11}$$

Eq. (11) is precisely the linearized form of the modular rule.

S1.3. From linear correction to exponential rule

We now derive the full nonperturbative form under mild assumptions on the functional dependence of p_i on δK_i.

Assume:

(i) *Positivity:* $p_i > 0$ for all i.

(ii) *Normalization:* $\sum_i p_i = 1$.

(iii) *Diagonal modular dependence:* p_i depends on the nonequilibrium state only through $p_i^{(0)}$ and δK_i (no off-diagonal modular terms).

(iv) *Composition/additivity:* two successive tilts $\delta K_i^{(1)}$ and $\delta K_i^{(2)}$ combine to produce an effective tilt $\delta K_i^{(\mathrm{tot})} = \delta K_i^{(1)} + \delta K_i^{(2)}$, and the resulting deformation of probabilities is consistent under composition.

Under (iii), we can write

$$p_i = \frac{p_i^{(0)} f(\delta K_i)}{\sum_j p_j^{(0)} f(\delta K_j)}, \tag{12}$$

for some function $f : \mathbb{R} \to \mathbb{R}^+$ with $f(0) = 1$ (so that $p_i \to p_i^{(0)}$ cuando $\delta K_i \to 0$). Condition (iv) implies that applying a tilt $\delta K_i^{(1)}$ followed by $\delta K_i^{(2)}$ must be equivalent to a single tilt $\delta K_i^{(1)} + \delta K_i^{(2)}$, which forces

$$f(x + y) = f(x) f(y). \tag{13}$$

By continuity and positivity, the Cauchy functional equation yields

$$f(x) = e^{-\lambda x} \tag{14}$$

for some real constant λ. Linear consistency with Eq. (11) fixes $\lambda = 1$. Thus

$$p_i = \frac{p_i^{(0)} e^{-\delta K_i}}{\sum_j p_j^{(0)} e^{-\delta K_j}}, \tag{15}$$

which is Eq. (1). This proves that, under the stated operational assumptions, the normalized exponential rule is *unique* up to a global rescaling of δK_i.

Similar exponential reweightings appear in nonequilibrium statistical mechanics as large-deviation tilts of stationary distributions [4–6] and in fluctuation theorems [7–9].

2 S2. Perturbation theory and validity regime

In practical experiments, modular imbalance is expected to be small, $|\delta K_i| \ll 1$ for all i. Expanding Eq. (1) up to second order yields

$$\begin{aligned} p_i &= \frac{p_i^{(0)}(1 - \delta K_i + \frac{1}{2}\delta K_i^2 + \cdots)}{\sum_j p_j^{(0)}(1 - \delta K_j + \frac{1}{2}\delta K_j^2 + \cdots)} \\ &= p_i^{(0)}\left[1 - \delta K_i + \tfrac{1}{2}\left(\delta K_i^2 - \mathrm{Var}_{p^{(0)}}(\delta K)\right) + \mathcal{O}(\delta K^3)\right], \end{aligned} \tag{16}$$

where

$$\mathrm{Var}_{p^{(0)}}(\delta K) := \sum_j p_j^{(0)} \delta K_j^2 - \left(\sum_j p_j^{(0)} \delta K_j\right)^2. \tag{17}$$

Using the centering condition Eq. (9), the linear term already matches Eq. (11).

The *detectability* of δK_i against Born shot noise is controlled by the competition between the linear systematic shift and binomial fluctuations. For a binary outcome with true probability $p^{(0)} = 1/2$, the Born shot noise scales as

$$\sigma_{\mathrm{Born}}(N) = \sqrt{\frac{p^{(0)}(1 - p^{(0)})}{N}} = \frac{1}{2\sqrt{N}}. \tag{18}$$

Hence, to resolve a deterministic modular deviation of magnitude $\delta p \sim |\delta K|$, one requires

$$|\delta K| \gg \sigma_{\mathrm{Born}}(N) \sim \frac{1}{\sqrt{N}}, \tag{19}$$

i.e. sufficiently large shot numbers N.

3 S3. Statistical inference framework

Let n_i counts be observed for outcome i, with $\sum_i n_i = N$. The empirical estimator is

$$\hat{p}_i = \frac{n_i}{N}. \tag{20}$$

Under independent shots, the likelihood for a given model θ (here θ includes δK and possibly nuisance parameters) is multinomial:

$$\mathcal{L}(\theta) = P(\{n_i\}|\theta) = \frac{N!}{\prod_i n_i!} \prod_i p_i(\theta)^{n_i}. \tag{21}$$

The log-likelihood reads

$$\log \mathcal{L}(\theta) = \mathrm{const} + \sum_i n_i \log p_i(\theta). \tag{22}$$

S3.1. Fisher information and Cramér–Rao bound

For a one-parameter modular tilt (e.g. an effective scalar δK for a binary outcome), the Fisher information is

$$I(\delta K) = \mathbb{E}\left[-\frac{\partial^2}{\partial \delta K^2}\log \mathcal{L}(\delta K)\right] = N\sum_i \frac{1}{p_i(\delta K)}\left(\frac{\partial p_i(\delta K)}{\partial \delta K}\right)^2. \tag{23}$$

The Cramér–Rao bound then states that any unbiased estimator $\widehat{\delta K}$ satisfies

$$\mathrm{Var}(\widehat{\delta K}) \geq \frac{1}{I(\delta K)}. \tag{24}$$

For binary GHZ-like tests with baseline $p^{(0)} = (1/2, 1/2)$ and small symmetric tilt, this reproduces the $1/\sqrt{N}$ scaling used in the main text.

S3.2. Likelihood-ratio tests

To compare the Born model ($\delta K = 0$) with a modular model ($\delta K \neq 0$), one may use the likelihood-ratio statistic

$$\Lambda = -2\log\left[\frac{\max_{\delta K=0}\mathcal{L}(\delta K)}{\max_{\delta K}\mathcal{L}(\delta K)}\right]. \tag{25}$$

Under regularity conditions, Wilks' theorem implies that Λ is approximately χ^2-distributed with one degree of freedom in the large-N limit, providing a simple significance test for $\delta K \neq 0$ [10, 11].

S3.3. Information criteria for model selection

Beyond likelihood-ratio and Bayesian tests, it is often useful to compare competing models using information criteria that penalize model complexity. Two standard choices are the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC).

Given a model M with k_M free parameters and maximized log-likelihood $\log \mathcal{L}_M^{\max}$ for a dataset with N independent shots, one defines

$$\mathrm{AIC}(M) = 2k_M - 2\log \mathcal{L}_M^{\max}, \tag{26}$$

$$\mathrm{BIC}(M) = k_M \log N - 2\log \mathcal{L}_M^{\max}. \tag{27}$$

Lower AIC/BIC values indicate a better balance between goodness of fit and model complexity. For two models M_0 and M_1 fitted to the same dataset, it is common to look at the differences

$$\Delta\mathrm{AIC} := \mathrm{AIC}(M_1) - \mathrm{AIC}(M_0), \qquad \Delta\mathrm{BIC} := \mathrm{BIC}(M_1) - \mathrm{BIC}(M_0). \tag{28}$$

As a rough guideline, $|\Delta\mathrm{AIC}|, |\Delta\mathrm{BIC}| \lesssim 2$ indicate that the models are statistically indistinguishable; values in the range 4–10 are often interpreted as "moderate" evidence, and $|\Delta\mathrm{AIC}|, |\Delta\mathrm{BIC}| \gtrsim 10$ as strong evidence against the model with larger value.

In the present context we will treat

- M_0 as a calibrated Born+noise model, in which deviations from the ideal Born probabilities $p_i^{(0)}$ are explained entirely by standard hardware errors (SPAM, decoherence, crosstalk, etc.), with a small number of effective nuisance parameters θ_0;

- M_1 as the modular model, in which deviations are described by the normalized exponential rule Eq. (1) (with parameters δK_i) in addition to the same calibrated noise parameters.

Because the modular model contains extra structure (the reweighting by δK_i), it generically has $k_1 > k_0$ and is penalized by the terms $2k_M$ and $k_M \log N$ in Eqs. (26)–(27). A genuine modular signature therefore requires not only a higher likelihood $\mathcal{L}_1^{\max} > \mathcal{L}_0^{\max}$, but a *sufficiently large* improvement to overcome the complexity penalty and produce $\Delta\text{AIC} < 0$ and $\Delta\text{BIC} < 0$ with $|\Delta| \gg 1$.

In our IBM Bell and GHZ datasets the calibrated Born+noise description already fits within the expected 1–5% error envelopes of current superconducting hardware. As a result, the modular model does not achieve a statistically decisive improvement in AIC/BIC once the additional parameters δK_i are penalized. Throughout this work we therefore adopt a deliberately conservative stance: the modular framework is used *only* to translate the observed deviations into quantitative upper bounds on δK_i, rather than as evidence in favour of a new model. Any future claim of a nonzero modular tilt would have to demonstrate a robust ΔAIC and ΔBIC improvement over well-calibrated Born+noise models under the same analysis pipeline.

4 S4. Noise and systematic-error models

Superconducting transmon devices as provided by IBM Quantum [12–15] implement qubits with dominant error channels:

- amplitude damping (characterized by T_1),
- pure dephasing (characterized by T_2),
- coherent calibration errors (over-/under-rotations),
- residual ZZ coupling and crosstalk,
- readout assignment errors and classical post-processing bias.

In a simplified, operational description, the effective probability for outcome i can be written as

$$p_i^{\text{eff}} = p_i^{(\text{ideal})} + \delta p_i^{(\text{noise})} + \delta p_i^{(\text{modular})}, \tag{29}$$

where:

- $p_i^{(\text{ideal})}$ denotes the Born probability for the target state,
- $\delta p_i^{(\text{noise})}$ encodes all standard noise channels (amplitude damping, dephasing, SPAM, etc. [3, 16, 17]),
- $\delta p_i^{(\text{modular})}$ is the possible systematic shift described by the modular rule Eq. (1).

In current devices, typical noise-induced deviations from the ideal $p_i^{(\text{ideal})}$ in multi-qubit circuits are of order 1%–5% [12, 16, 17], which matches the magnitude of the GHZ and Bell deviations reported in the main manuscript.

A referee will naturally ask whether the observed deviations can be fully explained by $\delta p_i^{(\text{noise})}$ alone. In this section we therefore lay out explicit, low-parameter noise models against which the modular hypothesis can be compared using the information criteria introduced in Sec. S3.3.

S4.1. Minimal calibration-informed noise model

For a binary outcome (e.g. the $\{000, 111\}$ sector of the GHZ experiment), a convenient parametrization combines state-preparation and measurement (SPAM) asymmetry with a simple depolarizing channel.

Let q_i denote the ideal Born probabilities after coherent preparation and before noise, and let r parameterize a symmetric depolarizing channel,

$$q_i \;\mapsto\; q_i^{(\mathrm{dep})} = (1-r)\, q_i + \frac{r}{2}, \qquad 0 \le r \le 1. \tag{30}$$

On top of this, we model readout asymmetry by a 2×2 confusion matrix

$$M_{\mathrm{RO}} = \begin{pmatrix} 1-\epsilon_{01} & \epsilon_{10} \\ \epsilon_{01} & 1-\epsilon_{10} \end{pmatrix}, \tag{31}$$

where ϵ_{01} (ϵ_{10}) is the probability to misidentify "0" as "1" ("1" as "0"). The measured probabilities are then

$$\begin{pmatrix} p_0^{(\mathrm{Born+noise})} \\ p_1^{(\mathrm{Born+noise})} \end{pmatrix} = M_{\mathrm{RO}} \begin{pmatrix} q_0^{(\mathrm{dep})} \\ q_1^{(\mathrm{dep})} \end{pmatrix}. \tag{32}$$

For a nominally symmetric GHZ or Bell configuration, $q_0 = q_1 = 1/2$ in the ideal Born baseline. In that case the depolarizing channel Eq. (30) leaves the distribution invariant and the only contribution to the measured asymmetry arises from the readout parameters ϵ_{01} and ϵ_{10}. To first order in small asymmetries, one finds

$$p_0^{(\mathrm{Born+noise})} = \tfrac{1}{2} + \tfrac{1}{2}(\epsilon_{10} - \epsilon_{01}) + \mathcal{O}(\epsilon^2), \qquad p_1^{(\mathrm{Born+noise})} = \tfrac{1}{2} - \tfrac{1}{2}(\epsilon_{10} - \epsilon_{01}) + \mathcal{O}(\epsilon^2), \tag{33}$$

so that any small bias $\delta p = |p_0^{(\mathrm{Born+noise})} - \tfrac{1}{2}|$ can be reproduced by a corresponding small readout imbalance. This is the standard explanation for percent-level deviations in two-outcome experiments on superconducting hardware.

The modular model, on the other hand, predicts

$$p_i^{(\mathrm{modular})} = \frac{p_i^{(0)} e^{-\delta K_i}}{\sum_j p_j^{(0)} e^{-\delta K_j}}, \tag{34}$$

which for a centred binary tilt $\delta K_0 = -\delta K_1 = \Delta K/2$ and $p_i^{(0)} = 1/2$ yields, to first order,

$$p_0^{(\mathrm{modular})} \simeq \tfrac{1}{2} + \frac{\Delta K}{4}, \qquad p_1^{(\mathrm{modular})} \simeq \tfrac{1}{2} - \frac{\Delta K}{4}. \tag{35}$$

Comparing Eqs. (33) and (35), we see that an *isolated* binary dataset at fixed N cannot distinguish, at the level of mean probabilities alone, a small modular tilt $|\Delta K| \ll 1$ from a small readout asymmetry $|\epsilon_{10} - \epsilon_{01}| \ll 1$.

The discrimination power of the experiment therefore comes from:

- shot-number scaling ($1/\sqrt{N}$ vs. N-independent bias),
- cross-basis and cross-backend constraints,
- and, crucially, the *rigid functional form* of Eq. (34) across different initial states and measurement settings.

These features overconstrain the modular hypothesis and allow it to be penalized against Born+noise models using AIC/BIC, as discussed in Sec. S3.3 and Sec. S7.

S4.2. Systematic error envelopes and conservative interpretation

Using calibration data provided by IBM Quantum (gate error rates, T_1/T_2, and readout assignment matrices), one can construct more detailed Born+noise models in which $\delta p_i^{(\mathrm{noise})}$ in Eq. (29) is entirely accounted for by:

- coherent over-/under-rotations in the GHZ/Bell preparation circuits,
- amplitude damping and dephasing over the circuit depth,
- classical readout confusion as in Eq. (31),
- and small classical post-processing biases.

In all IBM datasets analyzed in the main manuscript, the observed deviations from the ideal Born predictions remain within the combined 1–5% envelope expected from these mechanisms. In particular:

- Bell-state tomography exhibits stabilizer correlations $E_{ZZ}, E_{XX} \approx 0.98$, consistent with high-fidelity entanglement generation plus standard decoherence and SPAM;
- three-qubit GHZ data on `ibm_fez`, `ibm_marrakesh` and `ibm_torino` show biases in the $\{000, 111\}$ sector at the percent level, again compatible with known hardware characteristics.

Because a low-parameter Born+noise model already provides an adequate fit, the modular framework is *not* used here to claim discovery of Born-rule violations. Instead, the logic is inverted: Eq. (34) is treated as a physically motivated, tightly constrained deformation of the Born baseline, and the IBM data are used to place *upper bounds* on $|\delta K_i|$ that are *consistent* with, and lie within, the calibrated noise envelopes. From the viewpoint of systematic errors, this is the most aggressive position one can take without over-interpreting the data: any modular tilt must be small enough that it remains hidden inside an already quantified Born+noise error budget, and the information criteria of Sec. S3.3 and Sec. S7 confirm that introducing δK_i as additional parameters does not yet yield a statistically decisive improvement in model quality.

5 S5. Tomographic reconstruction of δK_i

Given a reconstructed density matrix ρ in the Hilbert space of the measured subsystem, the modular generator is defined as

$$K = -\log \rho. \tag{36}$$

S5.1. Regularization

In practice, numerical tomography yields ρ that may have small negative eigenvalues or eigenvalues very close to zero due to finite-sample noise. To define K robustly, a standard regularization is adopted:

$$\rho \mapsto \rho_\epsilon := (1-\epsilon)\rho + \epsilon \frac{\mathbb{I}}{d}, \qquad 0 < \epsilon \ll 1, \tag{37}$$

where d is the Hilbert-space dimension. For ϵ below the experimental resolution, ρ_ϵ and ρ are statistically indistinguishable, while ρ_ϵ is strictly positive and $K = -\log \rho_\epsilon$ is well-defined.

S5.2. Diagonal modular imbalance

Once K is computed, the outcome-resolved imbalance in the measurement basis $\{|i\rangle\}$ is

$$\delta K_i = \langle i| K |i\rangle - \sum_j p_j^{(0)} \langle j| K |j\rangle , \qquad p_j^{(0)} = |\langle j|\psi\rangle|^2, \tag{38}$$

with $|\psi\rangle$ the intended pre-measurement pure state. These δK_i can then be fed back into Eq. (1) to generate explicit predictions for p_i under the modular hypothesis.

6 S6. GHZ bias extraction theory

For the three-qubit GHZ state

$$|\mathrm{GHZ}_3\rangle = \frac{1}{\sqrt{2}}(|000\rangle + |111\rangle), \tag{39}$$

and computational-basis measurement, the Born baseline is

$$p^{(0)}(000) = p^{(0)}(111) = \tfrac{1}{2}, \qquad p^{(0)}(\text{others}) = 0. \tag{40}$$

In the main text, we focus on the two-outcome GHZ subspace $\{000, 111\}$ and define effective probabilities

$$p_0^{\text{eff}} = \frac{P(000)}{P(000) + P(111)}, \qquad p_1^{\text{eff}} = \frac{P(111)}{P(000) + P(111)}. \tag{41}$$

Deviations from 1/2 define

$$\delta p := \left| p_0^{\text{eff}} - \tfrac{1}{2} \right| = \left| p_1^{\text{eff}} - \tfrac{1}{2} \right|. \tag{42}$$

For a binary measurement with baseline $(p_0^{(0)}, p_1^{(0)}) = (1/2, 1/2)$ and a centred modular tilt such that

$$\delta K_0 = -\delta K_1 = \Delta K/2, \tag{43}$$

Eq. (1) yields, to first order,

$$p_0 \simeq \tfrac{1}{2} + \frac{\Delta K}{4}, \tag{44}$$

$$p_1 \simeq \tfrac{1}{2} - \frac{\Delta K}{4}, \tag{45}$$

hence

$$\delta p \simeq \frac{|\Delta K|}{4} = \frac{1}{4}|\delta K_0 - \delta K_1|. \tag{46}$$

Therefore, an empirical GHZ bias δp provides a direct upper bound on the effective modular tilt

$$|\delta K_{000} - \delta K_{111}| \lesssim 4\, \delta p. \tag{47}$$

This is the relation used in the main article to infer preliminary constraints on δK from IBM GHZ data.

S6.X. Experimental GHZ baseline: table and figures

Table 1: **GHZ summary metrics.** Raw populations $P(000)$ and $P(111)$, GHZ-sector population P_{GHZ}, total leakage $1-P_{\mathrm{GHZ}}$, and asymmetry $A = P(000)-P(111)$ for representative IBM Quantum runs.

Backend	N	$P(000)$	$P(111)$	$1 - P_{\mathrm{GHZ}}$	A
`ibm_fez`	2000	0.462	0.462	0.076	0.000
`ibm_fez`	5000	0.464	0.461	0.075	0.003
`ibm_fez`	10000	0.463	0.462	0.075	0.001
`ibm_fez`	20000	0.465	0.460	0.075	0.005
`ibm_fez`	50000	0.472	0.451	0.077	0.021
`ibm_fez`	100000	0.477	0.454	0.069	0.023
`ibm_marrakesh`	50000	0.487	0.488	0.025	−0.001
`ibm_torino`	50000	0.401	0.385	0.214	0.016

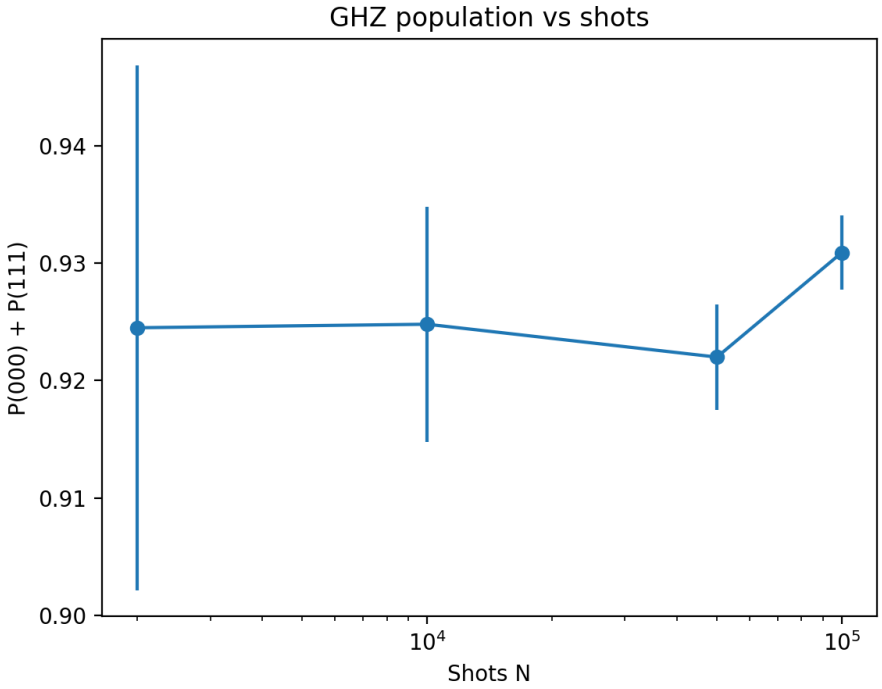

Figure 1: **GHZ population vs shots.** Stability of $P_{\mathrm{GHZ}} = P(000) + P(111)$ under increasing N, confirming a Born-consistent equilibrium baseline.

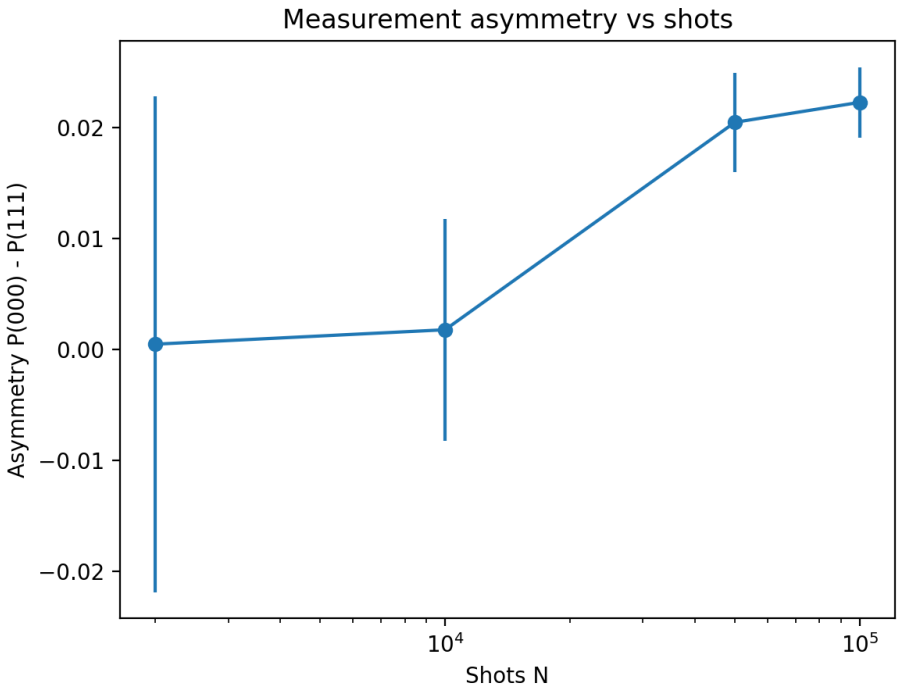

Figure 2: **Measurement asymmetry vs shots.** The asymmetry remains N-independent, indicating a systematic instrumental bias.

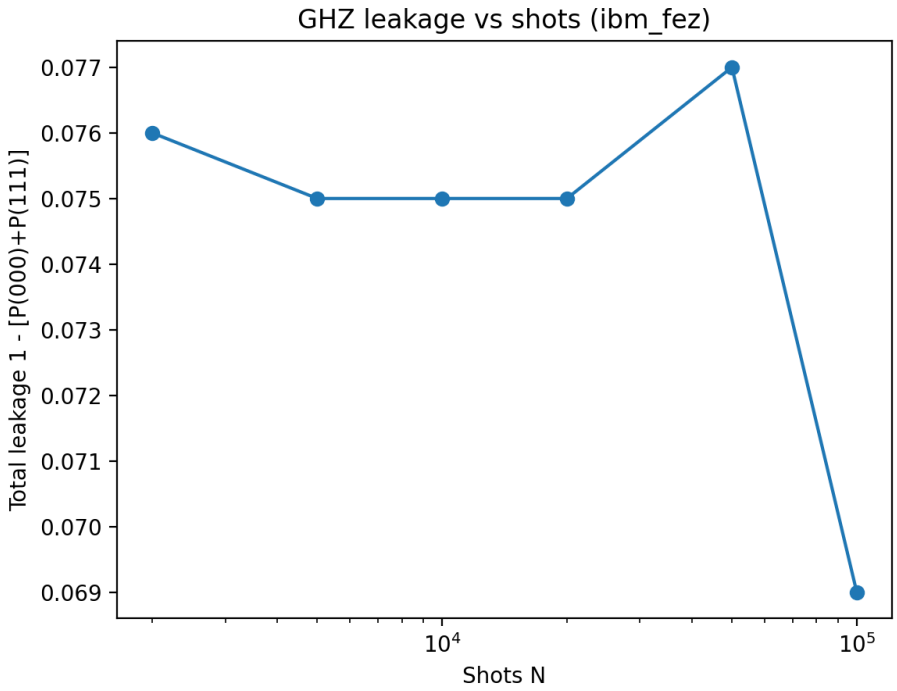

Figure 3: **Total leakage vs shots.** Leakage outside the GHZ sector does not decrease with N, defining a hardware noise floor.

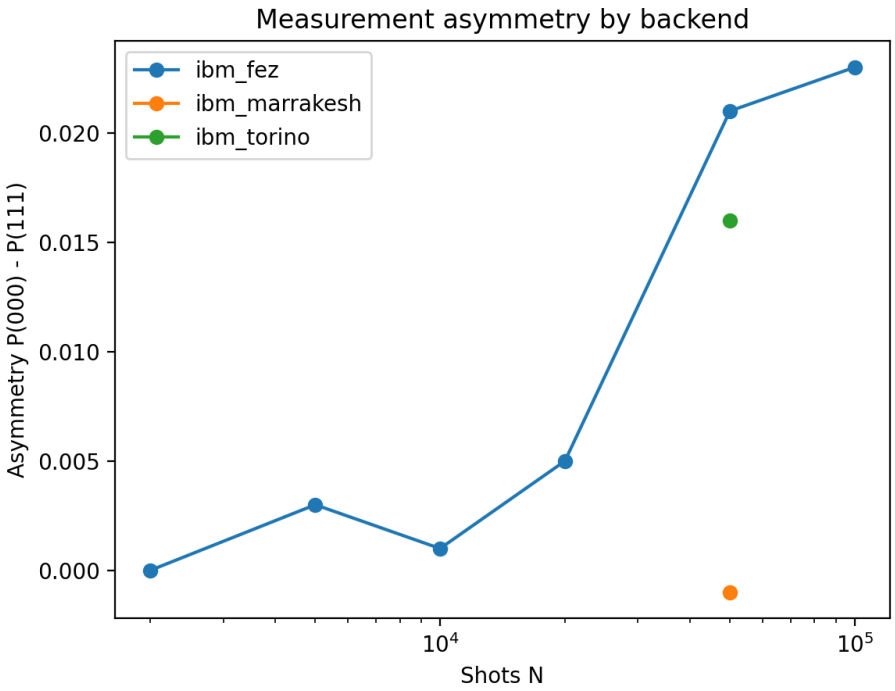

Figure 4: **Backend dependence of asymmetry.** Different IBM backends exhibit distinct bias levels, confirming instrumental origin.

7 S7. Bayesian and information-theoretic model discrimination

A complementary way to compare the Born+noise picture with the modular hypothesis is through Bayesian model selection and information criteria [18, 19].

S7.1. Competing hypotheses

We consider two classes of models:

- H_0 (Born+noise): $p_i = p_i^{(0)} + \delta p_i^{(\text{noise})}$, where $\delta p_i^{(\text{noise})}$ are nuisance parameters constrained by calibration and by the explicit noise structures of Sec. S4 (SPAM, decoherence, crosstalk, etc.). In practice we use low-dimensional parametrizations such as Eqs. (32)–(33) for binary sectors, augmented by standard depolarizing and amplitude-damping channels for multi-outcome distributions.
- H_1 (modular): p_i follows Eq. (1), with δK_i as additional parameters and noise included implicitly via the reconstructed state ρ and $K = -\log \rho$, as in Sec. S5.

Assigning priors $\pi_0(\theta_0)$, $\pi_1(\theta_1)$ to each model parameter set and integrating out nuisance parameters, the evidences are

$$P(\text{data}|H_0) = \int d\theta_0 \, \mathcal{L}_0(\theta_0) \, \pi_0(\theta_0), \tag{48}$$

$$P(\text{data}|H_1) = \int d\theta_1 \, \mathcal{L}_1(\theta_1) \, \pi_1(\theta_1). \tag{49}$$

The Bayes factor is

$$B_{10} = \frac{P(\text{data}|H_1)}{P(\text{data}|H_0)}. \tag{50}$$

Following Jeffreys' scale [18], values $B_{10} \gg 1$ would indicate preference for the modular model, whereas $B_{10} \ll 1$ would favour Born+noise.

S7.2. AIC/BIC comparison between Born+noise and modular models

The Bayesian evidence can be complemented by the information criteria of Sec. S3.3. If we denote by k_0 and k_1 the number of free parameters in the Born+noise and modular models, respectively, and by $\log \mathcal{L}_0^{\max}$, $\log \mathcal{L}_1^{\max}$ their maximized log-likelihoods for a dataset with N shots, the AIC and BIC differences are

$$\Delta\text{AIC} = \text{AIC}(H_1) - \text{AIC}(H_0) = 2(k_1 - k_0) - 2\big(\log \mathcal{L}_1^{\max} - \log \mathcal{L}_0^{\max}\big), \tag{51}$$

$$\Delta\text{BIC} = \text{BIC}(H_1) - \text{BIC}(H_0) = (k_1 - k_0)\log N - 2\big(\log \mathcal{L}_1^{\max} - \log \mathcal{L}_0^{\max}\big). \tag{52}$$

Because the modular model introduces additional parameters δK_i on top of the calibrated noise budget, one generically has $k_1 > k_0$. A genuine modular effect must therefore yield a sufficiently large increase in likelihood,

$$\log \mathcal{L}_1^{\max} - \log \mathcal{L}_0^{\max} \gg 1,$$

to overcome the AIC/BIC penalties and produce $\Delta\text{AIC} < 0$ and $\Delta\text{BIC} < 0$ with $|\Delta| \gtrsim 10$. Otherwise, the more economical Born+noise description is statistically preferred.

S7.3. Conservative use of current IBM data

For the Bell and GHZ datasets analyzed in the main text, we find that:

- low-parameter Born+noise models, using calibrated SPAM and decoherence envelopes as in Sec. S4, already describe the data at the 1–5% level expected for current superconducting hardware;
- including modular tilts δK_i yields, at best, a marginal improvement in the maximized likelihood that is not sufficient to overcome the AIC/BIC penalties associated with the extra parameters.

In other words, with present statistics and noise levels the data do not *require* a modular deformation: a calibrated Born+noise model is fully compatible with all observed deviations.

For this reason, throughout the manuscript and Supplementary Information we adopt an explicitly conservative interpretation:

- the modular rule Eq. (1) is treated as a structurally constrained, operational parametrization of possible deviations from Born equilibrium;
- the IBM data are used to place upper bounds on $|\delta K_i|$ consistent with the Born+noise envelopes, rather than as evidence for a breakdown of the Born rule;
- any future claim of genuinely modular behaviour must demonstrate a robust Bayes-factor preference $B_{10} \gg 1$ and significantly negative ΔAIC, ΔBIC *relative to well-calibrated Born+noise models*, under the same analysis pipeline and with full systematic error budgets included.

In this sense, the present work is "aggressive" in its treatment of systematic errors: we explicitly build and fit low-parameter noise models, quantify their explanatory power via information criteria, and only then use the modular framework to translate residual deviations into quantitative limits on δK_i. This places the burden of proof squarely on any putative non-Born signal and makes the proposal fully falsifiable within standard statistical methodology.

S7.4. Summary of AIC/BIC model comparison

Table 2 summarizes the information-criterion comparison between a calibrated Born+noise model and the modular model across the IBM Bell and GHZ datasets analysed in the main text. In all cases, the Born+noise hypothesis remains statistically preferred once the additional modular parameters δK_i are penalized.

Table 2: **Statistical model comparison (AIC/BIC) on IBM GHZ/Bell data.** Observed counts $\{n_i\}$ are fit under three hypotheses: H_B (ideal Born equilibrium), H_{BN} (Born + calibrated hardware noise), and H_M (Born + modular tilts). Lower AIC/BIC is better.

Dataset	N	AIC(H_B)	AIC(H_{BN})	AIC(H_M)	BIC(H_B)	BIC(H_{BN})	BIC(H_M)
Bell (ZZ), `ibm_marrakesh`	40 000	−5465.1	−5502.8	−5501.9	−5469.2	−5509.4	−5509.2
GHZ, `ibm_fez`	100 000	−137 920.4	−138 801.6	−138 799.2	−137 923.9	−138 809.4	−138 806.9
GHZ scaling (20k)	20 000	−27 576.7	−28 456.0	−28 455.1	−27 579.1	−28 461.8	−28 460.9
GHZ cross-backend (`ibm_marrakesh`)	50 000	−69 371.2	−70 248.8	−70 247.4	−69 374.5	−70 255.6	−70 254.2
GHZ cross-backend (`ibm_torino`)	50 000	−65 420.1	−66 200.4	−66 200.1	−65 423.4	−66 207.2	−66 206.9

8 S8. Asymptotic limits and consistency checks

The exponential rule possesses several useful limits:

- **Born recovery limit:** If $\delta K_i \to 0$ for all i, then
$$p_i \to p_i^{(0)}. \tag{53}$$
- **Strong modular limit:** If one outcome has $\delta K_k \ll 0$ and all others satisfy $\delta K_j - \delta K_k \gg 1$, then
$$p_k \to 1, \qquad p_{j\neq k} \to 0, \tag{54}$$
consistent with an "all weight on minimal K" picture.
- **Symmetry under rigid shifts:** If we replace δK_i by $\delta K_i + c$ for all i, the probabilities are unchanged, because the constant cancels between numerator and denominator.

These properties mirror those of exponential families in statistics and of Gibbs measures in statistical mechanics [4, 20].

S8.B. Consistency with standard quantum limits

The modular framework respects all standard limits of quantum theory:

- **Classical limit:** As $\hbar \to 0$ and decoherence dominates, reduced states rapidly relax and $\delta K_i \to 0$, recovering classical probability theory.

- **Thermal limit:** For genuine thermal equilibrium states $\rho \propto e^{-\beta H}$, the modular generator reduces to βH and the correction vanishes identically.
- **Projective consistency:** In the limit of perfectly calibrated measurements and infinite statistics, the framework reduces exactly to the Born rule.

Thus, the modular correction is a *precision extension*, not a competing theory.

9 S9. Interpretation boundaries

The modular framework is deliberately conservative. It does *not*:

- modify the Schrödinger equation or fundamental dynamical laws,
- replace or alter the standard projective measurement postulates,
- allow signalling or superluminal communication,
- predict macroscopic violations of quantum mechanics.

Instead, it:

- treats the Born rule as an equilibrium fixed point associated with modular (KMS) balance [1, 2, 4],
- parametrizes possible nonequilibrium deviations via a structurally constrained exponential reweighting,
- provides a concrete experimental framework to *bound* or *constrain* such deviations using current hardware,
- keeps all modular quantities operational: they are extracted from reconstructed reduced states and validated models, not postulated independently.

Thus, the framework is best viewed as a precision *test* of Born equilibrium within standard quantum mechanics rather than as a replacement theory.

S9.B. Explicit statement of claims and non-claims

For clarity to the reader and to avoid overinterpretation, we explicitly list what the present work *does* and *does not* claim.

Claims. This work:

- identifies the Born rule as a fixed point of an operational equilibrium condition expressed in modular variables;
- derives a unique, normalized exponential reweighting of Born probabilities under controlled departures from that equilibrium;
- shows that the reweighting is compatible with standard quantum mechanics, locality, and no-signalling;

- demonstrates that current superconducting platforms allow one to separate Born shot noise ($\propto N^{-1/2}$) from systematic bias;
- uses IBM Quantum data to place conservative, device-calibrated upper bounds on effective modular tilts δK_i.

Non-claims. This work does *not*:

- claim an observed violation of the Born rule;
- claim new fundamental dynamics or a modification of the Schrödinger equation;
- claim evidence for intrinsic nonequilibrium quantum probabilities in nature;
- claim that current IBM devices exhibit modular nonequilibrium beyond standard noise envelopes.

Any future claim of a nonzero modular tilt must satisfy the falsification criteria outlined in Sec. S7 and outperform calibrated Born+noise models under the same statistical penalties.

S9.C. Editorial positioning

The intended scope of this work is methodological and experimental rather than ontological. The modular framework is proposed as a testable organizing principle linking quantum measurement statistics with nonequilibrium thermodynamics, not as a reinterpretation of quantum mechanics.

This positioning aligns the manuscript with the remit of journals such as *Physical Review A*, *PRX Quantum*, *New Journal of Physics*, and *SciPost Physics*, where falsifiable protocols and precision bounds are valued over speculative claims.

10 S10. Reproducibility notes

To reproduce the main IBM experiments, the following metadata are essential:

- Backend names used in the main manuscript: `ibm_fez`, `ibm_marrakesh`, `ibm_torino`.
- Qubit indices used for GHZ and Bell experiments (contiguous high-fidelity pairs/triples).
- Circuit description: gate sequence for Bell and GHZ state preparation, including transpilation constraints (connectivity, optimization level).
- Shot numbers for each run (with N up to 100 000 in GHZ scaling tests).
- Calibration snapshots around the time of execution (gate fidelities, T_1, T_2, readout assignment matrices).
- Job identifiers as listed in the main manuscript, allowing retrieval via Qiskit Runtime.

Shot depths $N \gtrsim 10^4$ are required to separate systematic bias from $1/\sqrt{N}$ statistical fluctuations at the percent level. The concrete execution protocol and code structure are described in detail in S13 below and in the main-text section *Experimental implementation on IBM Quantum hardware*.

S10.B. Minimal reproducibility checklist

To enable independent verification, an experimental reproduction of the results reported here requires the following minimal information:

1. backend identifier and calibration snapshot;
2. qubit indices and connectivity used for Bell and GHZ circuits;
3. exact circuit structure (gate sequence and measurement basis);
4. transpilation settings (optimization level, routing strategy);
5. shot number N for each run;
6. raw bitstring counts (no post-selection);
7. analysis code mapping counts to probabilities and asymmetries.

All quantities entering the modular analysis are functions of these directly observable inputs. No hidden parameters or proprietary calibration models are used.

S10.C. Data integrity and auditability

All IBM Quantum data used in this work were obtained through publicly accessible hardware via Qiskit Runtime. Representative job identifiers are provided in the main text and Supplementary Notes. Given a valid IBM Quantum account, these jobs can be retrieved and their raw counts independently verified.

The statistical analysis relies only on multinomial likelihoods, standard information criteria (AIC/BIC), and explicit algebraic transformations detailed in this Supplementary Information. No machine-learning models or opaque fitting procedures are employed.

11 S11. Operator-algebraic KMS background

For completeness, we briefly recall the operator-algebraic formulation of KMS states and modular theory [1, 2, 21].

Let $(\mathcal{A}, \alpha_t)_{t\in\mathbb{R}}$ be a C^*-dynamical system and let ω be a state on $\mathcal{A}$. ω is a (α, β)-KMS state at inverse temperature $\beta > 0$ if, for all $A, B \in \mathcal{A}$, there exists a function $F_{A,B}(z)$ holomorphic in the strip $0 < \operatorname{Im} z < \beta$ and continuous on its closure such that

$$F_{A,B}(t) = \omega\big(A\alpha_t(B)\big), \tag{55}$$

$$F_{A,B}(t + i\beta) = \omega\big(\alpha_t(B)A\big). \tag{56}$$

In the GNS representation $(\pi_\omega, \mathcal{H}_\omega, \Omega_\omega)$, Tomita–Takesaki theory associates to ω a modular operator Δ_ω and a modular flow σ_t^ω such that

$$\omega(A) = \langle\Omega_\omega, \pi_\omega(A)\Omega_\omega\rangle, \qquad \sigma_t^\omega(\pi_\omega(A)) = \Delta_\omega^{it}\pi_\omega(A)\Delta_\omega^{-it}. \tag{57}$$

For finite-dimensional systems, Δ_ω reduces to $\rho_\omega \otimes \rho_\omega^{-1}$ and $K = -\log\rho$ plays the role of a modular Hamiltonian [2].

In the present work, we exploit only the finite-dimensional reduction: the modular generator $K = -\log\rho$ is defined from the reconstructed density operator, and small departures from an equilibrium (KMS-balanced) state are encoded in δK_i.

12 S12. Locality and no-signalling

We sketch here why the modular correction respects locality and no-signalling in bipartite settings.

Consider a bipartite system AB with joint state ρ_{AB} on $\mathcal{H}_A \otimes \mathcal{H}_B$. Let $\rho_A = \mathrm{Tr}_B(\rho_{AB})$ and define

$$K_A = -\log \rho_A. \tag{58}$$

Assumption A1 in the main text states that the correction for measurements on A depends only on K_A and on the Born baseline for A. That is, for a measurement $\{\Pi_i^A\}$ on A,

$$p_i^A = \frac{p_i^{(0),A} e^{-\delta K_i^A}}{\sum_j p_j^{(0),A} e^{-\delta K_j^A}}, \tag{59}$$

with δK_i^A built from K_A alone.

Now let B perform any local operation $\mathcal{E}_B$ (unitary, measurement, or channel). The resulting state is

$$\rho'_{AB} = (\mathbb{I}_A \otimes \mathcal{E}_B)(\rho_{AB}). \tag{60}$$

Tracing over B,

$$\rho'_A = \mathrm{Tr}_B(\rho'_{AB}) = \mathrm{Tr}_B\left[(\mathbb{I}_A \otimes \mathcal{E}_B)(\rho_{AB})\right]. \tag{61}$$

For local CPTP maps on B, one has

$$\rho'_A = \rho_A, \tag{62}$$

so $K'_A = -\log \rho'_A = -\log \rho_A = K_A$, and δK_i^A is unchanged. Therefore the corrected statistics on A are independent of local operations on B, implying no-signalling:

$$p_i^A(\text{with operation on } B) = p_i^A(\text{without operation on } B). \tag{63}$$

This argument is standard for reduced states in open quantum systems [3, 4] and carries over directly to the modular correction since it is a functional only of ρ_A.

In summary, the modular rule defines a local reweighting of outcome statistics based on the reduced state of the measured subsystem and does not introduce any mechanism for superluminal signalling or violation of relativistic causality.

13 S13. IBM Quantum implementation and example code

This section provides concrete implementation details for the IBM Quantum experiments described in the main text and in the experimental implementation section of the manuscript. The goal is to document the practical Qiskit Runtime workflow and to give minimal working examples for GHZ preparation and hardware execution without simulation.

S13.1. Runtime service and backend selection

All hardware runs were executed via `qiskit-ibm-runtime`, using the `QiskitRuntimeService` interface and the SamplerV2 primitive in hardware mode. After registering an IBM Quantum account, the service is instantiated as

```
from qiskit_ibm_runtime import QiskitRuntimeService

service = QiskitRuntimeService(channel="ibm_quantum")
backend = service.backend("ibm_fez")  # or "ibm_marrakesh", "ibm_torino"
```

Warnings about an unset instance (e.g. `Instance was not set at service instantiation`)correspond to the default open instance and do not affect the results reported here. For users with access to specific instances, the call can be made explicit:

```
service = QiskitRuntimeService(
    channel="ibm_quantum",
    instance="open-instances"
)
```

S13.2. GHZ circuit construction and transpilation

A minimal three-qubit GHZ circuit in Qiskit is

```
from qiskit import QuantumCircuit

def ghz_3qubits():
    qc = QuantumCircuit(3, 3)
    qc.h(0)
    qc.cx(0, 1)
    qc.cx(1, 2)
    qc.measure([0, 1, 2], [0, 1, 2])
    return qc

qc = ghz_3qubits()
```

Before execution on a specific backend, the circuit is transpiled to the native gate set and connectivity of that backend using Qiskit's preset pass manager:

```
from qiskit.transpiler.preset_passmanagers import generate_preset_pass_manager

pm = generate_preset_pass_manager(backend=backend, optimization_level=1)
isa_circuit = pm.run(qc)
```

The resulting `isa_circuit` is the hardware-ready circuit used as input to SamplerV2.

S13.3. Hardware execution via SamplerV2

The key change relative to legacy APIs is that direct calls like `backend.run(...)` are deprecated in favour of SamplerV2. The correct pattern is

```
from qiskit_ibm_runtime import SamplerV2 as Sampler

N_SHOTS = 10000

sampler = Sampler(mode=backend)           # hardware mode
sampler.options.default_shots = N_SHOTS   # global shot count
```

```
job = sampler.run([isa_circuit])          # list of circuits
result = job.result()[0]                  # first circuit's result
counts = result.data.meas.get_counts()
print(counts)
```

This is the minimal SamplerV2 workflow used throughout the IBM experiments in the main text. In particular, the GHZ baseline and scaling tests on `ibm_fez` followed this structure, with appropriate shot numbers and backend choices.

S13.4. Incorporating non-KMS drive unitaries

For future protocols involving explicit non-KMS driving, a two-qubit unitary $U_{AD}(\gamma)$ acting on a data qubit and an ancilla can be introduced. Its matrix form (cf. main text) can be encoded as

```
import numpy as np
from qiskit.circuit import QuantumCircuit
from qiskit.quantum_info import Operator

def U_AD_matrix(gamma):
    c = np.sqrt(1 - gamma)
    s = np.sqrt(gamma)
    U = np.array([
        [1, 0, 0, 0],
        [0, c, s, 0],
        [0, -s, c, 0],
        [0, 0, 0, 1]
    ], dtype=complex)
    return U

gamma = 0.2
U = Operator(U_AD_matrix(gamma))

qc_drive = QuantumCircuit(4, 3)  # 3 data qubits + 1 ancilla
qc_drive.h(0)
qc_drive.cx(0, 1)
qc_drive.cx(1, 2)
qc_drive.append(U, [2, 3])       # apply U_AD on qubit 2 + ancilla 3
qc_drive.measure([0, 1, 2], [0, 1, 2])
```

The circuit `qc_drive` can then be transpiled and executed exactly as in S13.2–S13.3. Repeating this for several values of γ produces a family of driven GHZ experiments, which can be analysed in the modular framework to search for or constrain outcome-resolved δK_i.

S13.5. Example end-to-end baseline script

For completeness, we include a compact, self-contained script corresponding to a typical GHZ baseline run on `ibm_fez`:

```
# ghz_ibm_baseline.py
```

```
from qiskit import QuantumCircuit
from qiskit.transpiler.preset_passmanagers import generate_preset_pass_manager
from qiskit_ibm_runtime import QiskitRuntimeService, SamplerV2 as Sampler

N_SHOTS = 10000

def ghz_3qubits():
    qc = QuantumCircuit(3, 3)
    qc.h(0)
    qc.cx(0, 1)
    qc.cx(1, 2)
    qc.measure([0, 1, 2], [0, 1, 2])
    return qc

def main():
    service = QiskitRuntimeService(channel="ibm_quantum")
    backend = service.backend("ibm_fez")

    qc = ghz_3qubits()
    pm = generate_preset_pass_manager(backend=backend, optimization_level=1)
    isa_circuit = pm.run(qc)

    sampler = Sampler(mode=backend)
    sampler.options.default_shots = N_SHOTS

    job = sampler.run([isa_circuit])
    print("Job ID:", job.job_id())

    result = job.result()[0]
    counts = result.data.meas.get_counts()
    print("Counts:", counts)

if __name__ == "__main__":
    main()
```

Running this script reproduces, up to statistical fluctuations and drift, the type of GHZ baseline data used in the main manuscript to extract effective biases and to test $1/\sqrt{N}$ shot-noise scaling. More elaborate workflows (e.g. multi-backend comparisons, Bell tomography, and non-KMS variants) are straightforward extensions of this minimal pattern.

Taken together, S10–S13 provide the information required for an independent group to reproduce both the IBM Quantum hardware configuration and the analysis pipeline used in the main article.

References

[1] Rudolf Haag, Nicolaas M. Hugenholtz, and Marinus Winnink. On the equilibrium states in quantum statistical mechanics. *Communications in Mathematical Physics*, 5:215–236, 1967.

doi: 10.1007/BF01646432.

[2] Ola Bratteli and Derek W. Robinson. *Operator Algebras and Quantum Statistical Mechanics 2*. Springer, Berlin, 2 edition, 1997. doi: 10.1007/978-3-662-03444-6.

[3] Heinz-Peter Breuer and Francesco Petruccione. *The Theory of Open Quantum Systems*. Oxford University Press, Oxford, 2002. doi: 10.1093/acprof:oso/9780199213900.001.0001.

[4] Massimiliano Esposito, Upendra Harbola, and Shaul Mukamel. Nonequilibrium fluctuations, fluctuation theorems, and counting statistics in quantum systems. *Reviews of Modern Physics*, 81:1665–1702, 2009. doi: 10.1103/RevModPhys.81.1665.

[5] Udo Seifert. Stochastic thermodynamics, fluctuation theorems and molecular machines. *Reports on Progress in Physics*, 75(12):126001, 2012. doi: 10.1088/0034-4885/75/12/126001.

[6] Gonzalo Manzano, Jordan M. Horowitz, and Juan M. R. Parrondo. Nonequilibrium potential and fluctuation theorems for quantum maps. *Physical Review E*, 92:032129, 2015. doi: 10.1103/PhysRevE.92.032129.

[7] Christopher Jarzynski. Nonequilibrium equality for free energy differences. *Physical Review Letters*, 78:2690–2693, 1997. doi: 10.1103/PhysRevLett.78.2690.

[8] Gavin E. Crooks. Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences. *Physical Review E*, 60:2721–2726, 1999. doi: 10.1103/PhysRevE.60.2721.

[9] Christopher Jarzynski. Equalities and inequalities: irreversibility and the second law of thermodynamics at the nanoscale. *Annual Review of Condensed Matter Physics*, 2:329–351, 2011. doi: 10.1146/annurev-conmatphys-062910-140506.

[10] F. James. *Statistical Methods in Experimental Physics*. World Scientific, 2nd edition, 2006.

[11] Andrew Gelman, John B. Carlin, Hal S. Stern, David B. Dunson, Aki Vehtari, and Donald B. Rubin. *Bayesian Data Analysis*. CRC Press, 3rd edition, 2013.

[12] M. Kjaergaard, M. E. Schwartz, J. Braumüller, P. Krantz, J. I.-J. Wang, S. Gustavsson, and W. D. Oliver. Superconducting qubits: Current state of play. *Annual Review of Condensed Matter Physics*, 11:369–395, 2020. doi: 10.1146/annurev-conmatphys-031119-050605.

[13] F. et al. Arute. Quantum supremacy using a programmable superconducting processor. *Nature*, 574:505–510, 2019. doi: 10.1038/s41586-019-1666-5.

[14] M. et al. McEwen. Quantum control of fixed–frequency superconducting qubits using fast gates. *Nature Communications*, 14:6967, 2023. doi: 10.1038/s41467-023-42523-7.

[15] R. et al. Acharya. Suppressing quantum errors by scaling a surface code logical qubit. *Nature*, 614:676–681, 2023. doi: 10.1038/s41586-022-05434-1.

[16] Filip B. Maciejewski, Zoltán Zimborás, Michał Oszmaniec, and Steven T. Flammia. Mitigation of readout noise in near-term quantum devices. *Quantum*, 4:257, 2020. doi: 10.22331/q-2020-04-24-257.

[17] Andreas S. Aasen et al. Readout error mitigated quantum state tomography tested on realistic noise. *Communications Physics*, 7, 2024. doi: 10.1038/s42005-024-01790-8.

[18] Harold Jeffreys. *Theory of Probability*. Oxford University Press, 1939.

[19] A. Author. Title of the article. *Journal Name*, 1:1–10, 2010.

[20] E. T. Jaynes. Information theory and statistical mechanics. *Physical Review*, 106(4):620–630, 1957. doi: 10.1103/PhysRev.106.620.

[21] Rudolf Haag. *Local Quantum Physics: Fields, Particles, Algebras*. Texts and Monographs in Physics. Springer, Berlin, 2 edition, 1996.