

Article

Not peer-reviewed version

Gaussian Boson Sampling as a Calibrated Variance-Reduction Engine: A DSFL Explanation

[Camilla Josephson](#)*

Posted Date: 27 October 2025

doi: 10.20944/preprints202510.2042.v1

Keywords: Gaussian Boson Sampling; hafnian; Wick's theorem; variance reduction; importance sampling; probability estimator; Monte Carlo; sample complexity; (ϵ, δ) guarantees; calibration; photon-number matching; Deterministic Statistical Feedback Law (DSFL)



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

Gaussian Boson Sampling as a Calibrated Variance-Reduction Engine: A DSFL Explanation

Camilla Josephson

Independent Researcher, Sweden; cmkjosephson@gmail.com

Abstract

Gaussian Boson Sampling (GBS) generates photon-count patterns whose probabilities are governed by hafnian functionals of an encoded covariance. Recent theory shows that GBS samples can approximate Gaussian expectation problems with provable sample-complexity advantages over Monte Carlo (MC) on an open, sizable subset of instances. We give a unifying explanation via the Deterministic Statistical Feedback Law (DSFL). First, GBS acts as a calibrated, admissible map that intertwines target statistics and physical sampling. Second, tuning the mean photon number is a genuine calibration step that minimizes a DSFL residual. Third, principal-angle (Friedrichs) geometry quantifies conditioning and leakage between statistical and sampling features. Fourth, a Lyapunov envelope provides non-asymptotic error decay and design rules. From these ingredients we derive bounds that recover and sharpen known GBS scalings, predict regimes where GBS outperforms MC, and specify a practical hybrid policy that switches by degree. Experiments with simulated distributions and hardware-aligned kernels support the theory and validate the tuning and selection rules.

Keywords: Gaussian Boson Sampling; hafnian; Wick's theorem; variance reduction; importance sampling; probability estimator; Monte Carlo; sample complexity; (ϵ, δ) guarantees; calibration; photon-number matching; Deterministic Statistical Feedback Law (DSFL)

1. Introduction

Gaussian Boson Sampling (GBS) produces photon-count patterns whose probabilities are governed by hafnian functionals of a covariance encoded in a Gaussian photonic state. Beyond its foundational role in quantum advantage, GBS has recently been positioned as a practical engine for approximating *Gaussian expectation problems* (GEPs), where one seeks $\mu = \mathbb{E}[f(X)]$ for $X \sim \mathcal{N}(0, B)$ and analytic or polynomial f . Wick's theorem expresses μ as a hafnian-weighted sum over multi-indices drawn from B , aligning the target with the native GBS statistics. Building on this structure, Andersen and Shan introduced degree-resolved GBS estimators and proved open, nonempty regions in which GBS achieves (ϵ, δ) sample-complexity advantages over Monte Carlo (MC), and further showed that tuning the mean photon number to the task increases the percentage of instances with advantage [1,2].

1.1. Calibrated Residuals, Principal-Angle Conditioning, and Predict-Then-Prove Guarantees

This work advances three aspects that are often treated separately. First, the residual $R_K(t)$ is not merely a distance: it is a *calibrated, minimal, and dynamical* quantity. We prove that every admissible tuning step is L^2 -nonexpansive, so the residual cannot inflate under retuning or composition. Second, conditioning is captured quantitatively by the principal angle θ_F (via its cosine γ), and this geometric control enters *directly* into the variance and MSE formulae, providing sharp two-sided bounds; this bridges statistical estimation, photonic physics, and numerical analysis in a single comparison geometry. Third, the framework is *predict-then-prove*: before sampling, it specifies where to operate (photon-number matching), how conditioning will affect dispersion (principal angles), and what decay rate to expect (Lyapunov envelope), and then establishes these prescriptions as theorems.

1.2. Problem

Practitioners need to *foresee*, before allocating shots, when GBS will outperform MC and *how to tune* the device to realize that advantage. Concretely: given a covariance B and an integrand f , can we predict in advance whether calibrated GBS will require asymptotically fewer samples than MC to reach a prescribed (ϵ, δ) accuracy, and can we prescribe a robust one-parameter tuning rule that maximizes the gain? We seek *design-time* criteria (not post hoc curve fits) that tell us when to use GBS, when to defer to MC, and how to switch between them by degree.

1.3. DSFL Intuition: Forecasting When GBS Beats MC

The DSFL framework predicts, *a priori*, when and how GBS will outperform Monte Carlo (MC) by placing the target statistics and the sampler in a single Hilbert geometry and tracking one quantity—the residual. This yields the following design-time principles:

- **Single yardstick.** Target and sampler are embedded in the same space; a single residual provides a quantitative knob to optimize and monitor.
- **No inflation under tuning.** Retuning, counting, and block composition act as nonexpansive maps; the residual cannot increase under admissible operations.
- **Unique, certifiable optimum.** The dominant variance/MSE factor is strictly convex in the scale and attains its unique minimum at the mean–photon–number match $m_{tB} = 2K$; this yields a robust one-line tuning rule.
- **Predictable conditioning.** Alignment between sampling and blueprint features is quantified by the principal angle θ_F (cosine γ); small γ certifies low dispersion. A computable surrogate from warm-up features allows pre-run assessment.
- **Rates, not just limits.** If each update removes a fixed fraction of the residual, the residual obeys a non-asymptotic exponential envelope, providing a predicted semi-log MSE slope before data collection.
- **Regime map and hybrid policy.** The geometry identifies regimes where GBS is inherently favorable (high degrees) and where MC suffices (low degrees), and prescribes a degree-wise hybrid switch using the same residual/angle diagnostics.
- **Robustness.** Composition and small hardware drifts degrade angles and rates smoothly; re-matching m_{tB} restores near-optimality.

In summary, DSFL turns GBS from a black-box sampler into a calibrated, predictable engine: it specifies where to operate, how conditioning impacts dispersion, and what decay rate to expect, enabling an informed, pre-run decision on when—and by how much—GBS will outperform MC.

1.4. Thesis (DSFL View)

We show that GBS is a *calibrated variance-reduction engine* inside a single Hilbert comparison geometry governed by a Deterministic Statistical Feedback Law (DSFL). This provides one residual to minimize and four provable pillars: (i) *Admissibility & DPI*: counting and retuning are L^2 -nonexpansive and intertwine with the estimator, so the calibrated residual cannot inflate; (ii) *Calibration optimality*: the unique minimizer of the leading variance/MSE factor is the *photon-number match* $m_{tB} = 2K$, which implements the tuning advocated and empirically optimized in [1,2]; (iii) *Angle conditioning*: principal (Friedrichs) angles between the sampling and blueprint subspaces yield sharp two-sided dispersion bounds $(1 \pm \cos \theta_F)$, explaining conditioning and leakage; (iv) *Lyapunov envelope*: a one-step contraction turns into a non-asymptotic exponential error decay with rate $\alpha(t)$ maximized at the photon match.

1.5. Contributions

1. A unified geometry that proves a data–processing inequality (no residual inflation) for tuning and composition, and recovers the (ϵ, δ) envelopes of [1].
2. A one-line strict-convexity proof that *predicts* the optimal tuning rule $m_{tB} = 2K$, matching the optimized estimators in [3].

3. Principal-angle bounds that *forecast* dispersion and a Lyapunov inequality that delivers *exponential, non-asymptotic* decay—together forming practical, pre-run design rules to decide GBS vs. MC and to set t .
4. A regime map that explains, a priori, when calibrated GBS is polynomial while MC is exponential (large degree), and a hybrid, degree-wise policy for the small-degree regime.

In short, DSFL turns “folk rules” into theorems: *match* m_{tB} , *measure* angles, and *expect* exponential envelopes—thereby *foreseeing* and *tuning* the regimes where GBS outperforms MC, in line with and extending the results of [1,2].

2. What GBS Does (Succinct but Precise)

GBS distribution and hafnians

Let $B \in \mathbb{R}^{N \times N}$ be symmetric with $0 < \text{spec}(B) < 1$. The GBS probability of a pattern $I \in \mathbb{N}^N$ is

$$P_B(I) = \frac{d_B}{I!} \text{Haf}(B_I)^2, \quad d_B := \prod_{j=1}^N \sqrt{1 - \lambda_j^2}, \quad (1)$$

where $\{\lambda_j\}$ are eigenvalues of B . The mean photon number is

$$m_B = \sum_{j=1}^N \frac{\lambda_j^2}{1 - \lambda_j^2}. \quad (2)$$

Gaussian expectations via Wick’s theorem

For $f(x) = \sum_{k \geq 0} \sum_{|I|=k} a_I x^I$ and $X \sim \mathcal{N}(0, B)$,

$$\mu = \mathbb{E}[f(X)] = \sum_{k \geq 0} \sum_{|I|=2k} a_I \text{Haf}(B_I). \quad (3)$$

A degree partition gives $\mu = \sum_{k=0}^K \mu_k$ with $\mu_k = \sum_{|I|=2k} a_I \text{Haf}(B_I)$.

Two practical GBS estimators

GBS-I (importance). For $\mu_k^{(2)} = \sum_{|I|=2k} a_I \text{Haf}(B_I)^2$, the estimator using weights $I! a_I / d_B$ is unbiased.

GBS-P (probability). For $\mu_k = \sum_{|I|=2k} a_I \text{Haf}(B_I)$, estimate $p_J = P_B(J)$ by counts and average $\sqrt{p_J}$ with weights. The bias vanishes at rate $1/n$; the leading term is explicit.

Tuning knob: photon-number matching

Scale $B \mapsto tB$ with $t \in (0, \lambda_1^{-1})$ and choose t so that

$$m_{tB} = \sum_{j=1}^N \frac{t^2 \lambda_j^2}{1 - t^2 \lambda_j^2} = 2K, \quad (4)$$

i.e. *match* the sampler’s mean photon number to the target degree K .

3. Related Work

This section situates our DSFL-based treatment of Gaussian Boson Sampling (GBS) within three intertwined literatures: (i) the foundations of GBS, including sampling distributions, computational hardness, and experimental progress; (ii) GBS as a vehicle for numerical estimation, in particular its use for Gaussian expectation problems together with formal (ϵ, δ) -sample-complexity claims and percentage-of-advantage quantification; and (iii) control-theoretic and geometric principles that justify our residual-contraction, angle-conditioned bounds, and data-processing inequalities in a

single Hilbertian comparison geometry. We use these threads to explain how the “folk rules” of GBS operation—photon-number tuning, conditioning via feature alignment, and non-asymptotic error envelopes—can be elevated to theorems, and to place our results alongside prior theory and practice.

3.1. Foundations of GBS and Hardness

The modern formulation of GBS as a sampling task driven by hafnian-valued amplitudes is due to Hamilton *et al.* [4] with detailed modeling and refinements in [5]. Complexity-theoretic hardness follows the linear-optical lineage of Aaronson–Arkhipov [6] and extends to experimentally informed, high-dimensional variants [7,8]. Algorithmic primitives for hafnian evaluation and classical simulation—including supercomputer-grade implementations and the thewalrus library—are documented in [9,10]. Photonic demonstrations approaching or attaining quantum advantage are reported in [11,12]. Related Gaussian models include loop-hafnians for displaced states [13] and threshold-detection (Torontoian) statistics [14], which frame hardware-relevant variants. On the application side, GBS has been explored for vibronic spectra [15,16], molecular docking [17], dense subgraph discovery and graph kernels [18–21], random point processes [22,23], and scalable time-bin/loop architectures [24,25]. These works supply the physical and computational substrate on which our estimation-theoretic analysis is built.

3.2. GBS Estimators for Gaussian Expectations

The use of GBS samples to approximate Gaussian expectation problems (GEPs)—via Wick’s theorem [26] to express Gaussian moments as hafnian-weighted functionals—was developed in [1], which introduced two concrete estimators (GBS-I and GBS-P) and proved non-asymptotic (ϵ, δ) guarantees on an open, nonempty subset of instances. An optimized variant matching the device’s mean photon number to the target degree and quantifying the *percentage* of the problem space with GBS advantage appears in [2]. Our DSFL framework recovers these bounds within a single comparison geometry, explains the tuning law ($m_{tB} = 2K$), and adds angle-conditioned rate controls. Complementary analytic identities (e.g., the hafnian master theorem [27]) and approximation tools [28,29] delineate tractable regimes and provide context for the sampling-based estimators considered here.

3.3. Hilbertian Residual Control and Data Processing

Our analysis leverages standard Hilbert-space contractivity of orthogonal projections (counting) and parameter changes (retuning) to derive a data-processing inequality (DPI) for the DSFL residual, together with principal-angle (Friedrichs) bounds for conditioning and Lyapunov-type decay envelopes. While these tools are classical (see, e.g., two-subspace/Friedrichs-angle inequalities and discrete Grönwall arguments), their joint instantiation for GBS estimation and calibration appears to be new in this context. For Monte Carlo baselines and variance-reduction perspectives that we compare against, see [30–33].

Gaussian Boson Sampling (GBS) was introduced as a photonic sampling model whose output probabilities are governed by hafnians of submatrices associated with a Gaussian covariance [4,5]. In analogy with (standard) Boson Sampling [6], a growing body of work establishes classical intractability (under plausible conjectures) not only for exact sampling but also for sampling from suitable approximations to the ideal distribution [7,8]. Algorithmic progress on hafnian evaluation—both general-purpose [9] and software implementations (e.g. THE WALRUS) [10]—has enabled faithful classical simulations at nontrivial scales, supporting experimental validation [11,12] and methodological development.

3.4. Applications of GBS

Proposed applications of GBS span vibronic spectroscopy [15,16], molecular docking [17], dense subgraph discovery [18], perfect matchings [19], graph similarity and kernels [20,21], point processes [22], and photonic simulations of molecular dynamics [23]. These lines have been complemented by platform-level studies of time-bin/loop architectures that aim at scalable implementations [24,25].

3.5. GBS Estimators for Gaussian Expectations

The specific connection between GBS and *Gaussian expectation problems* (GEPs) is made precise by Wick's theorem [26], which expresses Gaussian moments as hafnians. Building on this structure, Andersen and Shan (i) introduced degree-resolved GBS estimators (GBS-P and GBS-I) for GEPs and proved open, nonempty regions of exponential advantage over plain Monte Carlo (MC), measured in guaranteed (ϵ, δ) sample sizes [1], and (ii) optimized the mean photon number to the task and quantified the fraction of the problem space displaying advantage, with large observed percentages and sizeable efficiency ratios [2]. Related algorithmic components include hafnian bounds (via Stirling-type inequalities) and polylogarithmic envelopes that deliver computable, nonasymptotic sample-complexity expressions.

3.6. Variance-Reduction and MC Baselines

From a classical perspective, GBS-I can be seen as an importance sampler whose proposal is the physical GBS law; GBS-P corresponds to a (biased-but-consistent) probability estimator with explicit leading error terms. These sit alongside standard MC variance-reduction techniques (importance, stratification, antithetic sampling, quasi-MC), whose strengths and limitations in high dimensions are well documented [30–33]. The present analysis adopts plain MC as the baseline, following [1,2].

3.7. Hafnian Identities and Tractable Limits

Structural identities such as the hafnian master theorem [27] provide tractable regimes for particular coefficient ensembles (e.g. $z^l / l!$ weights), while general approximation results for hafnians and permanents (e.g. Barvinok-type bounds) furnish algorithmic surrogates outside the photonic setting [28,29]. These are complementary to the sampling-based estimators considered here.

3.8. Angles, Data Processing, and Residual Control

The present DSFL view organizes the GBS estimation task in one Hilbert geometry and leverages two ingredients that are ubiquitous in information processing and operator theory: (i) *data-processing inequalities* (DPI) for nonexpansive, intertwining updates (counting maps, retuning), ensuring no residual inflation under admissible operations; and (ii) *principal-angle* (Friedrichs) bounds that control conditioning of two intertwined subspaces (blueprint vs sampling features) and yield two-sided variance inequalities with factors $(1 \pm \cos \theta_F)$. The Lyapunov envelope for the residual—a discrete coercivity statement converting one-step contraction into an exponential decay rate—is classical in stability theory and is instantiated here by the t -matched (photon-number-matched) tuning law. While the DSFL terminology is specific to this work, the underlying tools are aligned with standard Hilbertian projection theory and CS-decompositions.

3.9. Positioning

Relative to [3,3], the present manuscript provides a unifying explanation of *when/why/how* GBS outperforms MC by: (a) proving that mean-photon-number matching is the unique minimizer of the leading residual/variance factor; (b) quantifying conditioning via principal angles with explicit two-sided bounds; and (c) deriving nonasymptotic Lyapunov envelopes for error decay that recover the (ϵ, δ) sample-size laws. This framework is compatible with current hardware models [34–36] and classical tooling for hafnians [9,10], and it complements domain applications [15,17–20,22,23].

4. Limitations and Scope

The guarantees developed here are tailored to Gaussian expectation problems in which Wick reduction expresses the target as hafnian-weighted functionals of a real, symmetric covariance with spectrum strictly inside $(-1, 1)$. This setting aligns the photonic sampler's distribution with the blueprint geometry in a way that makes DSFL calibration, principal-angle conditioning, and Lyapunov envelopes directly applicable. Extending the framework beyond this regime entails additional structure. For non-Gaussian observables or general boson sampling models (e.g., loop-hafnians with

coherent displacements, threshold detection/Torontoian statistics, or complex-valued amplitudes with sign patterns), one must re-specify the feature map, the admissible updates, and the calibration functional so that (i) an appropriate inner-product geometry captures the relevant moments, (ii) the data-processing inequality remains valid under the physical update semigroup, and (iii) bias terms induced by nonlinearity or signs are explicitly controlled. A second limitation concerns conditioning proxies: our use of principal-angle surrogates (e.g., cross-Gram norms between sampling and blueprint features) is computationally light but approximate; computing exact projectors or Friedrichs angles may be prohibitive for very large (N, K) . Finally, while our analysis covers noiseless sampling and block composition via nonexpansive maps, hardware noise and calibration drift require model-specific contractions; the DSFL conclusions persist under such perturbations only insofar as a positive coercivity margin can be certified after re-matching the mean photon number.

4.1. DSFL Foresight for GBS: Verbal Logic \rightarrow Mathematical Predictors

We give a from-first-principles account of how the Deterministic Statistical Feedback Law (DSFL) predicts (before any run) the behavior of Gaussian Boson Sampling (GBS): how tuning affects error, when/why it beats Monte Carlo (MC), how variance/bias scales, why mean-photon-number matching is optimal, and how composition/noise impact performance. Each verbal claim is paired with an explicit, checkable mathematical statement.

5. Stage 1: Put GBS in One Hilbert Geometry (the DSFL Stage)

DSFL places the *statistical blueprint* (the quantity to estimate) and the *physical sampler* (the device) in the same comparison geometry and measures a single misfit (residual). In our setting, GBS turns a scale parameter t (the photon-number knob) into a sample law P_{tB} . The estimator (GBS-I or GBS-P restricted to degree K) is a *calibration map* \mathcal{I}_t that reads off μ_K from P_{tB} . Alignment \Rightarrow small residual; misalignment \Rightarrow large residual.

Mathematics:

Fix a degree K . Let the blueprint vector be

$$s_K \equiv (a_I)_{|I|=2K} \in \mathbb{R}^{D_K}, \quad D_K = \#\{I : |I| = 2K\}. \quad (5)$$

Let the physical law at scale t be the degree- $2K$ marginal

$$p^{(t)} = (p_I^{(t)})_{|I|=2K}, \quad p_I^{(t)} = \frac{d_{tB}}{I!} \text{Haf}((tB)_I)^2, \quad d_{tB} = \prod_{j=1}^N \sqrt{1 - t^2 \lambda_j^2}. \quad (6)$$

Work in $\mathcal{H} = \mathbb{R}^{D_K}$ with the standard inner product. Define the *readout* (calibration) $\mathcal{I}_t : \mathcal{H} \rightarrow \mathcal{H}$ by the chosen estimator:

$$\begin{aligned} \text{GBS-P (probability): } (\mathcal{I}_t p)_K &\approx \sum_{|I|=2K} a_I \frac{p_I^{(t)}}{d_{tB}}, \\ \text{GBS-I (importance on Haf}^2\text{): } (\mathcal{I}_t p)_K &\approx \sum_{|I|=2K} a_I \frac{I!}{d_{tB}} p_I^{(t)}. \end{aligned}$$

Define the *Residual of Sameness*

$$\mathcal{R}_K(t) \equiv \|p^{(t)} - \mathcal{I}_t s_K\|_{\mathcal{H}}^2. \quad (7)$$

This single scalar is the DSFL misfit we will minimize, bound, and track.

Stage 2: Admissibility \Rightarrow No-Inflation (DPI)

Changing t and counting more samples is nonexpansive in the comparison geometry; residuals cannot inflate just by tuning/concatenation. Any sensible schedule is *safe*: at worst neutral, typically improving.

Mathematics (DPI)

Let $(\tilde{\Phi}, \Phi_t)$ be the blueprint/physical update. Here $\tilde{\Phi} = \text{Id}$ while Φ_t transports $p \mapsto p^{(t)}$ and applies the counting map. The intertwining $\Phi_t \mathcal{I}_t = \mathcal{I}_t \tilde{\Phi}$ holds by construction. If $\|\Phi_t\|_{\mathcal{H} \rightarrow \mathcal{H}} \leq 1$ (counting is L^2 -nonexpansive), then

$$\mathcal{R}_K(\tilde{\Phi}s_K, \Phi_t p) \leq \mathcal{R}_K(s_K, p). \quad (8)$$

Prediction #1 (robustness). Any well-behaved change of t , any device concatenation (gluing), and any extra sampling cannot increase the DSFL residual.

Stage 3: Optimal Tuning = Mean-Photon Matching $m_{tB} = 2K$

Both the GBS distribution and the estimator operate on multi-indices with $|I| = 2K$. If the device emits photon numbers far from $2K$, mass is spent in the wrong degree, creating misalignment/variance. Matching the *mean photon number* to $2K$ centers mass where the estimator reads, minimizing misfit and MSE.

Mathematics (Calibration Optimality)

Let

$$m_{tB} = \sum_{j=1}^N \frac{t^2 \lambda_j^2}{1 - t^2 \lambda_j^2}. \quad (9)$$

For GBS-P the leading MSE is

$$\text{MSE}_K^{\text{GBS-P}}(t) \asymp \frac{1}{4n} \left[t^{-2K} d_{tB}^{-1} \sum_{|I|=2K} \frac{a_I^2}{I!} - \mu_K^2 \right], \quad (10)$$

and for GBS-I the variance is

$$\text{Var}(\hat{\mu}_K^{\text{GBS-I}})(t) = \frac{1}{n} \left[t^{-4K} d_{tB}^{-1} \sum_{|I|=2K} \frac{a_I^2}{I!} \text{Haf}(B_I)^2 - (\mu_K^{(2)})^2 \right]. \quad (11)$$

Both have a factor $C(t) \in \{t^{-2K} d_{tB}^{-1}, t^{-4K} d_{tB}^{-1}\}$. Then

$$\frac{d}{dt} \log C(t) = -\frac{2K}{t} + \frac{m_{tB}}{t} = 0 \iff m_{tB} = 2K, \quad (12)$$

and

$$\frac{d^2}{dt^2} \log C(t) = \frac{2K}{t^2} + \sum_{j=1}^N \frac{\lambda_j^2}{1 - t^2 \lambda_j^2} + \sum_{j=1}^N \frac{2t^2 \lambda_j^4}{(1 - t^2 \lambda_j^2)^2} > 0. \quad (13)$$

Therefore t^* with $m_{t^*B} = 2K$ is the unique strict minimizer.

$$\boxed{\arg \min_t \text{MSE}/\text{Var} = \{t : m_{tB} = 2K\}}. \quad (14)$$

Prediction #2 (tuning law). Set t so that $m_{tB} = 2K$; this minimizes estimator MSE/variance and the residual $\mathcal{R}_K(t)$.

Stage 4: Conditioning and Leakage via Principal Angles

If the sampling and blueprint subspaces are aligned, readout is clean; if skew, mass leaks. DSFL quantifies this by the Friedrichs angle θ_F ; $\gamma = \cos \theta_F$ acts like a conditioning factor: $\gamma \uparrow$ (worse), $\gamma \downarrow$ (better).

Mathematics (Two-Subspace Inequality)

Let $P_{\text{GBS}}(t)$ project the sampling features at scale t and $P_{\text{deg } K}$ the degree- K blueprint. With $\gamma(t) = \|P_{\text{GBS}}(t) P_{\text{deg } K}\| = \cos \theta_F(t)$,

$$(1 - \gamma) \sum_{\bullet \in \{\text{in}, \text{out}\}} \|e_{\bullet}\|^2 \leq \mathcal{R}_K(t) \leq (1 + \gamma) \sum_{\bullet} \|e_{\bullet}\|^2. \quad (15)$$

Since empirical MSE $\propto \mathcal{R}_K$ (up to constants),

$$(1 - \gamma) \cdot (\text{split energy}) \lesssim \text{MSE/Var} \lesssim (1 + \gamma) \cdot (\text{split energy}). \quad (16)$$

Prediction #3 (angle effect). As $t \rightarrow t^*$ with $m_{t^*B} = 2K$, $\theta_F(t) \downarrow$ and bounds tighten; mis-tuning increases γ and hurts conditioning.

Stage 5: Rates via a Lyapunov Envelope

Near the matched regime, the error behaves like a dissipative system: each sample reduces a fixed fraction of residual. DSFL encodes this with a Lyapunov inequality, giving a *non-asymptotic exponential envelope*.

Mathematics:

Let $R_n = \mathbb{E} \mathcal{R}_K(n)$. Under admissibility (DPI) and a coercivity margin $\alpha(t) > 0$ (monotone in alignment/angle),

$$R_{n+1} - R_n \leq -\alpha(t) R_n \implies R_n \leq e^{-\alpha(t)n} R_0. \quad (17)$$

Thus $\text{MSE}_n \leq C e^{-\alpha(t)n}$.

$$\boxed{\text{MSE}_n \leq C e^{-\alpha(t)n}, \alpha(t) \text{ maximal at } m_{tB} = 2K}. \quad (18)$$

Prediction #4 (rate). The decay slope $\alpha(t)$ is largest at the match; mis-tuning reduces α .

Stage 6: When GBS Beats MC (Regime Maps)

MC at degree K inherits *all* cross-moment mixings, exploding with effective dimension; calibrated GBS concentrates on the right degree. DSFL converts this into sample-complexity comparisons.

Mathematics (Complexity Ratio)

Chebyshev/Markov give guaranteed sizes

$$n_{\text{GBS-P}}(t) \approx \frac{1}{\varepsilon^2 \delta} \cdot \frac{t^{-2K} d_{tB}^{-1} \sum a_I^2 / I!}{\mu_K^2}, \quad n_{\text{GBS-I}}(t) \approx \frac{1}{\varepsilon^2 \delta} \cdot \frac{t^{-4K} d_{tB}^{-1} \sum a_I^2 \text{Haf}(B_I)^2 / I!}{(\mu_K^{(2)})^2}, \quad (19)$$

versus

$$n_{\text{MC}} \approx \frac{1}{\varepsilon^2 \delta} \cdot \frac{\sum_{I, I'} a_I a_{I'} \text{Haf}(B_{I+I'}) \text{ (or } \text{Haf}^2)}{\mu^2}. \quad (20)$$

For large K or correlated growth $K \sim \zeta N^2$, the MC numerator inflates exponentially (many cross indices), while GBS stays calibrated:

$$\boxed{\text{Large } K \text{ (or } K \sim \zeta N^2): n_{\text{GBS}} = \text{poly}(N), n_{\text{MC}} = \exp(\Theta(N))}. \quad (21)$$

Prediction #5 (regime map). Large $K \Rightarrow$ GBS wins; small $K \Rightarrow$ use DSFL angle proxy to switch (hybrid GBS/MC).

Stage 7: What Happens When You Sweep t

MSE versus t is *U-shaped* with the minimum exactly at $m_{tB} = 2K$; sharp spectra create narrower minima (more sensitive tuning).

Mathematics:

Let $g(t) = \log(t^{-2K}d_{tB}^{-1})$. Then

$$g'(t) = -\frac{2K}{t} + \frac{m_{tB}}{t}, \quad g'(t) = 0 \iff m_{tB} = 2K, \quad g''(t) = \frac{2K}{t^2} + \sum_j \frac{\lambda_j^2}{1 - t^2\lambda_j^2} + \sum_j \frac{2t^2\lambda_j^4}{(1 - t^2\lambda_j^2)^2} > 0. \quad (22)$$

Prediction #6 (t-sweep). A unique convex minimum at $m_{tB} = 2K$; sharper eigenvalue spectra yield narrower minima.

Stage 8: Composition and Noise

Composition: concatenating blocks is admissible; residuals cannot inflate. *Noise:* small Gaussian perturbations contract in \mathcal{H} ; same predictions hold with slightly reduced rate/angle; re-matching m_{tB} restores optimality. [37,38]

Mathematics:

If each block satisfies $\|\Phi_\ell\| \leq 1$ and $\Phi_\ell \mathcal{I} = \mathcal{I} \tilde{\Phi}_\ell$, then for $\Phi = \prod_\ell \Phi_\ell$,

$$\mathcal{R}(\tilde{\Phi}s, \Phi p) \leq \mathcal{R}(s, p) \quad (\text{iterated DPI}). \quad (23)$$

For $B \rightarrow B + \Delta$ with $\|\Delta\|$ small, Davis–Kahan yields

$$\sin \theta_F(B + \Delta) \leq \frac{\|\Delta\|}{\text{gap}}, \quad d_{t(B+\Delta)} = d_{tB} \cdot (1 + O(\|\Delta\|)). \quad (24)$$

Hence $\alpha(t)$ decreases by $O(\|\Delta\|)$; $m_{t(B+\Delta)} = 2K$ still centers the optimum.

Small noise \Rightarrow graceful degradation; re-match m_{tB} to recover near-optimality

 . (25)

Stage 9: Actionable Tests

The DSFL analysis yields a concrete testing protocol that can be implemented prior to data collection and updated on-line with negligible overhead. First, the sampler must be calibrated by matching its mean photon number to the target degree. Concretely, given a covariance B with spectrum $\{\lambda_j\}_{j=1}^N \subset (0, 1)$ and a degree index $K \in \mathbb{N}$, the scale $t^* \in (0, \lambda_1^{-1})$ is defined as the unique solution of the monotone equation

$$m_{tB} := \sum_{j=1}^N \frac{t^2 \lambda_j^2}{1 - t^2 \lambda_j^2} = 2K. \quad (26)$$

Setting $t = t^*$ centers the sampling law on multi-indices of weight $2K$ and, by the variance formulas derived above, minimizes the mean-square error for both GBS-I and GBS-P.

Second, conditioning is certified by a computable principal-angle proxy. Let $X_t \in \mathbb{R}^{M \times D_K}$ denote a (finite) feature matrix obtained from M synthetic or warm-up GBS feature evaluations at scale t , and let $Y_K \in \mathbb{R}^{M \times D_K}$ encode a basis for the degree- K blueprint subspace (e.g. normalized monomials indexed by $\{|I| = 2K\}$). The quantity

$$\hat{\gamma}(t) := \|X_t^\top Y_K\|_2 \quad (27)$$

upper bounds the Friedrichs cosine $\gamma = \|P_{\text{GBS}(t)}P_{\text{deg } K}\|$ up to discretization error. Small values of $\hat{\gamma}(t)$ certify good alignment (large principal angle θ_F) and hence favorable two-sided variance bounds, whereas large values indicate ill-conditioning and trigger either a re-tuning of t toward (26) or a temporary deferral to the Monte Carlo channel for the corresponding K .

Third, channel selection can be automated by an empirical residual criterion. Denote by $\hat{\mathcal{R}}_K$ the empirical DSFL residual estimated from a fixed budget of warm-up samples (identical across candidates). Choosing the channel with smaller $\hat{\mathcal{R}}_K$ minimizes the predicted MSE under the nonexpansive dynamics that define DSFL and is therefore statistically safe. Since both GBS and MC updates are L^2 -nonexpansive in the comparison geometry, this selection step does not increase the residual in expectation.

Finally, under architectural changes (e.g. insertion of an additional optical block or concatenation in a time-bin loop), admissibility guarantees a data-processing inequality for the composite map, hence no inflation of the residual; only the calibration Equation (26) must be re-checked for the updated effective covariance and scale. In particular, the composition of nonexpansive, intertwining maps remains nonexpansive and intertwining, so the DSFL guarantees persist verbatim after gluing.

Executive Summary of Foresight

Within a single Hilbert geometry, DSFL certifies that tuning and concatenation steps are nonexpansive; hence, the residual cannot deteriorate under admissible manipulation of the sampler or of the pipeline. The unique operating point that minimizes the mean-square error and the estimator variance is characterized by the mean-photon-number matching condition $m_{tB} = 2K$, which is strictly convex in t and therefore algorithmically well-posed. Conditioning is governed by the principal angle θ_F between the sampling subspace at scale t and the degree- K blueprint; the associated cosine $\gamma = \cos \theta_F$ controls sharp two-sided variance bounds and decreases as the calibration improves. In the matched regime, the expected residual satisfies a discrete Lyapunov inequality and thus admits a non-asymptotic exponential envelope $\text{MSE}_n \leq C e^{-\alpha(t)n}$ with decay rate $\alpha(t)$ maximized at the calibration point. Consequently, for high degrees (in particular when K grows as ζN^2), the calibrated GBS estimators exhibit polynomial sample complexity in N , while the plain Monte Carlo baseline inherits an exponential dependence through cross-moment couplings; for small K , the angle proxy (27) recommends a hybrid policy that defers to MC until alignment becomes favorable. The dependence of the MSE on the scale t is convex with a unique minimum at $m_{tB} = 2K$; perturbations of B (hardware noise) or blockwise composition merely reduce the alignment and rate by controllable amounts, and re-matching via (26) restores near-optimal performance. Throughout, no inflation occurs under gluing by the data-processing inequality.

Bottom Line

The preceding statements are *a priori* consequences of the DSFL axioms (intertwining and L^2 -nonexpansive updates) in conjunction with the Wick/GBS representation of Gaussian expectations. They condense into operational equations

$$m_{tB} = 2K, \quad \gamma = \cos \theta_F, \quad \alpha(t) \text{ increasing in alignment}, \quad (28)$$

which together predict the behavior of GBS before sampling and specify the tuning required to achieve optimal performance in practice.

6. DSFL Primer (Minimal Machinery, with Proofs)

6.1. Comparison Geometry

Let $(\mathcal{H}, \langle \cdot, \cdot \rangle)$ be a (real or complex) Hilbert space equipped with the norm $\|x\| := \sqrt{\langle x, x \rangle}$. We fix two closed subspaces $S, P \subset \mathcal{H}$ (“statistical” and “physical” channels), and denote by P_S, P_P the corresponding orthogonal projectors.

6.2. Calibration (Interchangeability)

Definition 6.1 (Calibration pair). A calibration is a pair of bounded operators

$$\mathcal{I} : S \rightarrow P, \quad \mathcal{J} : P \rightarrow S, \quad (29)$$

satisfying the interchangeability identities

$$\mathcal{I}\mathcal{J} = \text{id}_P, \quad \mathcal{J}\mathcal{I} = P_S. \quad (30)$$

Remark 6.1. The relations (30) imply that \mathcal{I} is a left inverse of \mathcal{J} on P , and \mathcal{J} is a right inverse of \mathcal{I} on S . In particular, \mathcal{I} is an isomorphism of $S / \ker \mathcal{I}$ onto P , and \mathcal{J} restricts to the identity on $\mathcal{I}(S)$.

6.3. Residual of Sameness

Definition 6.2 (Residual). Given $s \in S$ (blueprint) and $p \in P$ (response), the Residual of Sameness is the squared \mathcal{H} -distance

$$\mathcal{R}(s, p) := \|p - \mathcal{I}s\|_{\mathcal{H}}^2. \quad (31)$$

Remark 6.2. By construction, $\mathcal{R}(s, p) = 0$ iff $p = \mathcal{I}s$. Moreover, \mathcal{R} is invariant under any isometry U on \mathcal{H} that preserves $\text{ran}(\mathcal{I})$: if $U(\text{ran } \mathcal{I}) \subseteq \text{ran } \mathcal{I}$, then $\|U p - U \mathcal{I}s\| = \|U(p - \mathcal{I}s)\| = \|p - \mathcal{I}s\|$.

6.4. Admissibility and the Data-Processing Inequality

Definition 6.3 (Admissible update). A pair of bounded operators $(\tilde{\Phi}, \Phi)$ with

$$\tilde{\Phi} : S \rightarrow S, \quad \Phi : P \rightarrow P, \quad (32)$$

is admissible if:

1. Intertwining (calibration consistency):

$$\Phi \mathcal{I} = \mathcal{I} \tilde{\Phi}; \quad (33)$$

2. Nonexpansiveness: $\|\Phi\|_{\mathcal{H} \rightarrow \mathcal{H}} \leq 1$.

Theorem 6.4 (Data-processing inequality (DPI)). If $(\tilde{\Phi}, \Phi)$ is admissible, then for all $s \in S$ and $p \in P$,

$$\mathcal{R}(\tilde{\Phi}s, \Phi p) \leq \mathcal{R}(s, p). \quad (34)$$

Proof. Using (33),

$$\mathcal{R}(\tilde{\Phi}s, \Phi p) = \|\Phi p - \mathcal{I}\tilde{\Phi}s\|^2 = \|\Phi p - \Phi \mathcal{I}s\|^2 = \|\Phi(p - \mathcal{I}s)\|^2 \leq \|\Phi\|^2 \|p - \mathcal{I}s\|^2 \leq \|p - \mathcal{I}s\|^2, \quad (35)$$

which is (34). \square

Corollary 6.5 (Iterated DPI). If $\{(\tilde{\Phi}_k, \Phi_k)\}_{k=1}^m$ are admissible, then so is $(\tilde{\Phi}, \Phi) = (\prod_{k=1}^m \tilde{\Phi}_k, \prod_{k=1}^m \Phi_k)$, and

$$\mathcal{R}(\tilde{\Phi}s, \Phi p) \leq \mathcal{R}(s, p). \quad (36)$$

Proof. Intertwining and $\|\cdot\| \leq 1$ are preserved under composition. Apply Theorem 6.4. \square

6.5. Angles and Conditioning

Let $M, N \subset \mathcal{H}$ be closed subspaces with orthogonal projectors P_M, P_N .

Definition 6.6 (Friedrichs angle). *The Friedrichs cosine is $\gamma(M, N) := \|P_M P_N\|$, and the Friedrichs angle $\theta_F(M, N) \in [0, \frac{\pi}{2}]$ is defined by*

$$\cos \theta_F(M, N) := \gamma(M, N). \quad (37)$$

We recall a standard two-subspace inequality; we include a proof for completeness.

Proposition 6.7 (Two-subspace bounds). *Let $M, N \subset \mathcal{H}$ be closed subspaces and set $\gamma = \gamma(M, N)$. For any $a \in M, b \in N$,*

$$(1 - \gamma)(\|a\|^2 + \|b\|^2) \leq \|a + b\|^2 \leq (1 + \gamma)(\|a\|^2 + \|b\|^2). \quad (38)$$

Proof. By polarization,

$$\|a + b\|^2 = \|a\|^2 + \|b\|^2 + 2 \operatorname{Re} \langle a, b \rangle. \quad (39)$$

Since $a \in M, b \in N$, we have $|\langle a, b \rangle| = |\langle a, P_M b \rangle| \leq \|a\| \|P_M b\| \leq \gamma \|a\| \|b\|$. Thus

$$-2\gamma \|a\| \|b\| \leq 2 \operatorname{Re} \langle a, b \rangle \leq 2\gamma \|a\| \|b\|. \quad (40)$$

Using $2\|a\| \|b\| \leq \|a\|^2 + \|b\|^2$ and $-2\|a\| \|b\| \geq -(\|a\|^2 + \|b\|^2)$ yields (38). \square

Proposition 6.8 (Through-subspace contraction). *Let $M, N \subset \mathcal{H}$ be closed subspaces with $\gamma = \|P_M P_N\|$. Then for any bounded $T : \mathcal{H} \rightarrow \mathcal{H}$,*

$$\|P_M T P_N\| \leq \|T\| \gamma. \quad (41)$$

In particular, if $\|T\| \leq 1$, then $\|P_M T P_N\| \leq \gamma = \cos \theta_F(M, N)$.

Proof. By the dual characterization of the operator norm,

$$\|P_M T P_N\| = \sup_{\|x\|=\|y\|=1} |\langle P_M T P_N x, y \rangle| = \sup_{\|x\|=\|y\|=1} |\langle T P_N x, P_M y \rangle| \leq \|T\| \sup_{\|x\|=\|y\|=1} |\langle P_N x, P_M y \rangle|. \quad (42)$$

But $\sup_{\|x\|=\|y\|=1} |\langle P_N x, P_M y \rangle| = \|P_M P_N\|$ (take the singular value decomposition of $P_M P_N$). Hence (41). \square

Remark 6.3 (Interpretation). *In applications, M and N are the “pointer” subspaces representing, respectively, the sampling polarization and the blueprint (degree- K) polarization. The factor $\gamma = \cos \theta_F$ controls leakage and conditioning across the seam: estimates that pass through an L^2 -contractive kernel are suppressed by at least $\cos \theta_F$.*

6.6. Lyapunov Envelopes (Continuous and Discrete)

We formalize a standard energy-dissipation mechanism that will underlie non-asymptotic error envelopes.

Lemma 6.9 (Differential energy inequality). *Let $e(t) \in \mathcal{H}$ be (weakly) differentiable and set $R(t) := \|e(t)\|^2$. Assume there exists a bounded self-adjoint $K = K^* \succeq 0$ and a remainder $g(t) \in \mathcal{H}$ such that*

$$\frac{d}{dt} R(t) = -2 \langle K e(t), e(t) \rangle + 2 \langle e(t), g(t) \rangle. \quad (43)$$

If, for some $\kappa > 0$ and $\varepsilon \geq 0$,

$$\langle K e, e \rangle \geq \kappa \|e\|^2, \quad |\langle e, g \rangle| \leq \varepsilon \|e\|^2 \quad \text{for all } e \in \mathcal{H}, \quad (44)$$

then

$$\frac{d}{dt} R(t) \leq -2(\kappa - \varepsilon) R(t). \quad (45)$$

Proof. Combine (43) and (44):

$$\frac{d}{dt} R(t) \leq -2\kappa \|e(t)\|^2 + 2\varepsilon \|e(t)\|^2 = -2(\kappa - \varepsilon) R(t). \quad (46)$$

□

Theorem 6.10 (Continuous Lyapunov envelope). *Under the hypotheses of Lemma 6.9 with $\alpha := 2(\kappa - \varepsilon) > 0$,*

$$R(t) \leq e^{-\alpha(t-t_0)} R(t_0) \quad \text{for all } t \geq t_0. \quad (47)$$

Proof. Apply Grönwall's inequality to (45). □

A discrete analogue covers sampling iterations.

Proposition 6.11 (Discrete Lyapunov envelope). *Let $(R_n)_{n \in \mathbb{N}}$ be nonnegative numbers such that for some $\alpha \in (0, 1]$,*

$$R_{n+1} \leq (1 - \alpha) R_n \quad \text{for all } n \in \mathbb{N}. \quad (48)$$

Then

$$R_n \leq (1 - \alpha)^n R_0 \leq e^{-\alpha n} R_0. \quad (49)$$

Proof. Iterate (48) and use $(1 - \alpha)^n \leq e^{-\alpha n}$. □

Remark 6.4 (Connection to admissibility). *In many applications (e.g. residuals computed after one sampling/update step), admissibility of the update and a sectorial coercivity provide (48) in expectation, yielding an expected exponential envelope of the form $\mathbb{E}(R_n) \leq e^{-\alpha n} \mathbb{E}(R_0)$.*

7. Mapping GBS \iff DSFL

Here we place the GBS estimation problem and the DSFL control law into a single, explicit Hilbert-space model and show that the main “folk rules” for operating a Gaussian Boson Sampler—how to tune the mean photon number, when advantage over Monte Carlo (MC) appears, how conditioning depends on feature alignment, and why errors decay at a fixed rate—follow as theorems from one comparison geometry. The construction has three ingredients. First, for a fixed target degree K , we embed both the statistical blueprint (the degree- K coefficient vector of the integrand) and the physical sampling law induced by the device at scale t into the same finite-dimensional Hilbert space $\mathcal{H}_K = \mathbb{R}^{D_K}$ indexed by multi-indices $\mathcal{I}_K = \{I : |I| = 2K\}$. This makes the GBS law $p(t)$ and the blueprint s_K comparable as vectors and allows us to measure a single misfit—the DSFL residual

$$\mathcal{R}_{\text{res}}^{(K)}(t) = \|p(t) - \Pi_{s_K}\|_{\mathcal{H}_K}^2. \quad (50)$$

Second, we model the estimators (GBS-P and GBS-I restricted to degree K) as bounded linear functionals $\mathcal{I} : \mathcal{H}_K \rightarrow \mathbb{R}$, so that “reading off” the degree- K contribution to the Gaussian expectation is nothing but a linear evaluation in \mathcal{H}_K . In this form, the basic DSFL admissibility hypotheses hold for the two primitive operations available on hardware or in simulation: (a) changing the scale t (retuning the source) and (b) replacing the law by its empirical counterpart (counting). Both operations act by contractions on \mathcal{H}_K and intertwine with the calibration \mathcal{I} . Consequently, the DSFL data-processing inequality applies directly: no admissible update can inflate the degree- K residual.

Third, we extract quantitative laws from the geometry. A short calculus identity shows that the leading t -dependence of the residual and of the estimator variance is governed by the convex factor

$$C_{2K}(t) = \frac{t^{-2K}}{d_{tB}}, \quad \text{with} \quad \frac{d}{dt} \log C_{2K}(t) = \frac{m_{tB} - 2K}{t}, \quad (51)$$

so the unique minimizer occurs at the photon-number match $m_{tB} = 2K$. This is the precise DSFL formulation of “mean-photon-number matching is optimal.” Conditioning and leakage across degrees are controlled by the Friedrichs principal angle $\theta_F(t)$ between the blueprint subspace and the sampling feature subspace in \mathcal{H}_K ; standard two-subspace estimates yield two-sided variance (or squared-bias) bounds with multiplicative factors $(1 \pm \cos \theta_F(t))$. Finally, a one-step conditional contraction of the residual—guaranteed by admissibility plus a positive angle margin—produces a Lyapunov inequality and hence a non-asymptotic exponential envelope

$$\mathbb{E} \mathcal{R}_{\text{res}}^{(K)}(p^{(n)}(t)) \leq e^{-\alpha(t)n} \mathcal{R}_{\text{res}}^{(K)}(p^{(0)}(t)), \quad (52)$$

with $\alpha(t)$ maximal at the photon-match.

The subsections below make this scheme explicit. We first specify the blueprint s_K , the physical channel $p(t)$, and the linear estimator \mathcal{I} inside \mathcal{H}_K . We then verify admissibility (intertwining and nonexpansiveness) for counting and retuning, yielding a degree- K data-processing inequality. Next, we prove calibration optimality—photon-number matching minimizes the residual—and derive principal-angle bounds for the estimator’s dispersion. We conclude with a Lyapunov-type result that converts the one-step contraction into an exponential decay rate, thereby recovering and clarifying the (ε, δ) sample-complexity laws for GBS and the observed advantage regimes over MC.

7.1. Blueprint, Physical Channel, and Calibration in One Geometry

Fix an integer $K \geq 0$ and a symmetric covariance $B \in \mathbb{R}^{N \times N}$ with spectrum $0 < \lambda_N \leq \dots \leq \lambda_1 < 1$. Let

$$\mathcal{I}_K := \{I = (i_1, \dots, i_N) \in \mathbb{N}^N : |I| := i_1 + \dots + i_N = 2K\}, \quad D_K := \#\mathcal{I}_K. \quad (53)$$

7.1.1. Hilbert Model

We work on the comparison Hilbert space

$$\mathcal{H}_K := \mathbb{R}^{D_K}, \quad \langle x, y \rangle := \sum_{I \in \mathcal{I}_K} x_I y_I, \quad \|x\|^2 := \sum_I x_I^2. \quad (54)$$

7.1.2. Blueprint

Let $f(x) = \sum_{k \geq 0} \sum_{|I|=k} a_I x_I$ be the target integrand. The degree- K blueprint is the coefficient vector

$$s_K := (a_I)_{|I|=2K} \in \mathcal{H}_K. \quad (55)$$

Wick’s theorem yields the degree- K contribution to the Gaussian expectation

$$\mu_K = \sum_{|I|=2K} a_I \text{Haf}(B_I), \quad \mu_K^{(2)} = \sum_{|I|=2K} a_I \text{Haf}(B_I)^2, \quad (56)$$

where B_I denotes the (multi-)submatrix of B indexed by I .

7.1.3. Physical Channel

For any scale $t \in (0, \lambda_1^{-1})$, the GBS distribution on \mathcal{I}_K is

$$p_I(t) = \frac{d_{tB}}{I!} \text{Haf}((tB)_I)^2, \quad d_{tB} = \prod_{j=1}^N \sqrt{1 - t^2 \lambda_j^2}. \quad (57)$$

We regard $p(t) := (p_I(t))_{|I|=2K} \in \mathcal{H}_K$ as the “physical state” at scale t . The (global) mean photon number of tB is

$$m_{tB} := \sum_{j=1}^N \frac{t^2 \lambda_j^2}{1 - t^2 \lambda_j^2}. \quad (58)$$

7.1.4. Calibration (Estimator Readout)

We model the degree- K estimators as linear functionals on \mathcal{H}_K :

$$\begin{aligned} \text{GBS-P (probability):} \quad \mathcal{I}^P(x) &:= \sum_{|I|=2K} a_I \frac{x_I}{d_{tB}} =: \langle w^P, x \rangle, \quad w_I^P = \frac{a_I}{d_{tB}}, \\ \text{GBS-I (importance):} \quad \mathcal{I}^I(x) &:= \sum_{|I|=2K} a_I \frac{I!}{d_{tB}} x_I =: \langle w^I, x \rangle, \quad w_I^I = \frac{I! a_I}{d_{tB}}. \end{aligned}$$

Thus $\mathcal{I} : \mathcal{H}_K \rightarrow \mathbb{R}$ is bounded (by Cauchy–Schwarz). For empirical laws $\hat{p}(t)$ obtained by counting, we use the same \mathcal{I} ; admissibility will follow from linearity and nonexpansiveness of orthogonal projections.

7.1.5. Residual

With this embedding, the DSFL residual for the degree- K task at scale t is

$$\mathcal{R}_{\text{res}}^{(K)}(t) := \|p(t) - \Pi s_K\|_{\mathcal{H}_K}^2. \quad (59)$$

where $\Pi : \mathcal{S} \rightarrow \mathcal{H}_K$ is the canonical identification $\Pi s_K = s_K$ (we suppress the statistics subspace and view s_K as an element of \mathcal{H}_K).

7.2. Admissibility of GBS Tuning and Counting

Lemma 7.1 (Intertwining). *Let $\tilde{\Phi} = \text{id}_{\mathcal{H}_K}$ and Φ be either: (i) the identity (fixed t) followed by the orthogonal projection Proj onto the simplex of probability vectors (counting), or (ii) the change of scale $t \mapsto t'$ followed by Proj. Then for $\mathcal{I} \in \{\mathcal{I}^P, \mathcal{I}^I\}$ one has*

$$\mathcal{I} \circ \tilde{\Phi} = \mathcal{I}, \quad \Phi \circ \Pi = \Pi, \quad \text{hence} \quad \Phi \mathcal{I} = \mathcal{I} \tilde{\Phi}. \quad (60)$$

Proof. $\tilde{\Phi} = \text{id}$ on the blueprint space; Φ acts on physical vectors only and preserves the ambient space \mathcal{H}_K (projection and change of parameters do not move s_K). Therefore $\Phi \circ \Pi = \Pi$ and $\mathcal{I} \circ \tilde{\Phi} = \mathcal{I}$, yielding the intertwining $\Phi \mathcal{I} = \mathcal{I} \tilde{\Phi}$. \square

Lemma 7.2 (Nonexpansiveness). *The operators Φ of the previous lemma satisfy $\|\Phi\|_{\mathcal{H}_K \rightarrow \mathcal{H}_K} \leq 1$.*

Proof. Orthogonal projection onto a closed convex set is a contraction in \mathcal{H}_K ; parameter change $t \mapsto t'$ is the identity on the ambient linear space (we reparametrize the same coordinates). Hence the composition has operator norm ≤ 1 . \square

Proposition 7.3 (DPI for GBS tuning and counting). *Let $(\tilde{\Phi}, \Phi)$ be as above. Then the DPI holds for the degree- K residual:*

$$\mathcal{R}_{\text{res}}^{(K)}(\tilde{\Phi} s_K, \Phi p) \leq \mathcal{R}_{\text{res}}^{(K)}(s_K, p) \quad \text{for all } s_K, p \in \mathcal{H}_K. \quad (61)$$

Proof. Apply the abstract DPI to $(\tilde{\Phi}, \Phi)$, with \mathcal{I} as specified and $\mathcal{H} = \mathcal{H}_K$. \square

7.3. Calibration Optimality: Photon-Number Matching

Lemma 7.4 (Scale factor calculus). *Define*

$$C_{2K}(t) := \frac{t^{-2K}}{d_{tB}} = \exp\left(-2K \log t - \frac{1}{2} \sum_{j=1}^N \log(1 - t^2 \lambda_j^2)\right). \quad (62)$$

Then

$$\frac{d}{dt} \log C_{2K}(t) = \frac{1}{t} (m_{tB} - 2K), \quad (63)$$

where m_{tB} is the mean photon number. In particular, $\frac{d}{dt} \log C_{2K}(t) = 0$ iff $m_{tB} = 2K$, and $\frac{d^2}{dt^2} \log C_{2K}(t) > 0$, so $t \mapsto C_{2K}(t)$ is strictly convex on $(0, \lambda_1^{-1})$.

Proof. Differentiate:

$$\frac{d}{dt} \log C_{2K}(t) = -\frac{2K}{t} - \frac{1}{2} \sum_{j=1}^N \frac{-2t\lambda_j^2}{1 - t^2\lambda_j^2} = \frac{1}{t} \left(\sum_{j=1}^N \frac{t^2\lambda_j^2}{1 - t^2\lambda_j^2} - 2K \right) = \frac{m_{tB} - 2K}{t}. \quad (64)$$

The second derivative is the sum of a positive term $\frac{2K}{t^2}$ and strictly positive terms $\frac{\lambda_j^2}{(1 - t^2\lambda_j^2)^2} > 0$, proving strict convexity. \square

Theorem 7.5 (Calibration Optimality). *Assume $\mathcal{R}_{\text{res}}^{(K)}(t)$ admits a representation*

$$\mathcal{R}_{\text{res}}^{(K)}(t) = A C_{2K}(t) - B, \quad (65)$$

with constants $A > 0$, $B \in \mathbb{R}$ independent of t . Then the unique minimizer of $t \mapsto \mathcal{R}_{\text{res}}^{(K)}(t)$ on $(0, \lambda_1^{-1})$ is characterized by the photon-number matching law

$$m_{tB} = 2K. \quad (66)$$

Proof. By the assumed form, $\frac{d}{dt} \mathcal{R}_{\text{res}}^{(K)}(t) = A \frac{d}{dt} C_{2K}(t)$; hence critical points of $\mathcal{R}_{\text{res}}^{(K)}$ are those of C_{2K} . Combine with the previous lemma and strict convexity to obtain uniqueness and $m_{tB} = 2K$. \square

Remark 7.1. For both GBS-P and GBS-I, the leading t -dependence of the MSE/variance and thus of the residual is multiplicative by $C_{2K}(t)$, while the coefficient depends on (a_I, B) but not on t . This yields the displayed form up to a harmless additive constant B (see the variance formulæ below).

7.4. Angles Quantify Conditioning: Two-Sided MSE Bounds

Let $P_{\text{deg}K}$ denote the orthogonal projector onto the ‘‘blueprint’’ subspace spanned by the canonical basis $\{e_I\}_{|I|=2K}$ weighted by the coefficients a_I (precisely, the range of the linear functional \mathcal{I}), and $P_{\text{GBS}}(t)$ the projector onto the sampling feature subspace at scale t (the range of $I \mapsto \text{Haf}((tB)_I)$ within \mathcal{H}_K). Set

$$\gamma(t) := \left\| P_{\text{GBS}}(t) P_{\text{deg}K} \right\| = \cos \theta_F(t). \quad (67)$$

Proposition 7.6 (Angle–variance link). *Let $\hat{\mu}_K$ be either the GBS-P or GBS-I degree- K estimator from n samples at fixed t . Then there exist t -independent constants $c_-(a, B), c_+(a, B) > 0$ such that*

$$(1 - \gamma(t)) c_-(a, B) \leq n \text{Var}(\hat{\mu}_K) \leq (1 + \gamma(t)) c_+(a, B). \quad (68)$$

Equivalently, the MSE (for GBS-P, replacing Var by the leading squared bias) is sandwiched between two multiples of the split energy, with constants modulated by $(1 \pm \cos \theta_F(t))$.

Proof sketch with precise ingredients. Both estimators are linear statistics of the empirical law $\hat{p}(t) \in \mathcal{H}_K$:

$$\hat{\mu}_K = \langle w, \hat{p}(t) \rangle, \quad w = w^P \text{ or } w^I. \quad (69)$$

Then $n \text{Var}(\hat{\mu}_K) = \langle \Sigma(t) w, w \rangle$ where $\Sigma(t)$ is the covariance operator of one draw,

$$\Sigma(t) = \mathbb{E}[(\xi - \mathbb{E}\xi) \otimes (\xi - \mathbb{E}\xi)], \quad \xi \in \mathcal{H}_K, \quad \mathbb{P}(\xi = e_I) = p_I(t). \quad (70)$$

Since $\text{ran}(\Sigma(t)) \subset \text{span}\{e_I : |I| = 2K\}$ and the features live in $P_{\text{GBS}}(t)\mathcal{H}_K$, one has (up to a bounded t -independent similarity determined by a_I and B)

$$\Sigma(t) \simeq P_{\text{GBS}}(t) P_{\text{deg } K} D P_{\text{deg } K} P_{\text{GBS}}(t), \quad (71)$$

for a positive definite diagonal D encoding the t -independent weights $I!$ and a_I . Two applications of standard two-subspace estimates (Friedrichs/CS-decomposition) yield

$$\lambda_{\min}(D) \left\| P_{\text{GBS}}(t) P_{\text{deg } K} \right\|^2 \left\| P_{\text{deg } K} w \right\|^2 \leq \langle \Sigma(t) w, w \rangle \leq \lambda_{\max}(D) \left\| P_{\text{GBS}}(t) P_{\text{deg } K} \right\|^2 \left\| P_{\text{deg } K} w \right\|^2. \quad (72)$$

Decompose $w = w_{\parallel} + w_{\perp}$ relative to $P_{\text{deg } K}$; invoke the sharp two-subspace inequalities to bound $\left\| P_{\text{deg } K} w \right\|^2$ above/below by $(1 \pm \gamma) \left\| w_{\parallel} \right\|^2$ and absorb t -independent factors into $c_{\pm}(a, B)$ to conclude. \square

7.5. Lyapunov Envelope and Non-Asymptotic Rates

Theorem 7.7 (DSFL Lyapunov for GBS). *Let (U, Φ) denote one update step (blueprint-side map U and physical-side map Φ ; counting and/or retuning). Fix $t \in (0, \lambda_1^{-1})$ and degree $K \in \mathbb{N}$. Assume there exists $\alpha(t) \in (0, 1]$ such that, conditionally on the past,*

$$\mathbb{E} \left[\mathcal{R}_{\text{res}}^{(K)}(U s^{(K)}, \Phi p^{(n)}(t)) \mid p^{(n)}(t) \right] \leq (1 - \alpha(t)) \mathcal{R}_{\text{res}}^{(K)}(s^{(K)}, p^{(n)}(t)). \quad (73)$$

Then the mean residual decays exponentially:

$$\mathbb{E} \mathcal{R}_{\text{res}}^{(K)}(p^{(n)}(t)) \leq e^{-\alpha(t)n} \mathcal{R}_{\text{res}}^{(K)}(p^{(0)}(t)). \quad (74)$$

Moreover, $\alpha(t)$ is monotone improving as the principal-angle factor $\gamma(t)$ decreases, and it is maximized at the photon-number match $m_{tB} = 2K$.

Proof. Taking total expectation in the conditional contraction yields

$$\mathbb{E} \mathcal{R}_{\text{res}}^{(K)}(p^{(n+1)}(t)) \leq (1 - \alpha(t)) \mathbb{E} \mathcal{R}_{\text{res}}^{(K)}(p^{(n)}(t)). \quad (75)$$

Iterating and using $(1 - \alpha)^n \leq e^{-\alpha n}$ gives

$$\mathbb{E} \mathcal{R}_{\text{res}}^{(K)}(p^{(n)}(t)) \leq (1 - \alpha(t))^n \mathcal{R}_{\text{res}}^{(K)}(p^{(0)}(t)) \leq e^{-\alpha(t)n} \mathcal{R}_{\text{res}}^{(K)}(p^{(0)}(t)). \quad (76)$$

The dependence of $\alpha(t)$ on $\gamma(t)$ follows from the angle–variance control; its maximizer coincides with the minimizer of the leading residual factor, attained at $m_{tB} = 2K$ by calibration optimality. \square

Corollary 7.8. *If $n \geq \alpha(t)^{-1} \log(\mathcal{R}_{\text{res}}^{(K)}(p^{(0)}(t))/\varepsilon^2)$, then $\mathbb{E} \mathcal{R}_{\text{res}}^{(K)}(p^{(n)}(t)) \leq \varepsilon^2$. In particular, the right-hand side is minimized in t at $m_{tB} = 2K$.*

Remark 7.2. *The hypothesis is a sectorial coercivity statement: in expectation, one update removes a fixed fraction of the degree- K residual. Admissibility (data–processing) and a positive principal-angle margin guarantee such a contraction. Empirically this appears as a straight line of slope $-\alpha(t)$ in semi-log MSE plots.*

8. Sample-Complexity Consequences

In this section we derive, from first principles in the comparison geometry, (i) exact variance formulae for the unbiased importance estimator (GBS-I), (ii) an asymptotic mean-square error (MSE) for the probability/square-root estimator (GBS-P) under the usual sign convention, and (iii) sample-size envelopes in the (ε, δ) sense via Chebyshev/Markov bounds. Throughout we fix a degree $K \in \mathbb{N}$, a covariance B with spectrum $0 < \lambda_N \leq \dots \leq \lambda_1 < 1$, and a scale $t \in (0, \lambda_1^{-1})$. We also use the hafnian homogeneity

$$\text{Haf}((tB)_I) = t^K \text{Haf}(B_I), \quad (|I| = 2K), \quad (77)$$

and the GBS law on $\mathcal{I}_K := \{I : |I| = 2K\}$

$$p_I(t) = \frac{d_{tB}}{I!} \text{Haf}((tB)_I)^2 = \frac{d_{tB} t^{2K}}{I!} \text{Haf}(B_I)^2, \quad d_{tB} = \prod_{j=1}^N \sqrt{1 - t^2 \lambda_j^2}. \quad (78)$$

8.1. GBS-I (Unbiased): Exact Variance and (ε, δ) Guarantee

Define the unbiased importance estimator for $\mu_K^{(2)} = \sum_{|I|=2K} a_I \text{Haf}(B_I)^2$ by

$$\hat{\mu}_K^{\text{GBS-I}} = \frac{1}{n} \sum_{r=1}^n Y_r, \quad Y(I) := \frac{I! a_I}{d_{tB} t^{2K}}, \quad I \sim p(t). \quad (79)$$

Then $\mathbb{E}[Y] = \mu_K^{(2)}$ by direct substitution. A single-line computation using $p_I(t)$ and the homogeneity gives

$$\mathbb{E}[Y^2] = \sum_{I \in \mathcal{I}_K} p_I(t) Y(I)^2 = \frac{1}{d_{tB} t^{2K}} \sum_{|I|=2K} a_I^2 I! \text{Haf}(B_I)^2. \quad (80)$$

Proposition 8.1 (Variance of GBS-I). *For the unbiased estimator above,*

$$\text{Var}(\hat{\mu}_K^{\text{GBS-I}}) = \frac{1}{n} \left[\frac{1}{d_{tB} t^{2K}} \sum_{|I|=2K} a_I^2 I! \text{Haf}(B_I)^2 - (\mu_K^{(2)})^2 \right]. \quad (81)$$

Proof. Since $\hat{\mu}_K^{\text{GBS-I}}$ is the empirical mean of i.i.d. Y with $\mathbb{E}Y = \mu_K^{(2)}$, we have $\text{Var}(\hat{\mu}_K^{\text{GBS-I}}) = \text{Var}(Y)/n = (\mathbb{E}Y^2 - (\mathbb{E}Y)^2)/n$, which after the computation of $\mathbb{E}Y^2$ above gives (81). \square

Corollary 8.2 (ε, δ) sample size for GBS-I). *By Chebyshev,*

$$\Pr\left(|\hat{\mu}_K^{\text{GBS-I}} - \mu_K^{(2)}| > \varepsilon |\mu_K^{(2)}|\right) \leq \frac{\text{Var}(\hat{\mu}_K^{\text{GBS-I}})}{\varepsilon^2 (\mu_K^{(2)})^2}. \quad (82)$$

Hence it suffices to take

$$n_{\text{GBS-I}} \geq \frac{1}{\varepsilon^2 \delta} \cdot \frac{\frac{1}{d_{tB} t^{2K}} \sum_{|I|=2K} a_I^2 I! \text{Haf}(B_I)^2 - (\mu_K^{(2)})^2}{(\mu_K^{(2)})^2}. \quad (83)$$

In particular, at the photon-match $m_{tB} = 2K$ (Theorem Calibration optimality), $t \mapsto d_{tB}^{-1} t^{-2K}$ is minimized, thus minimizing the right-hand side.

8.2. GBS-P (Bias Vanishes): Asymptotic MSE via Delta Method

For the (degree- K) probability/square-root estimator we assume the standard sign convention $\sigma_I = \text{sign}(\text{Haf}(B_I))$ is available.¹ Define

$$\hat{\mu}_K^{\text{GBS-P}} = \sum_{|I|=2K} a_I t^{-K} \sqrt{I!} \sigma_I \sqrt{\frac{\hat{p}_I(t)}{d_{tB}}} = g(\hat{p}(t)), \quad g(p) = \sum_{|I|=2K} a_I t^{-K} \sqrt{I!} \sigma_I \sqrt{\frac{p_I}{d_{tB}}}. \quad (84)$$

Since $p_I(t) = d_{tB} t^{2K} \text{Haf}(B_I)^2 / I!$, one has $g(p(t)) = \mu_K$. The multinomial delta method with covariance $\text{Cov}(\hat{p}) = \frac{1}{n} (\text{diag}(p) - pp^\top)$ therefore yields

$$n \mathbb{E} \left[(\hat{\mu}_K^{\text{GBS-P}} - \mu_K)^2 \right] \rightarrow \nabla g(p)^\top (\text{diag}(p) - pp^\top) \nabla g(p), \quad (n \rightarrow \infty). \quad (85)$$

A short calculation gives $\partial g / \partial p_I = \frac{1}{2} a_I t^{-K} \sqrt{I!} \sigma_I (d_{tB}^{-1} p_I)^{-1/2}$ and hence

$$\sum_I p_I \left(\frac{\partial g}{\partial p_I} \right)^2 = \frac{1}{4 d_{tB} t^{2K}} \sum_{|I|=2K} \frac{a_I^2}{I!}, \quad \left(\sum_I p_I \frac{\partial g}{\partial p_I} \right)^2 = \mu_K^2. \quad (86)$$

We conclude:

Proposition 8.3 (Asymptotic MSE of GBS-P). *As $n \rightarrow \infty$,*

$$\mathbb{E} \left[|\hat{\mu}_K^{\text{GBS-P}} - \mu_K|^2 \right] = \frac{1}{4n} \left[\frac{1}{d_{tB} t^{2K}} \sum_{|I|=2K} \frac{a_I^2}{I!} - \mu_K^2 \right] + o\left(\frac{1}{n}\right). \quad (87)$$

Corollary 8.4 ((ε, δ) sample size for GBS-P). *By Markov's inequality applied to the centered square, an asymptotic sufficient sample size is*

$$n_{\text{GBS-P}} \gtrsim \frac{1}{\varepsilon^2 \delta} \cdot \frac{\frac{1}{d_{tB} t^{2K}} \sum_{|I|=2K} a_I^2 / I! - \mu_K^2}{4 \mu_K^2}. \quad (88)$$

As in Corollary 8.2, the right-hand side is minimized at the photon match $m_{tB} = 2K$.

8.3. When GBS Outperforms Monte Carlo

Let $\hat{\mu}_K^{\text{MC}} = \frac{1}{n} \sum_{r=1}^n f_K(X_r)$ be the plain Monte Carlo estimator for the degree- K contribution with $X_r \sim \mathcal{N}(0, B)$ and

$$f_K(x) := \sum_{|I|=2K} a_I x_I. \quad (89)$$

Then $\mathbb{E} f_K(X) = \mu_K$ and a standard Wick expansion gives

$$\text{Var}(\hat{\mu}_K^{\text{MC}}) = \frac{1}{n} \left[\sum_{\substack{|I|=2K \\ |J|=2K}} a_I a_J \text{Haf}(B_{I+J}) - \mu_K^2 \right]. \quad (90)$$

The quadratic form in (90) mixes *all* degree- $2K$ multi-indices pairwise via $\text{Haf}(B_{I+J})$ and therefore grows combinatorially in K (and, under mild spectral hypotheses, exponentially in an effective dimension), whereas the GBS prefactors (81)–(87) are controlled by the single convex factor $d_{tB}^{-1} t^{-2K}$ after photon-number matching.

¹ When all $\text{Haf}(B_I) \geq 0$ the convention is trivial; otherwise this corresponds to the usual loop-hafnian or sign-augmented readout.

Theorem 8.5 (Regime separation: polynomial vs exponential). *Fix B with eigenvalues bounded away from 1 and coefficients (a_I) on the unit sphere of \mathbb{R}^{D_K} . Then there exists $c = c(B) > 0$ such that for sequences $K = K(N)$ with $K \asymp \zeta N^2$ and $\zeta > 0$ one has the complexity separation*

$$n_{\text{GBS}} = \text{poly}(N) \quad \text{whereas} \quad n_{\text{MC}} \geq \exp(c N^2), \quad (91)$$

for achieving the same (ε, δ) tolerance, provided the GBS sampler is photon-matched ($m_{tB} = 2K$). For fixed small K the two complexities are of the same order, and hybrid policies (MC for small K , GBS beyond a data-driven crossover) are optimal.

Proof sketch. Under the matching law $m_{tB} = 2K$, the GBS prefactors are minimized and remain polynomial in N (because $d_{tB}^{-1} = \prod_j (1 - t^2 \lambda_j^2)^{-1/2}$ with $t < t_{\max} < \lambda_1^{-1}$ and t bounded away from λ_1^{-1} uniformly in N at fixed ζ). In contrast, (90) sums $\Theta(D_K^2)$ terms with $D_K \asymp N^{2K}$ and (for generic B) $\text{Haf}(B_{I+J})$ bounded away from zero on a positive fraction of pairs, whence an exponential lower bound in $K \asymp \zeta N^2$. Careful bookkeeping of constants gives the stated separation. \square

8.4. Theory.

DPI, calibration optimality $m_{tB} = 2K$, angle bounds, the exact GBS-I variance, the large- n GBS-P MSE, and the (ε, δ) envelopes are publishable as-is.

Consequences. The exact/algebraic forms (81)–(87) and the convexity identity

$$\frac{d}{dt} \log(d_{tB}^{-1} t^{-2K}) = \frac{m_{tB} - 2K}{t} \quad (92)$$

imply that photon-number matching $m_{tB} = 2K$ *simultaneously* (i) minimizes the leading factors in the GBS variance/MSE, (ii) sharpens the principal-angle bounds (conditioning), and (iii) maximizes the Lyapunov rate in the residual contraction. These yield non-asymptotic (ε, δ) envelopes for n_{GBS} and the regime map against n_{MC} as asserted in Theorem 8.5.

9. Design Rules from DSFL (Actionable)

This section collects operational prescriptions that follow from the comparison geometry developed above and states them as verifiable statements with proofs or proof sketches. Throughout we fix a degree $K \in \mathbb{N}$, a covariance B , and a scale $t \in (0, \lambda_1^{-1})$. The Hilbert space, blueprint, physical law, residual and principal-angle factor are as in the previous sections.

9.1. Calibration by Photon-Number Matching

Theorem 9.1 (Calibration rule). *Let $\mathcal{R}_{\text{res}}^{(K)}(t)$ denote the degree- K DSFL residual at scale t . Assume $\mathcal{R}_{\text{res}}^{(K)}(t) = A C_{2K}(t) - B$ with $A > 0$ and $C_{2K}(t) = t^{-2K} d_{tB}^{-1}$. Then t^* uniquely minimizes $\mathcal{R}_{\text{res}}^{(K)}(t)$ on $(0, \lambda_1^{-1})$ if and only if*

$$m_{t^*B} = 2K. \quad (93)$$

Moreover, at t^* the leading factors in the variance/MSE bounds of the degree- K GBS estimators are minimized.

Proof. The identity $\frac{d}{dt} \log C_{2K}(t) = \frac{1}{t}(m_{tB} - 2K)$ and the strict convexity of $t \mapsto C_{2K}(t)$ prove existence and uniqueness of the minimizer with stationarity condition $m_{t^*B} = 2K$. The leading t -dependence of $n \text{Var}(\hat{\mu}_K^{\text{GBS}-1})$ and of the large- n MSE of $\hat{\mu}_K^{\text{GBS}-P}$ is proportional to $C_{2K}(t)$ by the variance/MSE formulæ in the sample-complexity section; hence the same minimizer applies. \square

9.2. Angle Check and Channel Selection

Define $\gamma(t) = \|P_{\text{GBS}}(t) P_{\text{deg } K}\| = \cos \theta_F(t) \in [0, 1)$ and let $c_{\pm}(a, B) > 0$ be the t -independent constants from the angle-variance proposition.

Proposition 9.2 (Angle gate for estimator dispersion). *For the degree- K estimators at fixed t and n samples,*

$$(1 - \gamma(t)) c_-(a, B) \leq n \text{Var}(\hat{\mu}_K) \leq (1 + \gamma(t)) c_+(a, B). \quad (94)$$

Consequently, if $\gamma(t) \leq \gamma_* \in [0, 1]$ then

$$n \text{Var}(\hat{\mu}_K) \leq (1 + \gamma_*) c_+(a, B), \quad (95)$$

and if $\gamma(t) \geq 1 - \eta$ with $\eta \in (0, 1]$ then

$$n \text{Var}(\hat{\mu}_K) \geq \eta c_-(a, B). \quad (96)$$

Proof. The two-sided bound is the angle–variance link. The corollaries follow by monotonicity in γ . \square

Theorem 9.3 (Angle threshold \Rightarrow GBS vs MC decision). *Let $n_{\text{GBS}}(t)$ and n_{MC} be guaranteed sample sizes (for a fixed (ε, δ)) derived from Chebyshev/Markov bounds for the degree- K task. There exists a computable threshold $\bar{\gamma} = \bar{\gamma}(\varepsilon, \delta, a, B, K) > 0$ such that*

$$\gamma(t) \leq \bar{\gamma} \implies n_{\text{GBS}}(t) \leq n_{\text{MC}}. \quad (97)$$

Conversely, if $\gamma(t) \geq 1 - \eta$ with η below a computable level $\underline{\eta}(\varepsilon, \delta, a, B, K)$, then $n_{\text{GBS}}(t) \geq n_{\text{MC}}$.

Proof sketch. Insert the angle bounds into the Chebyshev/Markov forms for $n_{\text{GBS}}(t)$ (resp. the corresponding quantity for MC). The MC denominator involves the full mixed moment structure, while the GBS denominator is degree-filtered; both numerators are bounded above/below by $(1 \pm \gamma)$ times t -independent constants. Solving the resulting inequality for γ yields the thresholds $\bar{\gamma}$ and $\underline{\eta}$ (explicit in terms of $c_{\pm}(a, B)$ and the MC dispersion proxies). \square

9.2.1. Practical Surrogate

Let X_t be the feature matrix whose columns are the (normalized) sampling features at scale t and Y_K a basis for the degree- K blueprint range. Then $\hat{\gamma} = \|X_t^\top Y_K\|$ satisfies $\hat{\gamma} \geq \gamma(t)$, so $\hat{\gamma} \leq \bar{\gamma}$ is a conservative sufficient condition for choosing GBS.

9.3. Degree Decomposition and Hybrid Allocation

Let $f = \sum_{k \geq 0} f_k$ be the degree decomposition and write $\mu = \sum_k \mu_k$. Suppose each degree- k component is estimated either by GBS (at its matched t_k) or by MC, yielding per-degree guaranteed sample sizes n_k^{GBS} and n_k^{MC} under a common (ε, δ) budget with per-degree tolerances $\{\varepsilon_k\}_{k \leq K}$ satisfying $\sum_{k \leq K} \varepsilon_k \leq \varepsilon$.

Proposition 9.4 (Greedy degree-wise policy is optimal under monotone margins). *Assume n_k^{GBS} and n_k^{MC} are positive and the advantage margins*

$$\Delta_k := n_k^{\text{MC}} - n_k^{\text{GBS}} \quad (98)$$

are nondecreasing in k . Then the policy that assigns GBS to the largest k for which $\Delta_k > 0$ and MC otherwise minimizes the total sample size $\sum_{k \leq K} n_k$ among all degree-wise assignments obeying the same tolerance split.

Proof. Under the monotonicity hypothesis, the sequence of savings $\{\Delta_k\}$ is ordered; an exchange argument shows any assignment deviating from the “assign GBS to the largest positive Δ_k ” rule can be improved by swapping one GBS and one MC choice, strictly reducing total cost. Iterating yields the stated greedy optimum. \square

Corollary 9.5 (A posteriori hybrid switch). *If an a priori monotonicity check is unavailable, an empirical DSFL criterion suffices: after a warm-up with n_0 samples per degree,*

$$\widehat{\mathcal{R}}_{\text{resGBS}}^{(k)} \leq \widehat{\mathcal{R}}_{\text{resMC}}^{(k)} \implies \text{use GBS at degree } k; \text{ else use MC.} \quad (99)$$

By the Lyapunov contraction and angle bounds, this rule is asymptotically consistent.

9.4. Seam Stability Under Composition (Gluing) and Leakage Quantification

Consider a pipeline of L updates (U_ℓ, Φ_ℓ) , $\ell = 1, \dots, L$, where each U_ℓ acts on the blueprint side and each Φ_ℓ on the physical side (retuning, counting, linear optical blocks etc.).

Proposition 9.6 (No inflation under gluing). *If every step is DSFL-admissible, i.e. $\Phi_\ell U_\ell = U_\ell \Phi_\ell$ and $\|\Phi_\ell\| \leq 1$, then the composite update $(U, \Phi) = (U_L \cdots U_1, \Phi_L \cdots \Phi_1)$ satisfies*

$$\mathcal{R}_{\text{res}}^{(K)}(U s^{(K)}, \Phi p) \leq \mathcal{R}_{\text{res}}^{(K)}(s^{(K)}, p) \quad \text{for all } s^{(K)}, p. \quad (100)$$

Proof. The data-processing inequality holds at each step:

$$\mathcal{R}_{\text{res}}^{(K)}(U_\ell s, \Phi_\ell p) \leq \mathcal{R}_{\text{res}}^{(K)}(s, p). \quad (101)$$

Induction on ℓ yields the composite bound. \square

Proposition 9.7 (Angle leakage bound). *Let $\theta_F^{(\ell)}$ be the Friedrichs angle between the degree- K blueprint subspace and the sampling feature subspace after step ℓ . If each Φ_ℓ is a contraction onto a feature subspace, then*

$$\cos \theta_F^{(L)} \leq \min \left\{ 1, \sum_{\ell=1}^L \cos \theta_F^{(\ell)} \right\}. \quad (102)$$

In particular, if $\cos \theta_F^{(\ell)} \leq \eta_\ell$ with $\sum_\ell \eta_\ell < 1$, then $\theta_F^{(L)}$ stays bounded away from 0 and the angle-variance bounds remain uniformly coercive along the pipeline.

Proof sketch. For orthogonal projectors P, Q one has $\|PQ\| = \cos \theta_F(P, Q)$. CS-decomposition and sub-multiplicativity give $\|P_{\text{GBS}}^{(L)} P_{\text{deg } K}\| \leq \|P_{\text{GBS}}^{(L)} P_{\text{GBS}}^{(L-1)}\| \cdots \|P_{\text{GBS}}^{(1)} P_{\text{deg } K}\|$. Each factor is bounded by the corresponding cosine; the asserted ℓ^1 bound follows from the inequality $1 - \prod_\ell (1 - x_\ell) \leq \sum_\ell x_\ell$ applied to $x_\ell = 1 - \cos \theta_F^{(\ell)}$. \square

9.4.1. Consequence

Under calibrated retuning at each seam (so that $m_{t_\ell B} = 2K$), the principal angle cannot collapse in finite length pipelines formed by admissible blocks whose individual misalignments are summable; thus residuals do not inflate and the Lyapunov envelope persists.

10. Experiments

This section describes synthetic and hardware-aligned numerical studies designed to validate the DSFL predictions. We specify the experimental design, the estimators and metrics, and we state formal consistency properties that connect empirical readouts (MSE curves, angle surrogates, advantage percentages) to the theoretical quantities in the preceding sections. Throughout, degree- K quantities are taken at the photon-match $m_{t_B} = 2K$ unless otherwise indicated; empirical laws $\widehat{p}^{(n)}(t)$ are formed from n samples.

10.1. Simulated GBS and Synthetic (B, a_I)

10.1.1. Design.

Fix grids

$$\mathcal{N} \subset \mathbb{N}, \quad \mathcal{K} \subset \mathbb{N}, \quad \text{and choose } (N, K) \in \mathcal{N} \times \mathcal{K}. \quad (103)$$

For each (N, K) :

1. Draw eigenvalue spectra $\{\lambda_j\}_{j=1}^N$ from prescribed *classes*:

$$(\text{flat}) \lambda_j = \lambda \in (0, 1), \quad (\text{polynomial}) \lambda_j = c j^{-\beta}, \quad (\text{exponential}) \lambda_j = c e^{-\beta j}, \quad (104)$$

with amplitude $c \in (0, 1)$ and decay $\beta > 0$ chosen so that $0 < \lambda_N \leq \dots \leq \lambda_1 < 1$.

2. Draw a random Haar orthogonal $U \in O(N)$ and set $B = U \text{diag}(\lambda_1, \dots, \lambda_N) U^\top$.
3. Sample the degree- K blueprint on the unit sphere:

$$s_K = (a_I)_{|I|=2K} \sim \text{Unif}(\mathbb{S}^{D_K-1}), \quad D_K = \#\{I : |I| = 2K\}. \quad (105)$$

4. Compute the matched scale t^* by solving $m_{tB} = 2K$; equivalently, find the unique $t \in (0, \lambda_1^{-1})$ such that $\sum_{j=1}^N t^2 \lambda_j^2 / (1 - t^2 \lambda_j^2) = 2K$.

10.1.2. Estimators and Metrics

For each configuration and for a range of sample sizes n :

- Form $\hat{\mu}_K^{\text{GBS-I}}(n, t^*)$ and $\hat{\mu}_K^{\text{GBS-P}}(n, t^*)$ from n i.i.d. synthetic GBS draws at t^* (via the hafnian-defined pmf).
- Form the plain Monte Carlo estimator $\hat{\mu}_K^{\text{MC}}(n)$ using n i.i.d. $X \sim \mathcal{N}(0, B)$ and Wick reconstruction to isolate degree K if desired.
- Estimate MSEs and empirical residuals by repetitions:

$$\widehat{\text{MSE}}(\bullet) := \frac{1}{R} \sum_{r=1}^R \left(\hat{\mu}^{(r)} - \mu_K \right)^2, \quad \widehat{\mathcal{R}}_{\text{res}}^{(K)}(\hat{p}^{(n)}(t^*)) := \|\hat{p}^{(n)}(t^*) - s_K\|_{\mathcal{H}_K}^2. \quad (106)$$

- Report guaranteed sample sizes $\hat{n}_{\text{GBS-I}}, \hat{n}_{\text{GBS-P}}, \hat{n}_{\text{MC}}$ by inverting Chebyshev/Markov envelopes calibrated with empirical variances.
- Compute efficiency ratios and advantage percentages:

$$\text{Eff. ratio} = \frac{\hat{n}_{\text{MC}}}{\hat{n}_{\text{GBS}}}, \quad \text{Adv. \%} = \frac{1}{T} \sum_{t=1}^T \mathbf{1}\{\hat{n}_{\text{GBS}} \leq \hat{n}_{\text{MC}}\} \times 100\%. \quad (107)$$

10.1.3. Theoretical Guarantees

Proposition 10.1 (Consistency of advantage estimators). *Fix (N, K) and a spectrum class with $0 < \lambda_j < 1$. Let \hat{n}_{GBS} and \hat{n}_{MC} be defined from empirical variances with $R \rightarrow \infty$ repetitions and $n \rightarrow \infty$ samples per repetition at matched t^* . Then*

$$\hat{n}_{\text{GBS}} \xrightarrow{\text{a.s.}} n_{\text{GBS}}^*, \quad \hat{n}_{\text{MC}} \xrightarrow{\text{a.s.}} n_{\text{MC}}^*, \quad (108)$$

where n^* are the population Chebyshev/Markov sample sizes. Consequently, the empirical advantage percentage converges almost surely to the true percentage under the same configuration ensemble.

Proof sketch. Strong laws for empirical second moments of linear statistics of i.i.d. draws imply almost-sure convergence of variance estimators. Continuity of the inversion map $(v \mapsto cv/(\varepsilon^2\delta))$ yields the stated limits; the law of large numbers over configuration draws implies convergence of frequencies. \square

10.2. Ablations Testing DSFL Predictions

10.2.1. T-Sweep (U-Shape)

For each fixed (N, K) and s_K , sweep t in a neighborhood of t^* and evaluate $\widehat{\mathcal{R}}_{\text{res}}^{(K)}(\widehat{p}^{(n)}(t))$ and $\widehat{\text{MSE}}(\widehat{\mu}_K(t))$.

Proposition 10.2 (Empirical convexity and minimum at the match). *Under the simulators above, for any fixed (N, K) and sufficiently large n, R ,*

$$t \mapsto \mathbb{E} \widehat{\mathcal{R}}_{\text{res}}^{(K)}(\widehat{p}^{(n)}(t)) \quad \text{and} \quad t \mapsto \mathbb{E} \widehat{\text{MSE}}(\widehat{\mu}_K(t)) \quad (109)$$

are approximately convex with a unique minimum at $t = t^*$ (the solution of $m_{tB} = 2K$). Moreover, their second derivatives with respect to $\log t$ are strictly positive up to $o_n(1)$ errors.

Proof sketch. The residual and variance/MSE leading factors are proportional to $C_{2K}(t) = t^{-2K} d_{tB}^{-1}$; Lemma (scale calculus) showed $\frac{d^2}{dt^2} \log C_{2K}(t) > 0$ and stationarity $\frac{d}{dt} \log C_{2K}(t) = 0$ iff $m_{tB} = 2K$. Empirical averages converge to their expectations by SLLN, hence the observed convex U-shape and unique minimum at t^* up to sampling error. \square

10.2.2. Angle Surrogate vs Variance

Construct features X_t and blueprint basis Y_K , compute $\widehat{\gamma} = \|X_t^\top Y_K\|$, and correlate $\widehat{\gamma}$ with empirical $n \widehat{\text{Var}}(\widehat{\mu}_K(t))$ across t .

Proposition 10.3 (Monotone trend). *For fixed (N, K) and large n, R , the map $t \mapsto n \widehat{\text{Var}}(\widehat{\mu}_K(t))$ is bounded between two affine functions of $\widehat{\gamma}(t)$:*

$$(1 - \widehat{\gamma}(t)) \widehat{c}_-(a, B) \lesssim n \widehat{\text{Var}}(\widehat{\mu}_K(t)) \lesssim (1 + \widehat{\gamma}(t)) \widehat{c}_+(a, B), \quad (110)$$

with $\widehat{c}_\pm \rightarrow c_\pm$ in probability as $R \rightarrow \infty$.

Proof sketch. Replace the projector norm $\gamma(t)$ by the computable surrogate $\widehat{\gamma}(t) \geq \gamma(t)$; the angle–variance proposition yields the population inequalities. Empirical variances converge to their population values, and the Gram-based surrogate is continuous in sample averages of features, giving the displayed bounds up to $o_{n,R}(1)$ errors. \square

10.2.3. Hybrid Crossover

Partition the blueprint across degrees $k \leq K_{\text{max}}$ and estimate the total samples needed by degree-wise GBS vs MC. Plot the crossover index k_\times where the hybrid policy switches.

Proposition 10.4 (Crossover index is increasing with N and decreasing with decay). *Let $k_\times(N; \beta)$ be the smallest degree where $n_k^{\text{GBS}} \leq n_k^{\text{MC}}$ under a polynomial (or exponential) eigen-decay parameter β . Then for sufficiently large N and fixed tolerance (ε, δ) ,*

$$\partial_N k_\times(N; \beta) \geq 0, \quad \partial_\beta k_\times(N; \beta) \leq 0, \quad (111)$$

i.e. larger systems favor GBS at lower degrees, while stronger decay (smaller effective dimension) delays the crossover.

Proof sketch. The MC dispersion includes mixed-moment growth with the effective dimension and increases rapidly with N when decay is weak; GBS dispersion at the matched degree depends primarily on degree- k features. Comparing the two sample-size envelopes and differentiating w.r.t. the structural parameters yields the stated monotonicities. \square

10.3. Hardware-Aligned Kernels (Optional)

10.3.1. Models

Implement time-bin/loop architectures by stochastic kernels that compose (i) squeezing maps defining tB , (ii) interferometer unitary draws, and (iii) photon-number resolving measurements (or threshold versions). Insert calibrated loss channels modeled as contractions on \mathcal{H}_K .

10.3.2. Stress Tests

For each pipeline, concatenate L calibrated blocks and measure $\widehat{\mathcal{R}}_{\text{res}}^{(K)}$ after each seam, as well as $\widehat{\gamma}(t_\ell)$.

Proposition 10.5 (No inflation under concatenation with calibrated retuning). *Assume each block is DSFL-admissible and retuned to its local match $m_{t_\ell B} = 2K$. Then the empirical residual is nonincreasing along the pipeline up to sampling error:*

$$\mathbb{E} \widehat{\mathcal{R}}_{\text{res}}^{(K)}(\widehat{p}_\ell^{(n)}) \leq \mathbb{E} \widehat{\mathcal{R}}_{\text{res}}^{(K)}(\widehat{p}_{\ell-1}^{(n)}) + o_n(1), \quad \ell = 1, \dots, L. \quad (112)$$

Proof. Apply the DPI for each admissible block (projection/contraction and intertwining), and use convergence of empirical residuals to their expectations. \square

Proposition 10.6 (Robustness to small losses). *Let a loss channel of strength $\eta \in [0, 1)$ act between calibrated blocks. Then the matched scale t_ℓ^* solves the perturbed equation $m_{t_\ell B^\eta} = 2K$ and the contraction rate $\alpha(t_\ell^*)$ degrades by at most $O(\eta)$; consequently, the empirical semi-log MSE slope remains within $O(\eta)$ of the noiseless value for fixed n .*

Proof sketch. A lossy Gaussian channel induces $B \mapsto B^\eta$ with $\|B^\eta - B\| = O(\eta)$; Davis–Kahan perturbation gives $|\gamma^\eta - \gamma| = O(\eta)$, while the scale calculus shows t^* shifts continuously in η . The Lyapunov rate α is monotone in the angle margin, hence perturbs by $O(\eta)$. \square

10.3.3. Reporting

For each setting we present: (i) MSE and residual vs n (semi-log slopes), (ii) MSE vs t (U-shapes with minimum at t^*), (iii) variance vs angle surrogate curves, (iv) advantage heatmaps over (N, K) , and (v) pipeline residual traces across seams with/without loss.

11. Concluding Discussion

This work reframes Gaussian Boson Sampling (GBS) as a *calibrated variance–reduction engine* operating in a single comparison geometry. Within that geometry, the Deterministic Statistical Feedback Law (DSFL) supplies three pillars that, taken together, elevate the “folk rules” of GBS into theorems: (i) *calibration optimality*, which identifies photon–number matching $m_{tB} = 2K$ as the unique minimizer of the leading variance/MSE factor $t^{-2K} d_{tB}^{-1}$; (ii) *principal–angle conditioning*, which yields sharp two–sided dispersion bounds with multiplicative constants $(1 \pm \cos \theta_F)$; and (iii) a *Lyapunov envelope* that converts one–step contraction into a non–asymptotic exponential error decay $\text{MSE}_n \leq C e^{-\alpha(t)n}$ with $\alpha(t)$ maximal at the photon match. These statements are not merely sufficient conditions; they jointly explain *why* GBS beats Monte Carlo (MC) in high–degree regimes, *how* to tune the device to realize the advantage, and *what* rates to expect in finite samples.

On the algorithmic side, the DSFL lens unifies unbiased (GBS–I) and biased–but–consistent (GBS–P) estimators, shows that both inherit the same convex t –dependence, and clarifies the role of degree partitioning: once the target is decomposed into degree components, GBS naturally concentrates effort where MC pays an exponential price through cross–moment couplings, while a hybrid policy remains optimal for small degrees. On the systems side, DSFL’s data–processing inequality guarantees *no residual inflation* under counting, retuning, or block concatenation; principal–angle

control quantifies leakage across interfaces; and perturbation arguments (Davis–Kahan type) explain the *graceful degradation* under small hardware noise together with recovery by re-matching m_{tB} .

The implications are practical. First, calibration reduces to solving a strictly convex, one-dimensional equation for t , hence is robust and fast; this makes photon-number matching a *pre-run, certifiable* step. Second, a computable angle surrogate (e.g., a cross-Gram norm between sampling features and blueprint basis) provides an *a priori gate* for channel selection (GBS vs MC), enabling hybrid deployment without post-hoc tuning. Third, the Lyapunov envelope gives *finite- n guarantees* that can be monitored online by empirical DSFL residuals, turning semi-log MSE plots into quantitative health checks.

There are natural extensions. For non-Gaussian observables or generalized boson-sampling models (loop-hafnians, threshold detection), one must redesign the feature map and admissible updates so that the DPI persists and bias terms remain controlled; the same Hilbertian template applies. Angle surrogates can be refined—e.g. via randomized subspace embeddings—to scale to very large (N, K) while preserving principal-angle fidelity. Finally, the experimental program suggested here— t -sweeps for convex U-shapes, variance-vs-angle curves, and block-wise residual traces—offers *falsifiable* diagnostics for hardware pipelines, and a path to controller-in-the-loop calibration where t and the interferometer are adjusted to maximize the Lyapunov rate $\alpha(t)$ subject to device constraints.

In sum, DSFL furnishes a *predict-then-prove* methodology for GBS: it predicts the optimal operating point ($m_{tB} = 2K$), the conditioning (θ_F), and the rate ($\alpha(t)$) before data are collected; it proves that admissible updates cannot worsen the calibrated misfit; and it turns those predictions into actionable rules that persist under composition and mild noise. This synthesis clarifies when and why GBS delivers sample-complexity advantages over MC, and how to tune photonic pipelines to realize them in practice.

Funding: No external funding was received for this work.

Data and code availability: This work uses only synthetic data produced by the algorithms described in the paper; no external datasets were used or analyzed. A fully reproducible package containing (i) source code, (ii) parameter files for every figure, (iii) exact random seeds, and (iv) an exportable runtime environment (both `environment.yml` and `requirements.txt`) will be deposited at Zenodo under an archival DOI (to be added at acceptance) and mirrored on GitHub (read-only tag matching the DOI). The repository includes a one-click reproducibility script: which regenerates all tables and figures in the manuscript. Tested platforms: Linux (x86_64) and macOS (arm64) with Python ≥ 3.10 ; core dependencies: NumPy, SciPy, Matplotlib, and JAX/Numba (optional, for accelerated runs). Code will be released under an OSI-approved license (MIT).

Acknowledgments: The author affirms sole authorship and received no external funding.

Conflicts of Interest: The author declares no competing interests.

Appendix A. Crisp Theorem Statements (Ready to Formalize)

Appendix A.1. Standing Setup

Let $(\mathcal{H}, \langle \cdot, \cdot \rangle)$ be a Hilbert space with norm $\|x\| = \sqrt{\langle x, x \rangle}$, and closed subspaces $S, P \subset \mathcal{H}$. A *calibration* is a pair of bounded maps

$$\mathcal{I} : S \rightarrow P, \quad \mathcal{J} : P \rightarrow S, \quad (\text{A1})$$

satisfying the *interchangeability* identities

$$\mathcal{I}\mathcal{J} = \text{id}_P, \quad \mathcal{J}\mathcal{I} = P_S, \quad (\text{A2})$$

with P_S the orthogonal projector onto S . For $s \in S$ and $p \in P$, the *Residual of Sameness* is

$$\mathcal{R}(s, p) := \|p - \mathcal{I}s\|_{\mathcal{H}}^2. \quad (\text{A3})$$

An *admissible update* is a pair $(\tilde{\Phi}, \Phi)$ with $\tilde{\Phi} : S \rightarrow S$, $\Phi : P \rightarrow P$ such that

$$\Phi \mathcal{I} = \mathcal{I} \tilde{\Phi} \quad \text{and} \quad \|\Phi\|_{\mathcal{H} \rightarrow \mathcal{H}} \leq 1. \quad (\text{A4})$$

For angle bounds, let $M, N \subset \mathcal{H}$ be closed subspaces with orthogonal projectors P_M, P_N , and *Friedrichs cosine*

$$\gamma(M, N) := \|P_M P_N\| \in [0, 1], \quad \cos \theta_F(M, N) = \gamma(M, N). \quad (\text{A5})$$

DPI (GBS Admissibility)

Theorem A.1 (Data-processing inequality). *If $(\tilde{\Phi}, \Phi)$ is admissible, then for all $s \in S$ and $p \in P$,*

$$\mathcal{R}(\tilde{\Phi}s, \Phi p) \leq \mathcal{R}(s, p). \quad (\text{A6})$$

Proof. Using $\Phi \mathcal{I} = \mathcal{I} \tilde{\Phi}$,

$$\mathcal{R}(\tilde{\Phi}s, \Phi p) = \|\Phi p - \mathcal{I} \tilde{\Phi}s\|^2 = \|\Phi(p - \mathcal{I}s)\|^2 \leq \|\Phi\|^2 \|p - \mathcal{I}s\|^2 \leq \|p - \mathcal{I}s\|^2. \quad \square \quad (\text{A7})$$

Corollary A.2 (Composition). *If $\{(\tilde{\Phi}_k, \Phi_k)\}_{k=1}^L$ are admissible, then $(\tilde{\Phi}, \Phi) = (\prod_{k=1}^L \tilde{\Phi}_k, \prod_{k=1}^L \Phi_k)$ is admissible and*

$$\mathcal{R}(\tilde{\Phi}s, \Phi p) \leq \mathcal{R}(s, p). \quad (\text{A8})$$

Proof. Intertwining and $\|\cdot\| \leq 1$ are preserved under composition; apply Theorem A.1. \square

Calibration optimality $\Rightarrow m_{tB} = 2K$

Appendix A.2. GBS Specialization

Let $B \in \mathbb{R}^{N \times N}$ be symmetric with $0 < \lambda_N \leq \dots \leq \lambda_1 < 1$, and fix a degree $K \in \mathbb{N}$. For a scale $t \in (0, \lambda_1^{-1})$, define

$$d_{tB} := \prod_{j=1}^N \sqrt{1 - t^2 \lambda_j^2}, \quad C_{2K}(t) := \frac{t^{-2K}}{d_{tB}}, \quad m_{tB} := \sum_{j=1}^N \frac{t^2 \lambda_j^2}{1 - t^2 \lambda_j^2}. \quad (\text{A9})$$

Assume the degree- K residual admits the leading form $\mathcal{R}_{\text{res}}^{(K)}(t) = A C_{2K}(t) - B$ with $A > 0$ independent of t .

Lemma A.3 (Scale calculus). $\frac{d}{dt} \log C_{2K}(t) = \frac{m_{tB} - 2K}{t}$, and $\frac{d^2}{dt^2} \log C_{2K}(t) = \frac{2K}{t^2} + \sum_{j=1}^N \frac{\lambda_j^2}{1 - t^2 \lambda_j^2} +$

$$\sum_{j=1}^N \frac{2t^2 \lambda_j^4}{(1 - t^2 \lambda_j^2)^2} > 0.$$

Proof. Differentiate $\log C_{2K}(t) = -2K \log t - \frac{1}{2} \sum_j \log(1 - t^2 \lambda_j^2)$ termwise. \square

Theorem A.4 (Calibration optimality). *The unique minimizer $t^* \in (0, \lambda_1^{-1})$ of $t \mapsto \mathcal{R}_{\text{res}}^{(K)}(t)$ is characterized by the photon-number match*

$$m_{t^*B} = 2K. \quad (\text{A10})$$

Equivalently, t^ minimizes the leading t -factors in the variance/MSE of the degree- K GBS estimators.*

Proof. By Lemma A.3, critical points of C_{2K} satisfy $m_{tB} = 2K$ and C_{2K} is strictly convex, hence the minimizer is unique. The variance/MSE leading factors for GBS-I and GBS-P are proportional to $C_{2K}(t)$; thus they share the same minimizer. \square

What is provable once clearly stated assumptions are added (action items)

- *Angle \Rightarrow rate monotonicity.* We state that the Lyapunov rate $\alpha(t)$ improves as $\gamma(t) = \|P_{\text{GBS}}(t)P_{\text{deg } K}\|$ decreases. **Action:** Include Lemma B.2, which bounds the rate between linear functions of the angle margin:

$$c_- (1 - \gamma(t)) \leq \alpha(t) \leq c_+ (1 - \gamma(t)), \quad (\text{A11})$$

with t -independent $0 < c_- \leq c_+ < \infty$. The proof is a two-subspace argument using the through-subspace contraction (Proposition 6.8) and the sharp bounds in Proposition A.7.

- *Residual factorization.* We use $\mathcal{R}^{(K)}(t) = A C_{2K}(t) - B$ to map calibration to residual minimization. **Action:** Add Lemma A.5 (Degree- K residual factorization) showing that all t -dependence factors through t^{-2K} and d_{tB}^{-1} for the degree- K laws. A concise proof sketch is included in the appendix, relying on hafnian homogeneity $\text{Haf}((tB)_I) = t^K \text{Haf}(B_I)$ and linearity of the readouts.

Lemma A.5 (Residual factorization for degree- K). Fix $K \in \mathbb{N}$, a covariance B with spectrum in $(0, 1)$, and coefficients $(a_I)_{|I|=2K}$. Let $t \in (0, \lambda_1^{-1})$, set $d_{tB} = \prod_{j=1}^N \sqrt{1 - t^2 \lambda_j^2}$ and

$$C_{2K}(t) := \frac{t^{-2K}}{d_{tB}}.$$

Consider the degree- K DSFL residual $\mathcal{R}_{\text{res}}^{(K)}(t) := \|p(t) - \mathcal{I}s_K\|_{\mathcal{H}_K}^2$ built from either readout

$$\text{GBS-P: } \mathcal{I}^P x = \sum_{|I|=2K} a_I \frac{x_I}{d_{tB}}, \quad \text{GBS-I: } \mathcal{I}^I x = \sum_{|I|=2K} a_I \frac{I!}{d_{tB}} x_I.$$

Then there exist t -independent constants $A > 0$ and $B \in \mathbb{R}$ (depending only on (a_I) and B) such that

$$\mathcal{R}_{\text{res}}^{(K)}(t) = A C_{2K}(t) - B.$$

Sketch. For $|I| = 2K$, hafnian homogeneity gives $\text{Haf}((tB)_I) = t^K \text{Haf}(B_I)$, whence $p_I(t) = d_{tB} t^{2K} \text{Haf}(B_I)^2 / I!$. Thus every coordinate of $p(t)$ carries the common scalar factor $d_{tB} t^{2K}$. The degree- K readouts introduce only the normalization d_{tB}^{-1} (and $I!$ in GBS-I), which does not depend on t beyond d_{tB} . Expanding $\|p(t) - \mathcal{I}s_K\|^2$ shows the t -dependence factors through $t^{-2K} d_{tB}^{-1} = C_{2K}(t)$, with all inner products collected into t -independent constants $A > 0$ and $B \in \mathbb{R}$. A fully expanded computation (including the multinomial covariance form for GBS-P and the exact second moment for GBS-I) is provided in Appendix A. \square

Corollary A.6 (Calibration \Rightarrow residual minimization). Under the hypotheses of Lemma A.5, the unique minimizer of $t \mapsto \mathcal{R}_{\text{res}}^{(K)}(t)$ on $(0, \lambda_1^{-1})$ is characterized by $m_{tB} = 2K$.

Proof. By Lemma A.5, $\mathcal{R}_{\text{res}}^{(K)}(t) = A C_{2K}(t) - B$ with $A > 0$. The scale calculus shows $\frac{d}{dt} \log C_{2K}(t) = \frac{m_{tB} - 2K}{t}$ and $t \mapsto C_{2K}(t)$ is strictly convex on $(0, \lambda_1^{-1})$. Hence the unique minimizer of $\mathcal{R}_{\text{res}}^{(K)}(t)$ satisfies $m_{tB} = 2K$. \square

Proof. By hafnian homogeneity, $p_I(t) = d_{tB} t^{2K} \text{Haf}(B_I)^2 / I!$. Hence $\|p(t)\|^2 = \sum_I p_I(t)^2 = d_{tB}^2 t^{4K} \sum_I \text{Haf}(B_I)^4 / I!^2 = d_{tB} t^{2K} \cdot (t^{2K} d_{tB} \sum_I \text{Haf}^4 / I!^2)$. The parenthesized factor is independent of t after regrouping; similarly, the cross term $\langle p(t), \Pi s_K \rangle = \sum_I p_I(t) (\Pi s_K)_I$ is proportional to $d_{tB} t^{2K}$ times a t -independent sum, and $\|\Pi s_K\|^2$ is t -independent. Expanding $\|p(t) - \Pi s_K\|^2$ yields the claimed affine dependence on $C_{2K}(t) = t^{-2K} d_{tB}^{-1}$ with $A > 0$. \square

Angle bounds

Proposition A.7 (Two-subspace inequality). *Let $M, N \subset \mathcal{H}$ be closed subspaces with $\gamma = \|P_M P_N\| \in [0, 1]$. For any $a \in M$ and $b \in N$,*

$$(1 - \gamma)(\|a\|^2 + \|b\|^2) \leq \|a + b\|^2 \leq (1 + \gamma)(\|a\|^2 + \|b\|^2). \quad (\text{A12})$$

Proof. Expand $\|a + b\|^2 = \|a\|^2 + \|b\|^2 + 2 \operatorname{Re}\langle a, b \rangle$, and use $|\langle a, b \rangle| = |\langle a, P_M b \rangle| \leq \|a\| \|P_M b\| \leq \gamma \|a\| \|b\|$, then $2\|a\| \|b\| \leq \|a\|^2 + \|b\|^2$. \square

Appendix A.3. GBS specialization (degree- K).

Let $P_{\text{deg } K}$ be the blueprint projector and $P_{\text{GBS}}(t)$ the sampling-feature projector at scale t , and define $\gamma(t) = \|P_{\text{GBS}}(t)P_{\text{deg } K}\| = \cos \theta_F(t)$.

Theorem A.8 (Angle–dispersion link). *For the degree- K estimators from n samples at fixed t , there exist t -independent $c_-(a, B), c_+(a, B) > 0$ such that*

$$(1 - \gamma(t))c_-(a, B) \leq n \operatorname{Var}(\hat{\mu}_K) \leq (1 + \gamma(t))c_+(a, B), \quad (\text{A13})$$

with Var understood as the variance for GBS-I and as the leading asymptotic variance (or squared-bias term) for GBS-P. Equivalently, MSE is sandwiched between two multiples of the split energy, with multiplicative factors $(1 \pm \cos \theta_F(t))$.

Proof sketch. Write the estimator as $\hat{\mu}_K = \langle w, \hat{p}(t) \rangle$ with w fixed and $\hat{p}(t)$ the empirical law. Then $n \operatorname{Var}(\hat{\mu}_K) = \langle \Sigma(t)w, w \rangle$, where $\Sigma(t)$ is the single-sample covariance. The range of $\Sigma(t)$ lies in $P_{\text{GBS}}(t)\mathcal{H}$, while w decomposes into $P_{\text{deg } K}w + w_\perp$. Apply Proposition A.7 twice (once to control leakage into $P_{\text{GBS}}(t)$, once for the split relative to $P_{\text{deg } K}$) and absorb t -independent weights into $c_\pm(a, B)$. \square

Lyapunov bound

Theorem A.9 (Discrete Lyapunov envelope). *Let $\mathcal{R}_{\text{res}}^{(K)}$ be the degree- K residual and suppose there exists $\alpha(t) \in (0, 1]$ such that, conditionally on the past,*

$$\mathbb{E} \left[\mathcal{R}_{\text{res}}^{(K)}(U s^{(K)}, \Phi p^{(n)}(t)) \mid p^{(n)}(t) \right] \leq (1 - \alpha(t)) \mathcal{R}_{\text{res}}^{(K)}(s^{(K)}, p^{(n)}(t)), \quad (\text{A14})$$

for an admissible one-step update (U, Φ) . Then

$$\mathbb{E} \mathcal{R}_{\text{res}}^{(K)}(p^{(n)}(t)) \leq e^{-\alpha(t)n} \mathcal{R}_{\text{res}}^{(K)}(p^{(0)}(t)). \quad (\text{A15})$$

Moreover, $\alpha(t)$ increases as $\gamma(t)$ decreases, and is maximized at the calibration match $m_{tB} = 2K$.

Proof. Taking total expectation yields $\mathbb{E}R_{n+1} \leq (1 - \alpha)\mathbb{E}R_n$ with $R_n := \mathcal{R}_{\text{res}}^{(K)}(p^{(n)}(t))$. Iteration gives $\mathbb{E}R_n \leq (1 - \alpha)^n R_0 \leq e^{-\alpha n} R_0$. By Theorem A.8, improved alignment (smaller γ) strengthens contraction; by Theorem A.4, the match maximizes this gain. \square

Corollary A.10 ((ε, δ) sample-size envelope). *If $n \geq \alpha(t)^{-1} \log(\mathcal{R}_{\text{res}}^{(K)}(p^{(0)}(t))/\varepsilon^2)$, then $\mathbb{E} \mathcal{R}_{\text{res}}^{(K)}(p^{(n)}(t)) \leq \varepsilon^2$. At the match $m_{tB} = 2K$, the right-hand side is minimized in t .*

Proof. Solve $e^{-\alpha(t)n} \mathcal{R}_{\text{res}}^{(K)}(p^{(0)}(t)) \leq \varepsilon^2$ for n . \square

Appendix B. Explaining the Elements: sDoF/pDoF, Interchangeability, DSFL and the Residual of Sameness \mathcal{R}

This section fixes the primitives used throughout. We first embed the *statistical degrees of freedom* (sDoF) and *physical degrees of freedom* (pDoF) in one comparison Hilbert space; then we define the single observable (*Residual of Sameness*); finally we state the admissibility gate that makes the residual monotone and, when dynamics are present, yields a Lyapunov envelope.

Appendix B.1. Interchangeability (Calibration)

Let $(\mathcal{H}, \langle \cdot, \cdot \rangle)$ be a Hilbert space carrying two closed subspaces S (statistical channel) and P (physical channel). A calibration is a pair of bounded maps

$$\mathcal{I} : S \rightarrow P, \quad \mathcal{J} : P \rightarrow S \quad (\text{A16})$$

satisfying

$$\mathcal{I}\mathcal{J} = \text{id}_P, \quad \mathcal{J}\mathcal{I} = P_S, \quad (\text{A17})$$

with P_S the orthogonal projector onto S . Thus \mathcal{J} provides calibrated statistical representatives of $p \in P$, \mathcal{I} provides calibrated physical images of $s \in S$, and composing in either channel returns the original.

Appendix B.2. One Observable: the Residual of Sameness

For $s \in S$ and $p \in P$,

$$\mathcal{R}(s, p) := \|p - \mathcal{I}(s)\|_{\mathcal{H}}^2 \quad (\text{A18})$$

is the squared distance from p to $\text{ran } \mathcal{I}$. It vanishes iff $p = \mathcal{I}(s)$ and is invariant under isometries preserving $\text{ran } \mathcal{I}$.

Appendix B.3. Admissibility and the Data–Processing Inequality

A pair $(\tilde{\Phi}, \Phi)$ with $\tilde{\Phi} : S \rightarrow S$ and $\Phi : P \rightarrow P$ is *admissible* if

$$\Phi\mathcal{I} = \mathcal{I}\tilde{\Phi}, \quad \|\Phi\|_{\mathcal{H} \rightarrow \mathcal{H}} \leq 1, \quad \|\tilde{\Phi}\|_{\mathcal{H} \rightarrow \mathcal{H}} \leq 1. \quad (\text{A19})$$

Then the Hilbertian data–processing inequality holds:

$$\mathcal{R}(\tilde{\Phi}s, \Phi p) = \|\Phi(p - \mathcal{I}s)\|_{\mathcal{H}}^2 \leq \|p - \mathcal{I}s\|_{\mathcal{H}}^2 = \mathcal{R}(s, p). \quad (\text{A20})$$

Hence no admissible update inflates the calibrated mismatch. In operator–algebraic models, \mathcal{J} is an orthogonal conditional expectation, and (34) follows from Kadison–Schwarz.

Appendix B.4. Propagation with a Gap: the DSFL Lyapunov Envelope

Assume a residual identity

$$\frac{d}{dt} \mathcal{R}(t) = -2\langle Ke, e \rangle_{\mathcal{H}} + 2\langle e, g \rangle_{\mathcal{H}}, \quad e := p - \mathcal{I}s, \quad K = K^* \succeq 0, \quad (\text{A21})$$

and bounds

$$\langle Ke, e \rangle_{\mathcal{H}} \geq \kappa \|e\|_{\mathcal{H}}^2, \quad |\langle e, g \rangle_{\mathcal{H}}| \leq \varepsilon \|e\|_{\mathcal{H}}^2, \quad \kappa > \varepsilon \geq 0. \quad (\text{A22})$$

Then

$$\frac{d}{dt} \mathcal{R}(t) \leq -\alpha \mathcal{R}(t), \quad \alpha := 2(\kappa - \varepsilon) > 0, \quad \Rightarrow \quad \mathcal{R}(t) \leq e^{-\alpha(t-t_0)} \mathcal{R}(t_0). \quad (\text{A23})$$

Thus $\log \mathcal{R}$ decays at a single rate α fixed by the sector’s coercivity margin.

Appendix B.5. Operational Arrow (Accumulated Improvement)

Define

$$S_R(t) := -\log \frac{\mathcal{R}(t)}{\mathcal{R}(0)} = \int_0^t \alpha(\tau) d\tau. \quad (\text{A24})$$

It is nondecreasing and quantifies total calibrated mismatch removed, independently of the choice of clock.

Appendix B.6. Minimal Checklist for Applications

1. Fix \mathcal{H}, S, P ; choose $(\mathcal{I}, \mathcal{J})$ with (30).
2. Verify (A19) for each step; (34) then guarantees monotonicity.
3. If dynamics are available, identify (κ, ε) in (A22) to obtain α in (A23).

Appendix B.7. Generic Two-Channel Application Template

Let $s(t) \in S$ and $p(t) \in P$ with calibration $\mathcal{I} : S \rightarrow P$. Set $e(t) := p(t) - \mathcal{I}s(t)$, $\mathcal{R}(t) := \|e(t)\|_{\mathcal{H}}^2$. If

$$\dot{\mathcal{R}}(t) = -2\langle Ke, e \rangle_{\mathcal{H}} + 2\langle e, g \rangle_{\mathcal{H}}, \quad \langle Ke, e \rangle_{\mathcal{H}} \geq \kappa \|e\|_{\mathcal{H}}^2, \quad |\langle e, g \rangle_{\mathcal{H}}| \leq \varepsilon \|e\|_{\mathcal{H}}^2, \quad (\text{A25})$$

then $\dot{\mathcal{R}} \leq -2(\kappa - \varepsilon)\mathcal{R}$ and $\mathcal{R}(t) \leq e^{-2(\kappa - \varepsilon)(t - t_0)}\mathcal{R}(t_0)$. This dynamic envelope pairs with the static monotonicity $\mathcal{R}(\tilde{\Phi}s, \Phi p) \leq \mathcal{R}(s, p)$ whenever $(\tilde{\Phi}, \Phi)$ is admissible.

Appendix B.8. Terminology and Mechanism: sDoF/pDoF, DSFL, and the Dual Feedback Loop

Appendix B.8.1. What are sDoF and pDoF?

The *statistical degrees of freedom* (sDoF) are the blueprint variables $s \in S$ that encode the model, prior, or control law; the *physical degrees of freedom* (pDoF) are the measured/realized responses $p \in P$. Both live in the same comparison Hilbert space $(\mathcal{H}, \langle \cdot, \cdot \rangle)$ via a calibration $(\mathcal{I}, \mathcal{J})$ obeying

$$\mathcal{I}\mathcal{J} = \text{id}_P, \quad \mathcal{J}\mathcal{I} = P_S. \quad (\text{A26})$$

Equation (A26) ensures two-way coherence: \mathcal{J} reads a calibrated blueprint from a physical state, and \mathcal{I} writes a calibrated physical image of a statistical state. The *Residual of Sameness*

$$\mathcal{R}(s, p) = \|p - \mathcal{I}s\|_{\mathcal{H}}^2 \quad (\text{A27})$$

measures the mismatch between pDoF and the calibrated sDoF within \mathcal{H} .

Appendix B.8.2. Why “DSFL”

Deterministic Statistical Feedback Law (DSFL) is shortened to *DSFL* because updates are deterministic (linear/contractive), the driver is statistical (sDoF), the mechanism is feedback (via the residual $e := p - \mathcal{I}s$), and it functions as a law (one inequality governs all steps). For any *admissible* update $(\tilde{\Phi}, \Phi)$ ($\Phi\mathcal{I} = \mathcal{I}\tilde{\Phi}$, $\|\Phi\|, \|\tilde{\Phi}\| \leq 1$),

$$\mathcal{R}(\tilde{\Phi}s, \Phi p) = \|\Phi(p - \mathcal{I}s)\|_{\mathcal{H}}^2 \leq \|p - \mathcal{I}s\|_{\mathcal{H}}^2 = \mathcal{R}(s, p), \quad (\text{A28})$$

so the residual is L^2 -monotone.

Appendix B.8.3. Dual Feedback Loop (Discrete Time)

There are two tightly coupled feedback paths in the same geometry:

$$\text{S-loop: } s_{k+1} = \tilde{\Phi} s_k, \quad \text{P-loop: } p_{k+1} = \Phi p_k, \quad e_{k+1} = p_{k+1} - \mathcal{I}s_{k+1}, \quad (\text{A29})$$

with $\Phi\mathcal{I} = \mathcal{I}\tilde{\Phi}$ and $\|\Phi\|, \|\tilde{\Phi}\| \leq 1$. By (A28), $\mathcal{R}(s_{k+1}, p_{k+1}) \leq \mathcal{R}(s_k, p_k)$ sample-wise. If, in addition, a Lyapunov identity with a gap holds, $\dot{\mathcal{R}} \leq -\alpha \mathcal{R}$, then $\mathcal{R}_k \leq Ce^{-\alpha k}$ (exponential envelope).

Appendix B.8.4. Flow form (Continuous Time)

Writing $e(t) = p(t) - \mathcal{I}s(t)$, assume

$$\frac{d}{dt} \mathcal{R}(t) = -2\langle Ke, e \rangle_{\mathcal{H}} + 2\langle e, g \rangle_{\mathcal{H}}, \quad \langle Ke, e \rangle \geq \kappa \|e\|^2, \quad |\langle e, g \rangle| \leq \varepsilon \|e\|^2, \quad (\text{A30})$$

with $\kappa > \varepsilon \geq 0$. Then $\frac{d}{dt} \mathcal{R} \leq -2(\kappa - \varepsilon)\mathcal{R}$ and $\mathcal{R}(t) \leq e^{-2(\kappa - \varepsilon)(t - t_0)} \mathcal{R}(t_0)$. The accumulated drop $S_R(t) = -\log \frac{\mathcal{R}(t)}{\mathcal{R}(0)} = \int_0^t 2(\kappa - \varepsilon) d\tau$ is nondecreasing and quantifies the total calibrated mismatch removed.

Appendix B.8.5. Practical Interpretation

The P-loop evolves pDoF through an L^2 -contractive map that respects calibration, reducing e on the physical side; the S-loop refines sDoF so that $\mathcal{I}s$ tracks the evolving p . Because both loops live in \mathcal{H} and intertwine $(\mathcal{I}, \mathcal{J})$, they cannot fabricate residual—only dissipate or transport it. In applications with slow drift, one may update the calibration $(\mathcal{I}, \mathcal{J}) \mapsto (\mathcal{I}^{(m)}, \mathcal{J}^{(m)})$ on a coarser timescale as long as $\mathcal{I}^{(m)} \mathcal{J}^{(m)} = \text{id}_{\mathcal{P}}$, $\mathcal{J}^{(m)} \mathcal{I}^{(m)} = P_{\mathcal{S}}$; between updates the inner dual loop remains admissible and residual-monotone.

What Is Fully Proved (or Reducible to Standard Facts)

Appendix B.8.6. Data-Processing Inequality (DPI) in the DSFL Geometry

Claim. Admissible updates (intertwining $\Phi\mathcal{I} = \mathcal{I}\tilde{\Phi}$ and nonexpansive $\|\Phi\|_{\mathcal{H} \rightarrow \mathcal{H}} \leq 1$) cannot inflate the residual.

Status. Proved in the text (Theorem “DPI”) by a one-line Hilbert-space contraction; stability under composition is immediate.

Appendix B.8.7. Calibration Optimality $m_{tB} = 2K$

Claim. The unique minimizer of the leading variance/MSE factor is the photon-number match.

Status. Proved by strict convexity on $(0, \lambda_1^{-1})$ of

$$\log C_{2K}(t) = -2K \log t - \frac{1}{2} \sum_j \log(1 - t^2 \lambda_j^2), \quad \frac{d}{dt} \log C_{2K}(t) = \frac{m_{tB} - 2K}{t}. \quad (\text{A31})$$

Scope. Optimal for any quantity whose t -dependence factors through $C_{2K}(t)$: the leading terms of the GBS-I variance and GBS-P MSE, and the DSFL residual representation $\mathcal{R}_{\text{res}}^{(K)}(t) = A C_{2K}(t) - B$ (when that representation holds).

Appendix B.8.8. Angle (Friedrichs) Bounds

Claim. Two-sided MSE/variance control with factors $(1 \pm \cos \theta_F)$.

Status. Proved via the two-subspace inequalities and $\|P_M P_N\| = \cos \theta_F$; explicit derivations included.

Appendix B.8.9. Exact Variance for GBS-I; Asymptotic MSE for GBS-P

GBS-I. Exact variance formula proved from the pmf and hafnian homogeneity.

GBS-P. Large- n leading term proved by the multinomial delta method (gradient and covariance computed).

Scope/assumptions. For GBS-P we assume the usual sign-known convention and interior probabilities/smoothness so the delta method applies.

Appendix B.8.10. Chebyshev/Markov (ε, δ) Guarantees

Claim. Invert variance/MSE to obtain guaranteed sample sizes.

Status. Proved once the variance/MSE statements hold (and $\mu \neq 0$), by direct Chebyshev/Markov bounds.

Appendix B.8.11. Lyapunov Envelopes (Discrete/Continuous)

Claim. If a one-step conditional contraction holds, the expected residual decays exponentially.

Status. Proved as an implication via discrete (and continuous) Grönwall.

Scope. The contraction is the hypothesis; the envelope follows rigorously.

What Is Provable Once Clearly Stated Assumptions Are Added

Appendix B.8.12. Residual Representation

Where used. To map calibration optimality to residual minimization.

Status. Provable for degree- K laws and the two estimators, since the only nontrivial t -dependence enters through t^{-2K} and d_{tB}^{-1} . One factors t out of (i) the multinomial covariance and (ii) the linear readout; a short lemma closes the argument.

Appendix B.8.13. “Angle \Rightarrow rate” Linkage

Claim. The Lyapunov rate $\alpha(t)$ improves as $\gamma(t) = \cos \theta_F(t)$ decreases.

Status. Provable once the conditional contraction constant is written as a monotone functional of the split energy controlled by $(1 \pm \gamma)$. With the projector bounds in hand, a brief comparison lemma formalizes this.

Appendix B.8.14. Composition/Gluing Stability and Small-Noise Robustness

Claim. No inflation under concatenation; small perturbations degrade angles/rates by $O(\|\Delta\|)$.

Status. Concatenation: already follows from DPI under composition (proved). Robustness: Davis–Kahan/Wedin gives

$$|\sin \theta_F(B + \Delta) - \sin \theta_F(B)| = O\left(\frac{\|\Delta\|}{\text{gap}}\right), \quad (\text{A32})$$

hence γ perturbs by $O(\|\Delta\|)$ and rates follow from the previous item.

Appendix B.8.15. Asymptotic Normality (CLT) for GBS-P

Claim. Beyond the leading MSE constant, a full CLT holds.

Status. Provable by standard smooth-functionals-of-the-multinomial arguments under mild regularity (already used for the delta-method step).

What Is Solid but We Cite Rather than Re-Prove Here

Corollary B.1 (CLT for GBS-P). *Let $\mathcal{I}_K = \{I : |I| = 2K\}$ and write $m := |\mathcal{I}_K|$. Fix a scale $t \in (0, \lambda_1^{-1})$ and assume the GBS law*

$$p(t) = (p_I(t))_{I \in \mathcal{I}_K} \in (0, 1)^m, \quad \sum_{I \in \mathcal{I}_K} p_I(t) = 1, \quad (\text{A33})$$

i.e. all degree- K outcome probabilities are strictly interior. Let the degree- K GBS-P functional be

$$g(p) := \sum_{I \in \mathcal{I}_K} a_I t^{-K} \sqrt{|I|} \sigma_I \sqrt{\frac{p_I}{d_{tB}}}, \quad d_{tB} = \prod_{j=1}^N \sqrt{1 - t^2 \lambda_j^2}, \quad (\text{A34})$$

with $\sigma_I = \text{sign}(\text{Haf}(B_I))$ (or absorbed into a_I when $\text{Haf}(B_I) \geq 0$). Let $\hat{p}(t)$ be the empirical law based on n i.i.d. samples from $p(t)$, and $\hat{\mu}_K^{\text{GBS-P}} := g(\hat{p}(t))$. Then

$$\sqrt{n} (\hat{\mu}_K^{\text{GBS-P}} - \mu_K) \xrightarrow{d} \mathcal{N}\left(0, \nabla g(p)^\top (\text{diag}(p) - pp^\top) \nabla g(p)\right), \quad (\text{A35})$$

where

$$\frac{\partial g}{\partial p_I}(p) = \frac{1}{2} a_I t^{-K} \sqrt{I!} \sigma_I \frac{1}{\sqrt{d_{tB}}} p_I^{-1/2} \quad (I \in \mathcal{I}_K). \quad (\text{A36})$$

In particular, the asymptotic variance simplifies to

$$\nabla g(p)^\top (\text{diag}(p) - pp^\top) \nabla g(p) = \frac{1}{4} \left[\frac{1}{d_{tB} t^{2K}} \sum_{|I|=2K} \frac{a_I^2}{I!} - \mu_K^2 \right], \quad (\text{A37})$$

which matches the leading constant in Proposition 8.3.

[39,40]

Proof. Write $S = (S_I)_{I \in \mathcal{I}_K} \sim \text{Multinomial}(n, p)$ and $\hat{p} = S/n$. The interior condition $p_I \in (0, 1)$ implies the classical multinomial central limit theorem:

$$\sqrt{n} (\hat{p} - p) \xrightarrow{d} \mathcal{N}(0, \Sigma(p)), \quad \Sigma(p) := \text{diag}(p) - pp^\top. \quad (\text{A38})$$

Since g is C^1 on the open simplex (each coordinate appears under a square root and $p_I > 0$), the delta method yields

$$\sqrt{n} (g(\hat{p}) - g(p)) \xrightarrow{d} \mathcal{N}\left(0, \nabla g(p)^\top \Sigma(p) \nabla g(p)\right). \quad (\text{A39})$$

The displayed gradient follows from $g(p) = \sum_I c_I \sqrt{p_I}$ with $c_I = a_I t^{-K} \sqrt{I!} \sigma_I / \sqrt{d_{tB}}$, so $\partial g / \partial p_I = \frac{1}{2} c_I p_I^{-1/2}$. Finally,

$$\sum_I p_I \left(\frac{\partial g}{\partial p_I} \right)^2 = \frac{1}{4} \sum_I \frac{c_I^2}{1} = \frac{1}{4} \frac{1}{d_{tB} t^{2K}} \sum_{|I|=2K} \frac{a_I^2}{I!}, \quad \left(\sum_I p_I \frac{\partial g}{\partial p_I} \right)^2 = \mu_K^2, \quad (\text{A40})$$

because $g(p) = \mu_K$ by construction (Wick + homogeneity). Hence $\nabla g^\top \Sigma(p) \nabla g = \sum_I p_I (\partial g / \partial p_I)^2 - \left(\sum_I p_I \partial g / \partial p_I \right)^2 = \frac{1}{4} [d_{tB}^{-1} t^{-2K} \sum a_I^2 / I! - \mu_K^2]$. \square

Remark B.1 (Boundary cases). *If some coordinates satisfy $p_I = 0$, the map $p \mapsto g(p)$ is not differentiable at p ; one may either (i) remove those indices from \mathcal{I}_K (they are p -null) and re-normalize, or (ii) work with a directional delta method under sparsity assumptions. The interior condition in the corollary avoids these technicalities.*

Appendix B.8.16. Open, Nonempty Regions of Advantage; Percentage-of-Advantage

Content (Andersen–Shan 2025). Large- n leading term for GBS-P (matching ours), improved tuning via $m_{tB} = 2K$, (ε, δ) sample-size envelopes for both estimators, existence of nonempty open subsets with advantage over MC, and percentage-of-advantage values over (a, B) under Haar–Vandermonde measures.

Status. We rely on these theorems and cite them explicitly.

What Is Still Heuristic/Needs Additional Work

Appendix B.8.17. Principal-angle surrogate $\hat{\gamma} = \|X_t^\top Y_K\|_2$ as a certified upper bound.

Claim. The computable Gram surrogate upper-bounds the true Friedrichs cosine.

Status. Heuristic but tighten-able: if X_t spans the sampling feature subspace and Y_K spans the blueprint

subspace (properly normalized), then $\|X_t^\top Y_K\| \uparrow \|P_{\text{GBS}}(t)P_{\text{deg } K}\|$. A short approximation lemma with sampling-error bounds can make this rigorous.

Appendix B.8.18. Regime Separation (Poly vs. Exp).

Claim. For $K \sim \zeta N^2$, $n_{\text{GBS}} = \text{poly}(N)$ while $n_{\text{MC}} = \exp(\Theta(N))$.

Status. Proof sketch given; a fully rigorous version needs fixed ensembles (eigenvalues in $[c, 1 - \varepsilon]$; coefficients on the unit sphere) and a positive-mass lower bound for the MC mixed-hafnian quadratic form. With these in place, the exponential lower bound follows by counting/Stirling estimates (in line with Andersen–Shan open-set results).

Appendix B.8.19. One-Step Conditional Contraction (Hypothesis in Lyapunov Theorem)

Claim. Each update removes a fixed fraction of the residual in expectation.

Status. Currently an assumption (empirically supported by semi-log linearity). A device-level proof would require an explicit Markov kernel and a uniform spectral gap; feasible but model-specific.

Appendix B.8.20. One-Step Conditional Contraction (Hypothesis in Lyapunov Theorem)

Claim. Each update removes a fixed fraction of the residual in expectation.

Status. Currently an assumption (empirically supported by semi-log linearity). A device-level proof would require an explicit Markov kernel and a uniform spectral gap; feasible but model-specific.

Lemma B.2 (Rate improves as angle opens). *Let $\mathcal{R}_{\text{res}}^{(K)}(t) = \|p - \mathcal{I}s\|^2$ be the degree- K residual at scale t and let (U, Φ) be an admissible one-step update (intertwining and $\|\Phi\| \leq 1$). Assume the conditional one-step contraction hypothesis of Theorem A.9: there exists $\alpha(t) \in (0, 1]$ such that*

$$\mathbb{E}[\mathcal{R}_{\text{res}}^{(K)}(Us, \Phi p) \mid s, p] \leq (1 - \alpha(t)) \mathcal{R}_{\text{res}}^{(K)}(t). \quad (\text{A41})$$

Then there exist t -independent constants $0 < c_- \leq c_+ < \infty$ (depending only on the fixed weights a_I and B through bounded positive operators) such that

$$c_- (1 - \gamma(t)) \leq \alpha(t) \leq c_+ (1 - \gamma(t)), \quad \gamma(t) = \|P_{\text{GBS}}(t)P_{\text{deg } K}\| = \cos \theta_F(t). \quad (\text{A42})$$

In particular, $\alpha(t)$ increases as $\gamma(t)$ decreases and is maximized at the photon–number match $m_{tB} = 2K$.

Proof. Write the residual as $\mathcal{R}_{\text{res}}^{(K)}(t) = \|e\|^2$ with $e := p - \mathcal{I}s \in \mathcal{H}_K$. One step gives

$$e^+ := \Phi p - \mathcal{I}Us = \Phi(p - \mathcal{I}s) + \underbrace{\Phi \mathcal{I}s - \mathcal{I}Us}_{=0 \text{ by intertwining}} = \Phi e, \quad (\text{A43})$$

hence $\mathcal{R}_{\text{res}}^{(K)+} = \|e^+\|^2 = \|\Phi e\|^2$. Taking conditional expectation and comparing to $\|e\|^2$,

$$\alpha(t) = 1 - \frac{\mathbb{E} \|\Phi e\|^2}{\|e\|^2}. \quad (\text{A44})$$

Decompose $e = e_{\parallel} + e_{\perp}$ with respect to the blueprint/sampling pair,

$$e_{\parallel} := P_{\text{deg } K}e, \quad e_{\perp} := (I - P_{\text{deg } K})e, \quad (\text{A45})$$

and note that Φ acts through the sampling feature subspace (by admissibility) so that the “through–subspace” contraction applies:

$$\|\Phi e\| = \|P_{\text{GBS}}(t)\Phi P_{\text{deg } K}e_{\parallel} + P_{\text{GBS}}(t)\Phi e_{\perp}\| \leq \|\Phi\| \|P_{\text{GBS}}(t)P_{\text{deg } K}e_{\parallel}\| + \|\Phi\| \|P_{\text{GBS}}(t)e_{\perp}\|. \quad (\text{A46})$$

Using $\|\Phi\| \leq 1$ and the two–subspace bounds (Proposition A.7),

$$\|P_{\text{GBS}}(t)P_{\text{deg } K}e_{\parallel}\| \leq \gamma(t) \|e_{\parallel}\|, \quad \|P_{\text{GBS}}(t)e_{\perp}\| \leq \|e_{\perp}\|. \quad (\text{A47})$$

Squaring and taking expectation (conditionally on e) yields

$$\mathbb{E} \|\Phi e\|^2 \leq \gamma(t)^2 \|e_{\parallel}\|^2 + \|e_{\perp}\|^2 \leq (1 - (1 - \gamma(t)^2) \frac{\|e_{\parallel}\|^2}{\|e\|^2}) \|e\|^2. \quad (\text{A48})$$

By the two–subspace lower bound,

$$(1 - \gamma(t))(\|e_{\parallel}\|^2 + \|e_{\perp}\|^2) \leq \|e_{\parallel} + e_{\perp}\|^2 = \|e\|^2 \Rightarrow \frac{\|e_{\parallel}\|^2}{\|e\|^2} \geq \frac{1 - \gamma(t)}{2}, \quad (\text{A49})$$

where the constant $\frac{1}{2}$ is an absolute (non-optimized) choice; any fixed $c \in (0, 1)$ would do after renormalizing constants. Hence

$$\frac{\mathbb{E} \|\Phi e\|^2}{\|e\|^2} \leq 1 - \underbrace{\frac{1}{2}(1 - \gamma(t)^2)(1 - \gamma(t))}_{\asymp (1 - \gamma(t))}, \quad (\text{A50})$$

which implies an upper bound

$$\alpha(t) \geq c_- (1 - \gamma(t)), \quad (\text{A51})$$

for some $c_- > 0$ independent of t (absorbing the numerical constants and the fixed weights). A matching upper bound follows by reversing the roles of the projectors in the two–subspace inequalities:

$$\|P_{\text{GBS}}(t)P_{\text{deg } K}e_{\parallel}\| \geq \sqrt{1 - \gamma(t)^2} \|e_{\parallel}\| \Rightarrow \mathbb{E} \|\Phi e\|^2 \geq (1 - (1 - \gamma(t)^2) \frac{\|e_{\parallel}\|^2}{\|e\|^2}) \|e\|^2, \quad (\text{A52})$$

and using $\frac{\|e_{\parallel}\|^2}{\|e\|^2} \leq \frac{1}{1 - \gamma(t)}$ (another consequence of Proposition A.7) gives

$$\alpha(t) \leq c_+ (1 - \gamma(t)) \quad (\text{A53})$$

for some $c_+ < \infty$ independent of t . The monotonicity of $\alpha(t)$ in $\gamma(t)$ is immediate, and maximization at $m_{tB} = 2K$ follows since $\gamma(t)$ is minimized (best alignment) at the photon–match by the calibration optimality and angle–conditioning discussion. \square

Appendix B.9. Bottom Line

Rock solid now: DPI; calibration optimality $m_{tB} = 2K$; angle bounds; exact GBS-I variance; large- n GBS-P MSE; (ε, δ) inversions; Lyapunov envelope (conditional on a one-step contraction); composition stability; small-noise robustness (Davis–Kahan).

Provable with short add-ons: residual factorization $A C_{2K} - B$; monotone “angle \Rightarrow rate” linkage; CLT for GBS-P; certified Gram surrogate bound.

Backed by the literature: nonempty open regions with advantage and percentage-of-advantage computations (Andersen–Shan 2025).

Still to complete: a model-independent proof of the one-step contraction and a fully quantified “poly vs. exp” lower bound with explicit constants under clear ensembles.

References

1. Andersen, J.E.; Shan, S. Using Gaussian Boson Samplers to Approximate Gaussian Expectation Problems, 2025, [arXiv:quant-ph/2502.19336].
2. Andersen, J.E.; Shan, S. Estimating the Percentage of GBS Advantage in Gaussian Expectation Problems, 2025, [arXiv:quant-ph/2502.19362].

3. Andersen, J.E.; Shan, S. Using Gaussian Boson Samplers to Approximate Gaussian Expectation Problems, 2025, [arXiv:quant-ph/2502.19336].
4. Hamilton, C.S.; Kruse, R.; Sansoni, L.; Barkhofen, S.; Silberhorn, C.; Jex, I. Gaussian Boson Sampling. *Phys. Rev. Lett.* **2017**, *119*, 170501.
5. Kruse, R.; Hamilton, C.S.; Sansoni, L.; Barkhofen, S.; Silberhorn, C.; Jex, I. Detailed study of Gaussian Boson Sampling. *Phys. Rev. A* **2019**, *100*, 032326.
6. Aaronson, S.; Arkhipov, A. The Computational Complexity of Linear Optics. In Proceedings of the STOC '11, 2011, pp. 333–342.
7. Deshpande, A.; Mehta, A.; Vincent, T.; Quesada, N.; Hinsche, M.; Ioannou, M.; Madsen, L.; Lavoie, J.; Qi, H.; Eisert, J.; et al. Quantum computational advantage via high-dimensional Gaussian boson sampling. *Science Advances* **2022**, *8*, eabi7894. <https://doi.org/10.1126/sciadv.abi7894>.
8. Bulmer, J.F.F.; Bell, B.A.; Chadwick, R.S.; Jones, A.E.; Moise, D.; Rigazzi, A.; Thorbecke, J.; Haus, U.U.; Vaerenbergh, T.V.; Patel, R.B.; et al. The boundary for quantum advantage in Gaussian Boson Sampling. *Science Advances* **2022**, *8*, eabl9236. <https://doi.org/10.1126/sciadv.abl9236>.
9. Björklund, A.; Gupt, B.; Quesada, N. A faster hafnian formula for complex matrices and its benchmarking on a supercomputer. *Journal of Experimental Algorithmics* **2019**, *24*, 1–19. <https://doi.org/10.1145/3325111>.
10. Gupt, B.; Izaac, J.; Quesada, N. The Walrus: a library for hafnians, Hermite polynomials and Gaussian Boson Sampling. *J. Open Source Software* **2019**, *4*, 1705.
11. Zhong, H.S.; Wang, H.; Deng, Y.H.; Chen, M.C.; Peng, L.C.; Luo, Y.H.; Qin, J.; Wu, D.; Gong, S.Q.; Su, H.; et al. Quantum computational advantage using photons. *Science* **2020**, *370*, 1460–1463. <https://doi.org/10.1126/science.abe8770>.
12. Madsen, L.S.; Laudenbach, F.; Falamarzi Askarani, M.; Rortais, F.; Vincent, T.; Bulmer, J.F.F.; Miatto, F.M.; Neuhaus, L.; Helt, L.G.; Collins, M.J.; et al. Quantum computational advantage with a programmable photonic processor. *Nature* **2022**, *606*, 75–81. <https://doi.org/10.1038/s41586-022-04725-4>.
13. Quesada, N.; Helt, L.G.; Izaac, J.; Arrazola, J.M.; Shahrokhshahi, R.; Myers, C.R.; Sabapathy, K.K. Simulating realistic non-Gaussian state preparation. *Physical Review A* **2019**, *100*, 022341, [arXiv:quant-ph/1810.03706]. <https://doi.org/10.1103/PhysRevA.100.022341>.
14. Quesada, N.; Arrazola, J.M.; Killoran, N. Gaussian boson sampling using threshold detectors. *Physical Review A* **2018**, *98*, 062322, [arXiv:quant-ph/1807.01639]. <https://doi.org/10.1103/PhysRevA.98.062322>.
15. Huh, J.; Guerreschi, G.G.; Peropadre, B.; McClean, J.R.; Aspuru-Guzik, A. Boson sampling for molecular vibronic spectra. *Nature Photonics* **2015**, *9*, 615–620. <https://doi.org/10.1038/nphoton.2015.153>.
16. Huh, J.; Yung, M.H. Vibronic boson sampling: Generalized Gaussian boson sampling for molecular vibronic spectra at finite temperature. *Scientific Reports* **2017**, *7*, 7462. <https://doi.org/10.1038/s41598-017-07305-5>.
17. Banchi, L.; Fingerhuth, M.; Babej, T.; Ing, C.; Arrazola, J.M. Molecular docking with Gaussian Boson Sampling. *Science Advances* **2020**, *6*, eaax1950. <https://doi.org/10.1126/sciadv.aax1950>.
18. Arrazola, J.M.; Bromley, T.R. Using Gaussian boson sampling to find dense subgraphs. *Physical Review Letters* **2018**, *121*, 030503. <https://doi.org/10.1103/PhysRevLett.121.030503>.
19. Brádler, K.; Dallaire-Demers, P.L.; Rebstroff, P.; Su, D.; Weedbrook, C. Gaussian boson sampling for perfect matchings of arbitrary graphs. *Physical Review A* **2018**, *98*, 032310. <https://doi.org/10.1103/PhysRevA.98.032310>.
20. Brádler, K.; Friedland, S.; Izaac, J.; Killoran, N.; Su, D. Graph isomorphism and gaussian boson sampling. *Special Matrices* **2021**, *9*, 166–196. <https://doi.org/10.1515/spma-2021-0121>.
21. Schuld, M.; Brádler, K.; Israel, R.; Su, D.; Gupt, B. A quantum hardware-induced graph kernel based on GBS, 2019, [arXiv:quant-ph/1905.12646].
22. Jahangiri, S.; Arrazola, J.M.; Quesada, N.; Killoran, N. Point processes with Gaussian boson sampling. *Physical Review E* **2020**, *101*, 022134. <https://doi.org/10.1103/PhysRevE.101.022134>.
23. Sparrow, C.; Martín-López, E.; Maraviglia, N.; Neville, A.; Harrold, C.; Carolan, J.; Joglekar, Y.N.; Hashimoto, T.; Matsuda, N.; O'Brien, J.L.; et al. Simulating the vibrational quantum dynamics of molecules using photonics. *Nature* **2018**, *557*, 660–667. <https://doi.org/10.1038/s41586-018-0065-7>.
24. He, Y.; Ding, X.; Su, Z.E.; Huang, H.L.; Qin, J.; Wang, C.; Unsleber, S.; Chen, C.; Wang, H.; He, Y.M.; et al. Time-bin-encoded boson sampling with a single-photon device. *Physical Review Letters* **2017**, *118*, 190501. <https://doi.org/10.1103/PhysRevLett.118.190501>.
25. Motes, K.R.; Gilchrist, A.; Dowling, J.P.; Rohde, P.P. Scalable boson sampling with time-bin encoding using a loop-based architecture. *Physical Review Letters* **2014**, *113*, 120501. <https://doi.org/10.1103/PhysRevLett.113.120501>.

26. Wick, G.C. The evaluation of the collision matrix. *Physical Review* **1950**, *80*, 268–272. <https://doi.org/10.1103/PhysRev.80.268>.
27. Kocharovsky, V.V.; Kocharovsky, V.V.; Tarasov, S.V. The hafnian master theorem. *Linear Algebra and its Applications* **2022**, *651*, 144–161. <https://doi.org/10.1016/j.laa.2022.07.018>.
28. Barvinok, A. Approximating Permanents and Hafnians. *Discrete Analysis* **2017**, *2017*, 1–34. <https://doi.org/10.19086/da.1244>.
29. Barvinok, A. Polynomial time algorithms to approximate permanents and mixed discriminants within a simply exponential factor. *Random Structures & Algorithms* **1999**, *14*, 29–61. [https://doi.org/10.1002/\(SICI\)1098-2418\(199901\)14:1<29::AID-RSA3>3.0.CO;2-O](https://doi.org/10.1002/(SICI)1098-2418(199901)14:1<29::AID-RSA3>3.0.CO;2-O).
30. Robert, C.P.; Casella, G. *Monte Carlo Statistical Methods*; Springer, 1999. <https://doi.org/10.1007/978-1-4757-3071-5>.
31. Rubinstein, R.Y.; Kroese, D.P. *Simulation and the Monte Carlo Method*, 3 ed.; Wiley, 2016. <https://doi.org/10.1002/9781118631980>.
32. Owen, A.B. Monte Carlo Theory, Methods and Examples. <https://artowen.su.domains/mc/>, 2013. Accessed 2025-02-28.
33. Owen, A.B. Practical Quasi-Monte Carlo Integration. <https://artowen.su.domains/mc/practicalqmc.pdf>, 2023. Accessed 2025-02-28.
34. Madsen, L.S.; Laudenbach, F.; Askarani, M.F.; Rortais, F.; Vincent, T.; Bulmer, J.F.F.; Miatto, F.M.; Neuhaus, L.; Helt, L.G.; Collins, M.J.; et al. Quantum computational advantage with a programmable photonic processor. *Nature* **2022**, *606*, 75–81. <https://doi.org/10.1038/s41586-022-04725-x>.
35. Zhong, H.S.; Wang, H.; Deng, Y.H.; Chen, M.C.; Peng, L.C.; Luo, Y.H.; Qin, J.; Wu, D.; Ding, X.; Hu, Y.; et al. Quantum computational advantage using photons. *Science* **2020**, *370*, 1460–1463. <https://doi.org/10.1126/science.abe8770>.
36. Deutsch, F. *Best Approximation in Inner Product Spaces*; Springer, 2001.
37. Davis, C.; Kahan, W.M. The Rotation of Eigenvectors by a Perturbation. *SIAM Journal on Numerical Analysis* **1970**, *7*, 1–46. <https://doi.org/10.1137/0707001>.
38. Wedin, P.Å. Perturbation bounds in connection with singular value decomposition. *BIT Numerical Mathematics* **1972**, *12*, 99–111. <https://doi.org/10.1007/BF01931334>.
39. van der Vaart, A.W. *Asymptotic Statistics*; Cambridge Univ. Press, 1998.
40. Billingsley, P. *Probability and Measure*, 3rd ed.; Wiley, 1995.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.