Asymptotic distinguishability of Haar-averaged measurement models

Ludmiła Marcinkowska[1], Łukasz Pawela[1], Marcin Markiewicz[1,2], and Zbigniew Puchała[1]

[1]Institute of Theoretical and Applied Informatics, Polish Academy of Sciences, ul. Bałtycka 5, 44-100 Gliwice, Poland
[2]International Centre for Theory of Quantum Technologies, University of Gdansk, 80-309 Gdańsk, Poland

We study discrimination problems generated by the same basic Haar-random measurement mechanism at two observational levels. First, we derive an explicit expression for the type-II error in the task of discriminating a Haar-random measure-and-prepare channel from the identity channel $\mathbb{I}$, using a coherence-sensitive entangled tester. Second, after passing to the induced classical measurement records, we compare two random measurement models: one induced by a single collective unitary of the form $U^{\otimes(n_1+n_2)}$ with $U \in U(d)$, and another induced by independent local unitaries $U_1^{\otimes n_1} \otimes U_2^{\otimes n_2}$. For the associated Haar-averaged aggregate histogram laws, in which the block of origin of each count is not retained, we obtain closed-form formulas and quantify their discrepancy through the total variation distance. We derive asymptotic expressions in the fixed-N, large-d regime, the fixed-d, large-N regime, the sparse joint-scaling regime $N = o(\sqrt{d})$, and the critical scaling regime $N/\sqrt{d} \to c$. We also identify the block-resolved pair-of-histograms law, showing that the aggregate total variation distance is a coarse-grained lower bound on the distinguishability available when block labels are retained.

1 Introduction and motivation

Distinguishing quantum processes is a central problem in quantum information theory, with key applications in device verification, quantum communication, and process tomography. Formally, it can be cast as a task of quantum hypothesis testing, in which a decision must be made between two candidate quantum channels based on a limited number of channel uses and the resulting measurement outcomes (see [1],[2],[3],[4],[5],[6],[7]). A particularly well-studied special case concerns the discrimination of unitary channels, where the structural properties of unitary operations — such as the role of entanglement, the feasibility of perfect discrimination, and the use of adaptive or multi-shot strategies — have been analyzed in detail (see [8],[9],[10],[11],[12]). The discrimination of quantum measurements, which is particularly relevant in experimentally accessible scenarios, has also been widely investigated, including single-shot and multi-shot settings as well as unambiguous discrimination strategies (see [13],[14],[15],[16],[17]).

A powerful approach to the study of typical properties of quantum channels relies on averaging over the Haar measure on the unitary group. In this setting, the Weingarten calculus provides explicit formulas for integrals of polynomial functions of unitary matrices and has become a standard tool in the analysis of random measurement operators, random quantum channels, and quantum networks (see [18],[19],[20]). Random quantum channels generated from Haar-random constructions and random POVMs induced by random isometries have been extensively studied in recent years, see [21],[22],[23].

Rather than characterizing channels through their full Choi representations, we adopt a more operational perspective that focuses directly on the statistics of measurement outcomes they induce. This approach emphasizes experimentally accessible data and avoids the overhead associated with full process tomography. It also aligns naturally with practical verification in which only measurement statistics are available.

The two discrimination problems studied in this paper are organized around the same basic object: a Haar-random unitary followed by a projective measurement, or equivalently a Haar-random measure-and-prepare channel. The first problem uses a coherence-sensitive quantum tester and asks how well this channel family can be distinguished from the identity channel $\mathbb{I}$. This provides a baseline at the channel level, before any restriction to classical records is imposed.

The second problem keeps the same random-measurement mechanism but changes the available data. Instead of testing the channel against the identity, we look at the classical outcome statistics produced after the measurement and ask whether these statistics reveal the structure of the underlying Haar randomness. More precisely, we compare a model generated by a single Haar-random unitary acting collectively (as a tensor power) on a system composed of $N = n_1 + n_2$ subsystems with a model generated by two independent Haar-random unitaries acting collectively on n_1 and n_2 subsystems. While both models produce exchangeable outcome statistics, their Haar-averaged histogram distributions differ in a nontrivial combinatorial manner.

A useful way to understand the statistical structure of random quantum measurements is through an analogy with classical occupancy processes. When a fixed input state is measured repeatedly in a basis obtained from a Haar-random unitary, each outcome can be viewed as drawing a symbol from a probability vector given by the squared amplitudes of a random column of the unitary matrix.

Conditioned on the unitary, the measurement outcomes are independent and follow a multinomial distribution. Averaging over the Haar measure therefore induces a classical random process in which the probability vector itself is random and uniformly distributed on the probability simplex. As a consequence, the resulting statistics can be interpreted as a classical occupancy problem: distributing n balls (measurement shots) into d bins (basis outcomes).

In this picture, the most informative events are collisions, i.e. repeated occurrences of the same outcome. When the number of samples is small compared to the Hilbert-space dimension, collisions are rare and the outcome statistics are close to those generated by independent sampling from a large alphabet. In contrast, when the number of samples grows, collisions become more frequent and reveal correlations imposed by the underlying measurement model.

This observation provides a natural mechanism for distinguishing different random measurement structures. In particular, correlations between subsystems in a collective Haar-random unitary influence the probability of collision events across blocks of measurement outcomes. By contrast, when the measurement is generated by independent unitaries acting collectively on separate subsystems, such cross-block correlations are suppressed. As we show below, the distinguishability between these two models can therefore be traced directly to differences in their collision statistics.

Thus the first part fixes the channel-level benchmark, while the main technical part of the paper analyzes what remains visible after one passes to classical measurement statistics. In that setting, the central quantity is the total variation distance (TVD) between the Haar-averaged aggregate histogram laws. The TVD provides an operationally meaningful measure of statistical distinguishability for the block-unresolved experiment in which only the total histogram is observed. We also record the corresponding block-resolved law for the ordered pair of histograms, which gives the operational distinguishability when the block labels are retained. Our analysis shows that the distinguishability between collective and block-collective random measurements is governed by collision statistics, leading to four distinct asymptotic regimes depending on the scaling of the number of samples N and the Hilbert-space dimension d.

1.1 Contributions

In this work we provide a detailed analysis of discrimination problems for Haar-random quantum measurement models, moving from a channel-level test to the classical statistics induced by the same random measurement construction. Our main contributions are as follows:

1. We derive an explicit formula for the Haar-averaged type-II error probability in the problem of distinguishing a Haar-random measure-and-prepare channel from the identity channel using a

maximally entangled input and a two-outcome tester. This establishes a channel-level benchmark for the random measurement family, expressed as a unitary integral evaluated via Weingarten calculus in Theorem 1.

2. We show that Haar-averaged measurement statistics can be mapped to a classical occupancy problem, where outcome histograms follow a Dirichlet-induced distribution on the probability simplex. This provides a transparent probabilistic interpretation of quantum measurement statistics in terms of collisions.

3. For both the collective and block-collective random measurement models, we derive closed-form expressions for the Haar-averaged aggregate histogram laws, and show that the block model induces a nontrivial combinatorial deformation governed by constrained histogram decompositions.

4. We express the aggregate total variation distance TVD between the two models in terms of a combinatorial functional counting admissible histogram decompositions. This identifies collision events as the fundamental mechanism responsible for statistical distinguishability.

5. We derive the exact block-resolved pair-of-histograms law. This clarifies the operational meaning of the aggregate statistic: the total-histogram TVD is a coarse-grained lower bound on the TVD obtainable from block-resolved data, and in the fixed-d, large-N regime the block-resolved TVD tends to one.

6. We analyze the behavior of the aggregate TVD in several asymptotic regimes:

   - for fixed N and $d \to \infty$, we show that the distinguishability is governed by rare collision events and vanishes with the dimension,
   - for fixed d and $N \to \infty$, we derive a simplex-limit formula expressed as an explicit integral over the probability simplex,
   - in the sparse joint-scaling regime $N = o(\sqrt{d})$, we recover the same collision-driven asymptotic behaviour,
   - in the critical scaling regime $N/\sqrt{d} \to c$, we prove that the collision count converges to a Poisson random variable and derive an exact expression for the limiting aggregate TVD as an expectation over this distribution (Theorem 27), providing a complete description of the transition between sparse and dense sampling regimes.

2 Distinguishing a measurement channel from the identity channel $\mathbb{I}$

In this section we analyze the problem of distinguishing the identity channel $\mathbb{I}$ from a random quantum channel implemented as a measurement in a collectively rotated computational basis, using a fixed two-outcome tester and a maximally entangled input state. Our goal is to compute the Haar-averaged probability of a type II error and to express it in terms of a Haar integral over the unitary group. For each $U \in U(d)$ we define a measure-and-prepare channel $Q_U^{(N)}$ acting on $(\mathbb{C}^d)^{\otimes N}$, defined as follows. Let $U \in U(d)$ be a Haar-random unitary and let $\{|k\rangle\}_{k=0}^{d-1}$ denote the computational basis of $\mathbb{C}^d$. For a multi-index $\mathbf{k} = (k_0, \ldots, k_{N-1}) \in [d]^N$, define, in the Heisenberg picture,

$$\begin{aligned} E_{U,k} &:= U^\dagger |k\rangle\langle k| U, \\ E_{U,\mathbf{k}} &:= E_{U,k_0} \otimes \cdots \otimes E_{U,k_{N-1}}. \end{aligned} \tag{1}$$

The channel $Q_U^{(N)}$ is then given by

$$Q_U^{(N)} : \rho \mapsto \sum_{\mathbf{k} \in [d]^N} \left(\operatorname{Tr} E_{U,\mathbf{k}}\, \rho\right) |\mathbf{k}\rangle\langle\mathbf{k}|, \tag{2}$$

where $\{|\mathbf{k}\rangle\}$ denotes the standard product basis of $(\mathbb{C}^d)^{\otimes N}$. Equivalently

$$Q_U^{(N)}(\rho) = \Delta\left(U^{\otimes N}\,\rho\,U^{\dagger\otimes N}\right), \tag{3}$$

where

$$\Delta(\rho) := \sum_k \langle k|\,\rho\,|k\rangle\,|k\rangle\,\langle k|\,. \tag{4}$$

This channel corresponds to a measurement in the rotated computational basis followed by a classical encoding of the outcome.

We consider the hypothesis

$$\begin{aligned} H_0 &: X = \mathbb{I}, \\ H_1 &: X = Q_U^{(N)} \neq \mathbb{I}, \end{aligned} \tag{5}$$

and a bipartite system $\mathcal{H}_A \otimes \mathcal{H}_B \cong (\mathbb{C}^d)^{\otimes N} \otimes (\mathbb{C}^d)^{\otimes N}$. The unknown channel $X \in \{\mathbb{I}, Q_U^{(N)}\}$ acts on subsystem $\mathcal{H}_A$, while the identity channel acts on $\mathcal{H}_B$. The discrimination strategy consists of preparing a maximally entangled state between $\mathcal{H}_A$ and $\mathcal{H}_B$, $|\psi\rangle = \frac{1}{\sqrt{d^N}}\sum_{\mathbf{k}}|\mathbf{kk}\rangle$, followed by a joint measurement on the output state.

The two-outcome measurement $\{\Pi_0, \Pi_1\}$ is performed jointly on $\mathcal{H}_A \otimes \mathcal{H}_B$ and is defined as

$$\begin{aligned} \Pi_0 &= \frac{1}{d^N}\sum_{\mathbf{k},\mathbf{l}} |\mathbf{kk}\rangle\langle\mathbf{ll}| = |\psi\rangle\langle\psi|\,, \\ \Pi_1 &= \mathbf{1} - \Pi_0. \end{aligned} \tag{6}$$

Let us denote the resulting output states by

$$\rho_0 := |\psi\rangle\langle\psi|\,, \qquad \rho_U := (Q_U^{(N)} \otimes \mathbb{I})(|\psi\rangle\langle\psi|). \tag{7}$$

The measurement outcome corresponding to Π_0 is interpreted as accepting H_0, while the outcome corresponding to Π_1 is interpreted as accepting H_1. The type-I and type-II errors are then given by

$$\begin{aligned} p_I &= \mathrm{Tr}(\rho_0 \Pi_1), \\ p_{II} &= \mathrm{Tr}(\rho_U \Pi_0), \end{aligned} \tag{8}$$

where the type-I error corresponds to incorrectly identifying the channel as different from the identity when it is in fact the identity, and the type-II error corresponds to failing to detect a difference when the channel is $Q_U^{(N)}$.

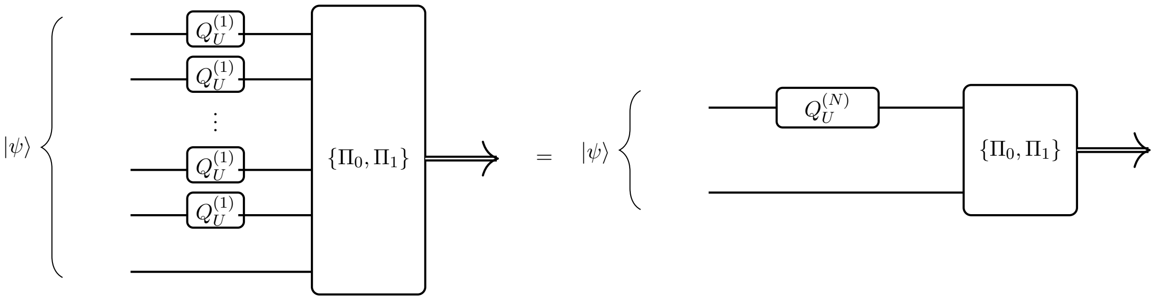

Figure 1: Discrimination protocol between the identity channel $\mathbb{I}$ and the Haar-random measurement channel $Q_U^{(N)}$. The maximally entangled state $|\psi\rangle$ is prepared across N subsystems. The channel $Q_U^{(N)}$ acts collectively on these subsystems, followed by the joint measurement $\{\Pi_0, \Pi_1\}$.

Theorem 1. *For the above discrimination strategy the following holds:*

1. *If $X = \mathbb{I}$, the outcome corresponding to Π_0 is obtained with probability* 1.

2. *If $X = Q_U^{(N)}$, then the Haar-averaged type–II error*

$$\overline{p}_{II} := \int p_{II}(U)dU \tag{9}$$

is given by

$$\overline{p}_{II} = \frac{1}{d^{2N}} \sum_{\sigma,\tau \in S_N} d^{c(\sigma\vee\tau)} \mathrm{Wg}(\sigma^{-1}\tau, d) \tag{10}$$

where Wg *denotes the Weingarten function and $c(\sigma \vee \tau)$ denotes the number of blocks in the join of the set partitions induced by the cycle decompositions of σ and τ.*

Proof. The first statement follows from the definition of Π_0 and $|\psi\rangle$;

$$\mathrm{Tr}\, |\psi\rangle\langle\psi|\, \Pi_0 = 1. \tag{11}$$

For $X = Q_U^{(N)}$, define the Haar integral

$$I_{N,d} := \sum_{\mathbf{k}} \int \left| \langle \mathbf{k}|\, U^{\otimes N}\, |\mathbf{k}\rangle \right|^2 dU. \tag{12}$$

Then

$$\overline{p}_{II} = \frac{1}{d^{2N}} I_{N,d}. \tag{13}$$

A straightforward computation using (2)-(7) shows that

$$I_{N,d} = \sum_{\mathbf{k}} \int \langle \mathbf{k}|\, E_{U,\mathbf{k}}\, |\mathbf{k}\rangle\, dU = \int \Big(\sum_{j=0}^{d-1} |U_{jj}|^2 \Big)^N dU. \tag{14}$$

Using Weingarten calculus, we obtain

$$I_{N,d} = \sum_{\sigma,\tau\in S_N} d^{c(\sigma\vee\tau)} \mathrm{Wg}(\sigma^{-1}\tau, d). \tag{15}$$

The reduction to the Haar integral is purely algebraic; details are provided in Appendix A.1. □

Proposition 2 (Large-d type-II asymptotics)**.** *Fix $N \geq 1$. As $d \to \infty$, with constants in the $O(\cdot)$ terms depending only on N, the following asymptotic expansion holds:*

$$I_{N,d} = 1 + \frac{\binom{N}{2}}{d} + O(d^{-2}), \tag{16}$$

$$\overline{p}_{II} = \frac{1}{d^{2N}} \left(1 + \frac{\binom{N}{2}}{d} + O(d^{-2}) \right) \xrightarrow{d\to\infty} 0. \tag{17}$$

Hence, for fixed N, the Haar-averaged type-II error vanishes asymptotically.

The proof is based on the Weingarten expansion and an analysis of the contributions of permutations in S_N. It is presented in Appendix A.3. We have thus derived an exact formula for the Haar-averaged type-II error in discriminating a Haar-random measure-and-prepare channel family $\{Q_U^{(N)}\}_{U\in U(d)}$ from the identity $\mathbb{I}$. The result reduces to a unitary integral that can be evaluated using Weingarten calculus. This first result uses the coherence of a maximally entangled input to detect the loss of coherence caused by the random measurement channel. We now turn to the complementary situation in which the observed data are already classical measurement outcomes. In that regime the identity channel is no longer the point of comparison; instead, the question is whether the outcome statistics retain enough information to distinguish a shared Haar-random measurement from independent blockwise Haar-random measurements.

3 Discrimination via Haar-averaged measurement statistics

In this section we study the classical statistical counterpart of the channel discrimination problem above. The same operation, a Haar-random unitary followed by a computational-basis measurement, now produces a classical outcome record. We compare two random measurement models acting on a composite system of $N = n_1 + n_2$ subsystems of local dimension d. The models differ in the structure of the underlying random channels: a single collective channel versus two independent block-collective channels. Our goal is to characterize and compare the Haar-averaged distributions of measurement outcomes induced by these models, and to quantify their aggregate distinguishability using the total variation distance. We also record how this statistic changes when block-resolved histograms are available.

3.1 Measurement channels

We fix the computational basis $\{|j\rangle\}_{j=0}^{d-1}$ of $\mathbb{C}^d$ and denote by

$$|\mathbf{0}\rangle := |0\rangle^{\otimes N} \tag{18}$$

the product input state on $\mathcal{H}^{\otimes N}$. Measurement outcomes are labeled by strings

$$\mathbf{k} = (k_0, \dots, k_{N-1}) \in \{0, \dots, d-1\}^N, \tag{19}$$

with the corresponding basis vectors

$$|\mathbf{k}\rangle := |k_0\rangle \otimes \cdots \otimes |k_{N-1}\rangle\,. \tag{20}$$

We consider measurement channels obtained by applying a random unitary, followed by a projective measurement in the computational basis. The classical post-processing is described by the dephasing map Δ

$$\Delta(\rho) := \sum_{\mathbf{k}} \langle\mathbf{k}|\,\rho\,|\mathbf{k}\rangle\, |\mathbf{k}\rangle\langle\mathbf{k}|\,, \tag{21}$$

which discards all off-diagonal terms and produces a classical outcome distribution.

Definition 3 (Collective unitary measurement channel)**.** *Let $N = n_1 + n_2$. For fixed unitary $U \in U(d)$, define*

$$Q_U^{(N)}(\rho) := \Delta\left(U^{\otimes N}\,\rho\,U^{\dagger\otimes N}\right), \tag{22}$$

and write

$$Q_U^{(N)}(\rho) = \sum_{\mathbf{k}} \mathrm{tr}\Big(E_{U,\mathbf{k}}^{(N)}\,\rho\Big)\, |\mathbf{k}\rangle\langle\mathbf{k}|\,, \quad \textit{where} \quad E_{U,\mathbf{k}}^{(N)} := U^{\dagger\otimes N}\,|\mathbf{k}\rangle\langle\mathbf{k}|\,U^{\otimes N}. \tag{23}$$

Averaging over the Haar measure gives

$$\overline{Q}^{(N)}(\rho) = \sum_{\mathbf{k}} \mathrm{tr}\left(\overline{E}_{\mathbf{k}}^{(N)}\rho\right) |\mathbf{k}\rangle\langle\mathbf{k}|\,, \qquad \overline{E}_{\mathbf{k}}^{(N)} := \int E_{U,\mathbf{k}}^{(N)} dU. \tag{24}$$

Definition 4 (Two-block-collective unitary measurement channel)**.** *For fixed $U, V \in U(d)$, define*

$$Q_{U,V}^{(n_1,n_2)}(\rho) := \Delta\left[\left(U^{\otimes n_1} \otimes V^{\otimes n_2}\right)\rho\left(U^{\dagger\otimes n_1} \otimes V^{\dagger\otimes n_2}\right)\right] \tag{25}$$

where U and V are independent Haar-distributed unitaries in $U(d)$. Equivalently,

$$Q_{U,V}^{(n_1,n_2)}(\rho) = \sum_{\mathbf{l},\mathbf{m}} \mathrm{tr}\Big(E_{U,\mathbf{l}}^{(n_1)} \otimes E_{V,\mathbf{m}}^{(n_2)}\,\rho\Big)\, |\mathbf{lm}\rangle\langle\mathbf{lm}|\,. \tag{26}$$

We write

$$\overline{Q}^{(N)} := \mathbb{E}_U\left[Q_U^{(N)}\right], \qquad \overline{Q}^{(n_1,n_2)} := \mathbb{E}_{U,V}\left[Q_{U,V}^{(n_1,n_2)}\right] \tag{27}$$

for the corresponding Haar-averaged channels. In the remainder of this section we analyze the histogram distributions induced by $Q_U^{(N)}$ and $Q_{U,V}^{(n_1,n_2)}$ on the input state $|\mathbf{0}\rangle$, and then average these laws over the Haar measure. Exploiting permutation invariance, we first derive an exact formula for the total variation distance between the resulting Haar-averaged aggregate histogram laws, where the observer retains only the total counts. We then compare this block-unresolved statistic with the block-resolved statistic given by the ordered pair of histograms.

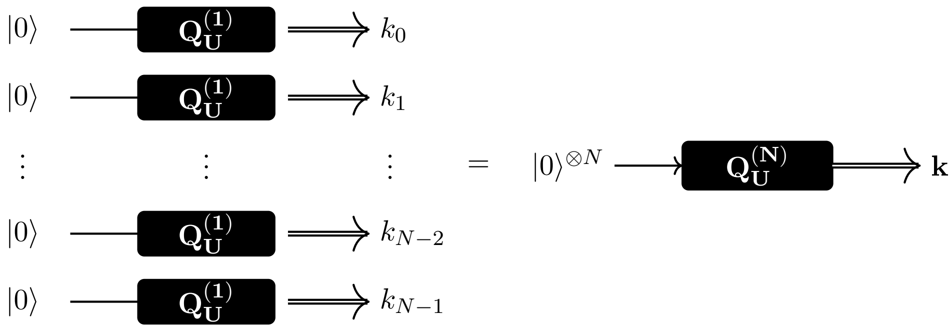

Figure 2: Equivalent representations of the collective model. The multipartite implementation using N copies of $Q_U^{(1)}$ is equivalent to the collective channel $Q_U^{(N)}$ acting on the product input state $|0\rangle^{\otimes N}$, producing the classical outcome vector $\mathbf{k} = (k_0, \ldots, k_{N-1})$.

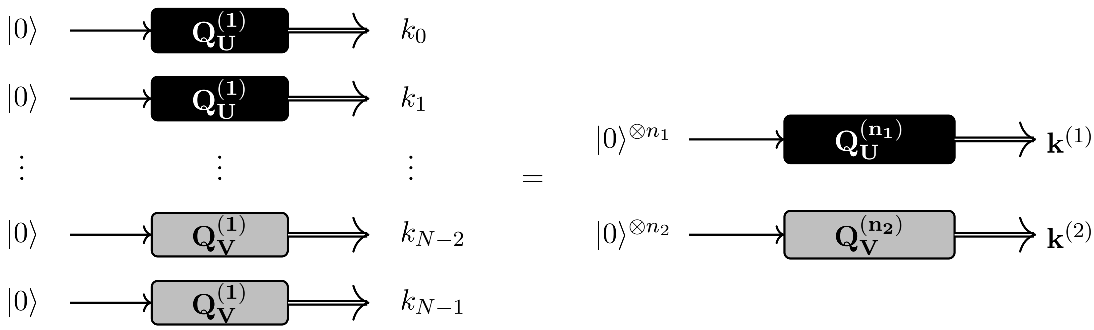

Figure 3: Equivalent representations of the block model. The multipartite implementation using n_1 copies of $Q_U^{(1)}$ and n_2 copies of $Q_V^{(1)}$ is equivalent to two collective channels $Q_U^{(n_1)}$ and $Q_V^{(n_2)}$ acting independently on the corresponding subsystems, producing the classical outcome vectors $\mathbf{k}^{(1)}$ and $\mathbf{k}^{(2)}$. In the aggregate experiment, the observer retains only the total histogram obtained after forgetting the block assignment of the outcomes.

3.2 Histogram reduction for a single random unitary

In this section, we show that the output statistics of the random measurement channel can be reduced to a distribution over histograms, which significantly simplifies the analysis.

Lemma 5 (Permutation invariance)**.** *Let $\mathbf{k}, \mathbf{k}' \in \{0, \ldots, d-1\}^N$ be two outcome strings such that $\mathbf{k}'$ is obtained from $\mathbf{k}$ by a permutation of subsystems. Then*

$$\operatorname{tr}\left(\overline{E}_{\mathbf{k}}^{(N)} |\mathbf{0}\rangle\langle\mathbf{0}|\right) = \operatorname{tr}\left(\overline{E}_{\mathbf{k}'}^{(N)} |\mathbf{0}\rangle\langle\mathbf{0}|\right). \tag{28}$$

Consequently, the probability of observing an outcome $\mathbf{k}$ depends only on the multiplicities of the symbols appearing in $\mathbf{k}$.

Proof. Let $\sigma \in S_N$ be a permutation such that $\mathbf{k}' = \sigma(\mathbf{k})$, and denote by P_σ the corresponding permutation operator on $\mathcal{H}^{\otimes N}$. Using the Haar invariance and the fact that $P_\sigma \,|\mathbf{0}\rangle = |\mathbf{0}\rangle$, we obtain

$$\overline{E}^{(N)}_{\mathbf{k}'} = P_\sigma \,\overline{E}^{(N)}_{\mathbf{k}}\, P_\sigma^\dagger, \tag{29}$$

which implies

$$\mathrm{tr}\Big(\overline{E}^{(N)}_{\mathbf{k}}\,|\mathbf{0}\rangle\langle\mathbf{0}|\Big) = \mathrm{tr}\Big(\overline{E}^{(N)}_{\mathbf{k}'}\,|\mathbf{0}\rangle\langle\mathbf{0}|\Big). \tag{30}$$

The detailed permutation-based argument is given in Appendix B.1. □

Motivated by Lemma 5, we introduce the histogram associated with an outcome string $\mathbf{k}$. We denote by Λ the random histogram of measurement outcomes obtained from the channel $Q_U^{(N)}$ acting on the input state $|\mathbf{0}\rangle$. Thus, Λ is a random variable taking values in the set of integer vectors $\lambda = (\lambda_0, \dots, \lambda_{d-1}) \in \mathbb{N}^d$ such that $\sum_{j=0}^{d-1} \lambda_j = N$. A particular realization of this random variable will be denoted by λ.

Definition 6 (Outcome histogram)**.** *Given an outcome string* $\mathbf{k} = (k_0, \dots, k_{N-1})$*, the associated histogram is the vector*

$$\begin{aligned} \lambda &= (\lambda_0, \dots, \lambda_{d-1}), \\ \lambda_j &:= \#\{r : k_r = j\}, \end{aligned} \tag{31}$$

satisfying $\sum_{j=0}^{d-1} \lambda_j = N$*.*

Lemma 7 (Conditional multinomial distribution)**.** *Conditioned on a fixed unitary* U*, the histogram* Λ *of outcomes produced by* $Q_U^{(N)}$ *acting on* $|\mathbf{0}\rangle$ *follows a multinomial distribution with parameters* (N, W)*, where*

$$W = (W_0, W_1, \dots, W_{d-1}), \quad \textit{and} \quad W_j := |\langle j|U|0\rangle|^2 \quad \textit{are the probabilities of observing outcome } j.$$

That is,

$$\mathbb{P}(\Lambda = \lambda \,|\, U) = \frac{N!}{\prod_{j=0}^{d-1} \lambda_j!} \prod_{j=0}^{d-1} W_j^{\lambda_j}. \tag{32}$$

Since the measurement outcomes are independent and identically distributed according to the probability vector W, the claim follows directly from standard multinomial statistics. A detailed derivation is provided in Appendix B.2.

Proposition 8 (Haar-averaged histogram distribution)**.** *Let* $\mathbb{P}(\Lambda = \lambda)$ *denote the Haar-averaged distribution of the histogram* Λ *obtained from* $Q_U^{(N)}$ *acting on* $|\mathbf{0}\rangle$*. Then* $\mathbb{P}(\Lambda = \lambda)$ *is uniform over all compositions of* N *into* d *parts:*

$$\mathbb{P}(\Lambda = \lambda) = \frac{(d-1)!N!}{(N+d-1)!}. \tag{33}$$

Proof. We compute the Haar average:

$$\mathbb{P}(\Lambda = \lambda) = m(\lambda)\mathbb{E}_U\left[\prod_{j=0}^{d-1} W_j^{\lambda_j}\right] = m(\lambda)\int \prod_{j=0}^{d-1} W_j^{\lambda_j}\, dU. \tag{34}$$

where $m(\lambda)$ is the number of distinct sequences $\mathbf{k}$ with counts specified by λ, i.e.

$$m(\lambda) := \binom{N}{\lambda_0, \dots, \lambda_{d-1}} = \frac{N!}{\prod_j \lambda_j!}. \tag{35}$$

For a fixed vector $\lambda = (\lambda_0, \ldots, \lambda_{d-1})$, denote

$$I(\lambda) := \int \prod_{j=0}^{d-1} W_j^{\lambda_j} dU. \tag{36}$$

For a Haar-distributed unitary U, the probability vector $W = (W_0, \ldots, W_{d-1})$ is uniformly distributed over the probability simplex. Thus, the vector W has the Dirichlet distribution $\mathrm{Dir}(1, \ldots, 1)$ on the simplex Δ_{d-1}, since under the Haar measure all probability configurations are equally likely. Here

$$\Delta_{d-1} = \{p \in \mathbb{R}^d : p_i \geq 0, \sum_{i=0}^{d-1} p_i = 1\}. \tag{37}$$

Therefore

$$\mathbb{P}(\Lambda = \lambda) = m(\lambda) I(\lambda) = \frac{(d-1)!N!}{(N+d-1)!}. \tag{38}$$

The explicit value of $I(\lambda)$ is calculated in Appendix B.3. □

Since the distribution is uniquely determined, we denote

$$P^{(N)}(\lambda) := \mathbb{P}(\Lambda = \lambda). \tag{39}$$

Lemma 9 (Blockwise Haar–averaged histogram distribution)**.** *Consider the random measurement channel $Q_{U,V}^{(n_1,n_2)}$ acting on the product input state*

$$|\mathbf{0}\rangle = |0\rangle^{\otimes n_1} \otimes |0\rangle^{\otimes n_2}. \tag{40}$$

Let $\lambda^{(1)}$ and $\lambda^{(2)}$ denote the outcome histograms associated with the first and second block, respectively. Then the following holds:

1. *Conditioned on fixed unitaries U and V, the outcomes within each block are independent and identically distributed, and the corresponding histograms follow multinomial distributions.*

2. *After averaging over the Haar measure, the joint distribution of $(\lambda^{(1)}, \lambda^{(2)})$ factorizes because the two Haar-distributed unitaries are independent:*

$$\mathbb{P}(\Lambda^{(1)} = \lambda^{(1)}, \Lambda^{(2)} = \lambda^{(2)}) = P^{(n_1)}(\lambda^{(1)}) \cdot P^{(n_2)}(\lambda^{(2)}) = \frac{n_1!n_2!(d-1)!(d-1)!}{(n_1+d-1)!(n_2+d-1)!}, \tag{41}$$

for all histograms satisfying $\sum_j \lambda_j^{(1)} = n_1$ and $\sum_j \lambda_j^{(2)} = n_2$.

In particular, all admissible pairs of histograms occur with equal probability.

Proof. The result follows from the permutation invariance of the Haar measure and the independence of the two unitary blocks. A detailed derivation, including the multinomial structure and the Haar integration leading to the Dirichlet distribution, is given in Appendix B.4. □

Proposition 10 (Histogram decompositions)**.** *Fix a histogram $\lambda = (\lambda_0, \ldots, \lambda_{d-1})$ with $\sum_{j=0}^{d-1} \lambda_j = N$. Let $L_\lambda(n_1, n_2)$ denote the number of vectors*

$$\lambda^{(1)} = (k_0, \ldots, k_{d-1}) \in \mathbb{N}^d \tag{42}$$

satisfying

$$\sum_{j=0}^{d-1} k_j = n_1, \quad \textit{and} \quad 0 \leq k_j \leq \lambda_j \textit{ for all } j. \tag{43}$$

Equivalently, $L_\lambda(n_1,n_2)$ counts the number of ways in which the histogram λ can be decomposed as $\lambda=\lambda^{(1)}+\lambda^{(2)}$ with $\sum_j \lambda_j^{(1)}=n_1$ and $\sum_j \lambda_j^{(2)}=n_2$. Then $L_\lambda(n_1,n_2)$ admits the inclusion–exclusion representation

$$L_\lambda(n_1,n_2)=\sum_{S\subseteq\{0,\dots,d-1\}}(-1)^{|S|}\binom{n_1-\sum_{j\in S}(\lambda_j+1)+d-1}{d-1}, \tag{44}$$

where $\binom{a}{b}=0$ for $a<b$.

The formula follows from a standard inclusion–exclusion argument for integer compositions with upper bounds. Details are given in Appendix B.5.

Proposition 11 (Effective aggregate histogram distribution)**.** *Let $P^{(n_1,n_2)}(\lambda)$ denote the Haar-averaged distribution of the total outcome histogram Λ of size $N=n_1+n_2$ obtained from $Q_{U,V}^{(n_1,n_2)}$ acting on the input state $|\mathbf{0}\rangle=|0\rangle^{\otimes n_1}\otimes|0\rangle^{\otimes n_2}$. This is the block-unresolved distribution obtained after the coarse-graining*

$$(\lambda^{(1)},\lambda^{(2)})\longmapsto\lambda=\lambda^{(1)}+\lambda^{(2)}. \tag{45}$$

Then

$$P^{(n_1,n_2)}(\lambda):=\frac{n_1!n_2!(d-1)!(d-1)!}{(n_1+d-1)!(n_2+d-1)!}L_\lambda(n_1,n_2), \tag{46}$$

where $L_\lambda(n_1,n_2)$ is defined in Proposition 10.

Proof. We can obtain the effective distribution $P^{(n_1,n_2)}(\lambda)$ for $\lambda=\lambda^{(1)}+\lambda^{(2)}$ by summing over all admissible pairs $(\lambda^{(1)},\lambda^{(2)})$ satisfying this constraint:

$$P^{(n_1,n_2)}(\lambda)=\sum_{\lambda^{(1)},\lambda^{(2)}}\mathbf{1}_{\{\lambda=\lambda^{(1)}+\lambda^{(2)}\}}\mathbb{P}(\Lambda^{(1)}=\lambda^{(1)},\Lambda^{(2)}=\lambda^{(2)}). \tag{47}$$

Since each pair contributes the same constant value, we obtain from Lemma 9

$$P^{(n_1,n_2)}(\lambda)=\frac{n_1!n_2!(d-1)!(d-1)!}{(n_1+d-1)!(n_2+d-1)!}L_\lambda(n_1,n_2). \tag{48}$$

This quantity counts the number of ways in which the total histogram λ can be decomposed into two component histograms $\lambda^{(1)}$ and $\lambda^{(2)}$ of sizes n_1 and n_2, respectively. □

3.3 Total variation distance

Corollary 12 (Aggregate total variation distance)**.** *We quantify the discrepancy between the single–block distribution $P^{(N)}(\lambda)$ and the block–separated aggregate distribution $P^{(n_1,n_2)}(\lambda)$ by the total variation distance*

$$\mathrm{TVD}\left(P^{(N)},P^{(n_1,n_2)}\right)=\frac{1}{2}\sum_\lambda|P^{(N)}(\lambda)-P^{(n_1,n_2)}(\lambda)|. \tag{49}$$

Let

$$M=\binom{N+d-1}{d-1},\qquad M_1=\binom{n_1+d-1}{d-1},\qquad M_2=\binom{n_2+d-1}{d-1}. \tag{50}$$

and define

$$\overline{M}=\frac{M_1M_2}{M}. \tag{51}$$

Using Proposition 11 and the fact that

$$P^{(N)}(\lambda)=\frac{1}{M},\qquad P^{(n_1,n_2)}(\lambda)=\frac{L_\lambda(n_1,n_2)}{M_1M_2},\qquad \frac{P^{(n_1,n_2)}(\lambda)}{P^{(N)}(\lambda)}=\frac{L_\lambda(n_1,n_2)}{\overline{M}} \tag{52}$$

we obtain

$$\mathrm{TVD}\left(P^{(N)},P^{(n_1,n_2)}\right)=\frac{1}{2}\mathbb{E}_{\lambda\sim P^{(N)}}\left[\left|1-\frac{L_\lambda(n_1,n_2)}{\overline{M}}\right|\right]. \tag{53}$$

Although both distributions admit explicit formulas, the evaluation of the aggregate total variation distance requires averaging the combinatorial quantity $L_\lambda(n_1, n_2)$ over all histograms. This motivates the asymptotic analysis carried out in the next section.

3.4 Block-resolved histograms and operational interpretation

The preceding aggregate total variation distance refers to the statistic

$$\lambda = \lambda^{(1)} + \lambda^{(2)}. \tag{54}$$

This is the natural statistic when the measurement record preserves outcome labels but not the block of origin of each occurrence. If the observer instead has access to the ordered pair of block histograms $(\lambda^{(1)}, \lambda^{(2)})$, then the relevant experiment is different and generally more informative.

Let

$$\mathcal{C}_{m,d} := \left\{ a \in \mathbb{N}^d : \sum_{j=0}^{d-1} a_j = m \right\}, \qquad |\mathcal{C}_{m,d}| = \binom{m+d-1}{d-1}. \tag{55}$$

For $a \in \mathcal{C}_{n_1,d}$ and $b \in \mathcal{C}_{n_2,d}$, write $\lambda = a + b$. In the shared-unitary model, corresponding to the collective channel split into two observed blocks, both histograms are sampled from the same Haar-random probability vector

$$W = (W_0, \dots, W_{d-1}) \sim \mathrm{Dirichlet}(1, \dots, 1). \tag{56}$$

Conditioned on W, the two block histograms are independent multinomials with parameters (n_1, W) and (n_2, W):

$$P(a, b \mid W) = \left(\frac{n_1!}{\prod_j a_j!} \prod_j W_j^{a_j} \right) \left(\frac{n_2!}{\prod_j b_j!} \prod_j W_j^{b_j} \right). \tag{57}$$

Hence

$$P(a, b \mid W) = \frac{n_1! n_2!}{\prod_j a_j! b_j!} \prod_{j=0}^{d-1} W_j^{a_j + b_j}. \tag{58}$$

Averaging over the Haar measure and using the Dirichlet integral from Appendix B.3, we obtain

$$\int_{\Delta_{d-1}} \prod_j W_j^{a_j+b_j} \, dU = (d-1)! \frac{\prod_j (a_j + b_j)!}{(N+d-1)!}. \tag{59}$$

Therefore

$$P_{\mathrm{sh}}(a, b) = \frac{n_1! n_2! (d-1)!}{(N+d-1)!} \prod_{j=0}^{d-1} \frac{(a_j + b_j)!}{a_j! b_j!}. \tag{60}$$

Equivalently,

$$P_{\mathrm{sh}}(a, b) = \frac{1}{M} \frac{\prod_j \binom{a_j + b_j}{a_j}}{\binom{N}{n_1}}. \tag{61}$$

In the independent-block model, the two Haar averages factorize. By the uniform histogram law derived earlier, each block histogram is uniformly distributed on its own composition set $P(a) = \frac{1}{M_1}$, $P(b) = \frac{1}{M_2}$. Since the two blocks are independent,

$$P_{\mathrm{ind}}(a, b) = \frac{1}{M_1 M_2}. \tag{62}$$

Thus the block-resolved total variation distance is

$$\mathrm{TVD}\left(P_{\mathrm{sh}}, P_{\mathrm{ind}}\right) = \frac{1}{2} \sum_{a,b} |P_{\mathrm{sh}}(a,b) - P_{\mathrm{ind}}(a,b)| = \frac{1}{2} \sum_{a \in \mathcal{C}_{n_1,d}} \sum_{b \in \mathcal{C}_{n_2,d}} \left| \frac{1}{M} \frac{\prod_j \binom{a_j+b_j}{a_j}}{\binom{N}{n_1}} - \frac{1}{M_1 M_2} \right|. \tag{63}$$

This distance is also the operational distance for full block-labeled outcome strings after Haar averaging: both models are exchangeable within each block, so conditional on the pair (a, b) all strings with those histograms have the same probability under both hypotheses, and the likelihood ratio depends only on (a, b). Since under the independent-block model the pair (a, b) is uniformly distributed on $\mathcal{C}_{n_1,d} \times \mathcal{C}_{n_2,d}$, we may rewrite the total variation distance as an expectation with respect to the uniform law.

$$\mathrm{TVD}\left(P_{\mathrm{sh}}, P_{\mathrm{ind}}\right) = \frac{1}{2}\,\mathbb{E}_{a,b\,\mathrm{uniform}}\left[\left|\frac{M_1 M_2}{M\binom{N}{n_1}}\prod_{j=0}^{d-1}\binom{a_j+b_j}{a_j} - 1\right|\right]. \tag{64}$$

The aggregate formulas above are recovered by applying the coarse-graining map $(a, b) \mapsto a+b$. Indeed, by Vandermonde's identity,

$$\sum_{a+b=\lambda}\prod_{j=0}^{d-1}\binom{\lambda_j}{a_j} = \binom{N}{n_1}, \tag{65}$$

and therefore

$$\sum_{a+b=\lambda} P_{\mathrm{sh}}(a, b) = \frac{1}{M} = P^{(N)}(\lambda), \qquad \sum_{a+b=\lambda} P_{\mathrm{ind}}(a, b) = \frac{L_\lambda(n_1, n_2)}{M_1 M_2} = P^{(n_1,n_2)}(\lambda). \tag{66}$$

The map $(a, b) \mapsto a + b = \lambda$ forgets the information about which block produced a given outcome count. In other words, it is a coarse-graining from the pair histogram to the aggregate histogram. Since total variation distance cannot increase under deterministic post-processing, we obtain

$$\mathrm{TVD}\left(P^{(N)}, P^{(n_1,n_2)}\right) \le \mathrm{TVD}\left(P_{\mathrm{sh}}, P_{\mathrm{ind}}\right). \tag{67}$$

Thus the total-histogram TVD studied in the sequel is a block-unresolved distinguishability, or equivalently a lower bound on the distinguishability available from block-resolved data.

The extra information in (a, b) can also be seen conditionally on the total histogram. Under the shared-unitary model,

$$P_{\mathrm{sh}}(a \mid \lambda) = \frac{\prod_j \binom{\lambda_j}{a_j}}{\binom{N}{n_1}}, \tag{68}$$

which is a multivariate hypergeometric law: the first block looks like a random subset of n_1 samples from the N aggregate samples. Under the independent-block model,

$$P_{\mathrm{ind}}(a \mid \lambda) = \frac{1}{L_\lambda(n_1, n_2)}. \tag{69}$$

Hence the two conditional laws coincide for collision-free histograms but differ once some $\lambda_j \ge 2$. For instance, if $\lambda_j = 2$, the shared-unitary model assigns relative weights $1, 2, 1$ to the local decompositions $(a_j, b_j) = (2, 0), (1, 1), (0, 2)$, while the independent-block model weights admissible decompositions uniformly. The pair statistic therefore detects whether a collision occurred inside a single block or across the two blocks, whereas the aggregate statistic only records that a collision occurred.

4 Asymptotics of TVD between $P^{(N)}$ and $P^{(n_1,n_2)}$

4.1 Overview of asymptotic regimes

In this section we investigate the asymptotic behaviour of the aggregate total variation distance between the Haar–averaged total-histogram laws $P^{(N)}(\lambda)$ and $P^{(n_1,n_2)}(\lambda)$ in the joint limit $N, d \to \infty$. Thus the statistic in this section is the block-unresolved histogram $\lambda = \lambda^{(1)} + \lambda^{(2)}$. Depending on the relative growth of the number of copies N and the Hilbert space dimension d, we distinguish several asymptotic regimes.

- for fixed N and $d \to \infty$, we show that the aggregate total variation distance vanishes asymptotically with leading term $\frac{n_1 n_2}{d}$.
- for fixed d and $N \to \infty$, we show that the aggregate total variation distance converges to a nontrivial simplex-limit depending on d and $\alpha = \frac{n_1}{N}$.
- In the sparse joint scaling regime $N = o(\sqrt{d})$, we recover the same leading collision-driven asymptotic $\frac{n_1 n_2}{d}$,
- in the critical scaling regime $N/\sqrt{d} \to c$, we obtain an exact limiting formula for the aggregate TVD in terms of a Poisson-distributed collision count, leading to a closed probabilistic expression for the distinguishability.

Each regime is analyzed separately in the subsections below. Depending on the relative scaling of N and d, the distinguishability exhibits qualitatively different asymptotic behaviours. The regimes analyzed in this section are summarized in the table below.

Statistic	**Regime**	**Limiting behaviour**	**Main result**
Aggregate TVD	fixed N, $d \to \infty$	$\frac{n_1 n_2}{d} + O(d^{-2})$	Proposition 13
Aggregate TVD	fixed d, $N \to \infty$	simplex integral	Proposition 15
Aggregate TVD	$N = o(\sqrt{d})$	$\frac{n_1 n_2}{d} + O\left(\frac{N^4}{d^2}\right)$	Proposition 20
Aggregate TVD	$N/\sqrt{d} \to c$	$\frac{1}{2}\mathbb{E}\left\|1 - e^{\alpha(1-\alpha)c^2}(1-\alpha(1-\alpha))^K\right\|$ $K \sim \text{Poisson}(\text{c}^2)$	Theorem 27
Block-resolved TVD	fixed N, $d \to \infty$	$\frac{n_1 n_2}{d} + O(d^{-2})$	Proposition 28
Block-resolved TVD	$N = o(\sqrt{d})$	$\frac{n_1 n_2}{d} + O\left(\frac{N^4}{d^2}\right)$	Proposition 28
Block-resolved TVD	$N/\sqrt{d} \to c$	$\frac{1}{2}\,\mathbb{E}\left[\left\|e^{-\alpha(1-\alpha)c^2} 2^C - 1\right\|\right]$ $C \sim \text{Poisson}(\alpha(1-\alpha)c^2)$	Theorem 29
Block-resolved TVD	fixed d, $N \to \infty$	tends to 1	Corollary 30

Throughout this section we use the representation of the aggregate total variation distance in terms of the histogram functional $L_\lambda(n_1, n_2)$ derived in the previous section. Recall that

$$\text{TVD}\left(P^{(N)}, P^{(n_1,n_2)}\right) = \frac{1}{2}\,\mathbb{E}_{\lambda \sim P^{(N)}}\left[\left|1 - \frac{L_\lambda(n_1,n_2)}{\overline{M}}\right|\right], \qquad \overline{M} = \frac{M_1 M_2}{M}, \tag{70}$$

where the expectation is taken with respect to the uniform distribution over stars–and–bars histograms and

$$M = \binom{N+d-1}{d-1}, \qquad M_1 = \binom{n_1+d-1}{d-1}, \qquad M_2 = \binom{n_2+d-1}{d-1}. \tag{71}$$

4.2 Fixed N, large d for aggregate model

We analyze the behavior of the aggregate total variation distance between the two Haar–averaged sampling models:

$$\text{TVD}\left(P^{(N)}, P^{(n_1,n_2)}\right),$$

in the limit of large local Hilbert-space dimension d, while N is fixed. Due to symmetry under Haar averaging, the probability of an N-shot aggregate measurement record depends only on its histogram $\lambda = (\lambda_0, \dots, \lambda_{d-1})$ of occupation numbers. We evaluate the aggregate TVD by separating histograms according to their collision structure, because collision events constitute the dominant source of distinguishability between the two models. We split histograms by the number of repeated occupations, distinguishing collision-free and single-collision cases:

$$S_0 = \{\lambda : \lambda_j \in \{0,1\},\ \sum_j \lambda_j = N\},$$
$$S_1 = \{\lambda : \text{exactly one}, \lambda_j = 2,\ \text{all others in } \{0,1\}\}. \tag{72}$$

Histograms with at least two collisions contribute only $O(d^{-2})$ to the aggregate TVD. Let

$$S_{\geq 2} = \{\lambda : \lambda \text{ contains at least two collision pairs}\}. \tag{73}$$

For fixed N, every additional collision reduces by one the number of freely chosen occupied bins. Hence

$$|S_{\geq 2}| = O(d^{N-2}). \tag{74}$$

Moreover, uniformly in λ,

$$P^{(N)}(\lambda) = O(d^{-N}), \qquad P^{(n_1,n_2)}(\lambda) = O(d^{-N}). \tag{75}$$

Therefore

$$\sum_{\lambda \in S_{\geq 2}} \left| P^{(N)}(\lambda) - P^{(n_1,n_2)}(\lambda) \right| = O(d^{N-2}) O(d^{-N}) = O(d^{-2}). \tag{76}$$

Proposition 13 (Fixed N, large d)**.** *Fix integers $N \geq 1$ and $n_1, n_2 \geq 0$ with $n_1 + n_2 = N$. As $d \to \infty$, we obtain*

$$\mathrm{TVD}\left(P^{(N)}, P^{(n_1,n_2)}\right) = \mathrm{TVD}_{S_0} + \mathrm{TVD}_{S_1} + O(d^{-2}) = \frac{n_1 n_2}{d} + O(d^{-2}) \tag{77}$$

Proof. We decompose the set of histograms according to the number of collisions. Collision-free histograms S_0 and single-collision histograms S_1 give the leading contributions, while all histograms outside $S_0 \cup S_1$ contribute only $O(d^{-2})$. The detailed combinatorial evaluation is given in Appendix C. □

Hence, the aggregate distinguishability between collective and block-separated Haar-random measurements decays inversely with d, and is proportional to the number of cross-block outcome pairs $n_1 n_2$, reflecting the suppression of global correlations in the large-dimension limit.

4.3 Fixed d, large N: simplex limit ($n_1 = \lfloor \alpha N \rfloor$ and $\alpha \in (0,1)$ fixed) for aggregate model

In this section we study the asymptotic behavior of the aggregate total variation distance in the regime where the local dimension d is fixed and the number of measurement outcomes N tends to infinity. We consider a splitting $n_1 = \lfloor \alpha N \rfloor$, $n_2 = N - n_1$ with $\alpha \in (0,1)$ fixed. As $N \to \infty$ with d fixed, the aggregate total variation distance admits a finite limit depending only on d and α. We denote this limit by

$$\mathrm{TVD}_\infty(d, \alpha) := \lim_{N \to \infty} \mathrm{TVD}\left(P^{(N)}, P^{(n_1,n_2)}\right). \tag{78}$$

Below we show that this limit exists and can be expressed explicitly. We then evaluate the general limiting formula for low-dimensional cases.

4.3.1 Asymptotic expansion of $L_\lambda(n_1, n_2)$

For fixed λ and large N, $L_\lambda(n_1, n_2)$ counts the lattice points in a truncated simplex.

$$\mathcal{S}(\lambda) = \{x \in \mathbb{R}^d_{\geq 0} : \sum_j x_j = n_1,\ x_j \leq \lambda_j\} \tag{79}$$

Let $\ell = \frac{\lambda}{N} \in \Delta_{d-1}$, $t = \frac{x}{n_1} \in \Delta_{d-1}$. Then $\mathcal{S}(\lambda)$ is n_1^{d-1} times the fixed polytope

$$\left\{ t \in \Delta_{d-1} : t_j \leq u_j(\ell) := \min\left(1, \frac{\ell_j}{\alpha}\right) \right\}. \tag{80}$$

It is convenient to express $L_\lambda(n_1, n_2)$ in terms of generating functions. Writing $\lambda = (\lambda_0, \ldots, \lambda_{d-1})$, we have

$$L_\lambda(n_1, n_2) = [x^{n_1}] \prod_{j=0}^{d-1} (1 + x + \cdots + x^{\lambda_j}). \tag{81}$$

This representation allows us to recover the inclusion–exclusion formula and provides an alternative route to the asymptotic analysis used below.

Lemma 14 (Truncated-simplex asymptotics)**.** *Fix $d \geq 2$ and $\alpha \in (0, 1)$. Let $N \to \infty$, $n_1 = \lfloor \alpha N \rfloor$, $n_2 = N - n_1$, and let $\lambda = \lambda(N)$ be any histogram with $\sum_j \lambda_j = N$ and $\ell = \lambda/N$. Then $L_\lambda(n_1, n_2)$ admits the asymptotic expansion*

$$L_\lambda(n_1, n_2) = \frac{n_1^{d-1}}{(d-1)!}\, \varphi_\alpha(\ell) + O(n_1^{d-2}), \tag{82}$$

uniformly in λ, where

$$\varphi_\alpha(\ell) = \sum_{S \subseteq [d]} (-1)^{|S|} \left(1 - \sum_{j \in S} \min\left\{1, \frac{\ell_j}{\alpha}\right\}\right)_+^{d-1} \tag{83}$$

is the normalized $(d-1)$-dimensional volume of the continuous polytope obtained by scaling x and λ. Its explicit inclusion–exclusion expression is derived below.

Proof. See Appendix D □

Proposition 15 (Fixed d, large N)**.** *Fix $d \geq 2$ and $\alpha \in (0, 1)$. Let $N \to \infty$, $n_1 = \lfloor \alpha N \rfloor$, and $n_2 = N - n_1$. Then*

$$\mathrm{TVD}\left(P^{(N)}, P^{(n_1,n_2)}\right) = \frac{1}{2} \mathbb{E}_{\ell \sim \mathrm{Unif}(\Delta_{d-1})} \left[\left|1 - \frac{\varphi_\alpha(\ell)}{(1-\alpha)^{d-1}}\right|\right] + o(1). \tag{84}$$

In particular, the above limit exists and is given by

$$\mathrm{TVD}_\infty(d, \alpha) = \frac{1}{2} \mathbb{E}_{\ell \sim \mathrm{Unif}(\Delta_{d-1})} \left[\left|1 - \frac{\varphi_\alpha(\ell)}{(1-\alpha)^{d-1}}\right|\right]. \tag{85}$$

Proof. To pass to the limit, we use the representation of the aggregate total variation distance as an average over histograms

$$\mathrm{TVD}\left(P^{(N)}, P^{(n_1,n_2)}\right) = \frac{1}{2} \frac{1}{M} \sum_\lambda \left|1 - \frac{L_\lambda(n_1, n_2)}{\overline{M}}\right|. \tag{86}$$

For fixed d, we also have

$$M \sim \frac{N^{d-1}}{(d-1)!} + O\left(N^{d-2}\right), \quad M_i \sim \frac{n_i^{d-1}}{(d-1)!} + O\left(N^{d-2}\right) \quad \text{for} \quad i = 1, 2. \tag{87}$$

Hence

$$\overline{M} = \frac{M_1 M_2}{M} = \frac{n_1^{d-1}}{(d-1)!} (1-\alpha)^{d-1} (1 + o(1)). \tag{88}$$

Since d is fixed, the inclusion-exclusion formula contains finitely many terms only. Moreover,

$$0 \leq \left(1 - \sum_{j \in S} \frac{\ell_j}{\alpha}\right)_+^{d-1} \leq 1, \tag{89}$$

uniformly in $\ell \in \Delta_{d-1}$. Hence the remainder term in the asymptotic expansion of $L_\lambda(n_1, n_2)$ is uniform over all histograms λ. Combining the asymptotic expansion of $L_\lambda(n_1, n_2)$ from Lemma 14 with the asymptotics of $\overline{M}$, we obtain

$$\frac{L_\lambda(n_1, n_2)}{\overline{M}} = \frac{\varphi_\alpha(\ell)}{(1-\alpha)^{d-1}} + o(1), \tag{90}$$

uniformly over all λ such that $\ell = \frac{\lambda}{N} \in \Delta_{d-1}$. Since the absolute value is a continuous function, it follows that

$$\left|1 - \frac{L_\lambda(n_1, n_2)}{\overline{M}}\right| = \left|1 - \frac{\varphi_\alpha(\ell)}{(1-\alpha)^{d-1}}\right| + o(1). \tag{91}$$

Therefore,

$$\frac{1}{M}\sum_\lambda \left|1 - \frac{L_\lambda(n_1, n_2)}{\overline{M}}\right| = \frac{1}{M}\sum_\lambda \left|1 - \frac{\varphi_\alpha(\ell)}{(1-\alpha)^{d-1}}\right| + o(1). \tag{92}$$

We now interpret the sum as a Riemann sum over the simplex. Histograms λ satisfying $\sum_j \lambda_j = N$ correspond to lattice points in the dilated simplex $N \cdot \Delta_{d-1}$. After normalization $\ell = \frac{\lambda}{N}$, these points form a grid in Δ_{d-1} with mesh size $\frac{1}{N}$, which becomes dense as $N \to \infty$. Since all histograms are equally weighted, the normalized counting measure on the lattice points converges to the uniform probability measure on Δ_{d-1}. Moreover, the function

$$\ell \mapsto \left|1 - \frac{\varphi_\alpha(\ell)}{(1-\alpha)^{d-1}}\right| \tag{93}$$

is bounded and piecewise continuous on the compact simplex Δ_{d-1}. Therefore, by the standard Riemann-sum approximation theorem,

$$\frac{1}{M}\sum_\lambda \left|1 - \frac{\varphi_\alpha(\ell)}{(1-\alpha)^{d-1}}\right| \longrightarrow \mathbb{E}_{\ell\sim \mathrm{Unif}(\Delta_{d-1})}\left[\left|1 - \frac{\varphi_\alpha(\ell)}{(1-\alpha)^{d-1}}\right|\right]. \tag{94}$$

Hence

$$\mathrm{TVD}\left(P^{(N)}, P^{(n_1,n_2)}\right) = \frac{1}{2}\mathbb{E}_{\ell\sim \mathrm{Unif}(\Delta_{d-1})}\left[\left|1 - \frac{\varphi_\alpha(\ell)}{(1-\alpha)^{d-1}}\right|\right] + o(1). \tag{95}$$

□

Remark 16. *The model is invariant under exchanging the two blocks of sizes n_1 and n_2. Hence the aggregate total variation distance depends only on the unordered pair $\{\alpha, 1-\alpha\}$:*

$$\mathrm{TVD}_\infty(d, \alpha) = \mathrm{TVD}_\infty(d, 1-\alpha) \tag{96}$$

4.3.2 Low-dimensional examples

Case $d = 2$

Theorem 17. *Fix $\alpha \in (0, 1)$. In the fixed-d, large-N regime with $d = 2$, a normalized histogram can be parametrized as $\ell = (t, 1-t)$ with $t \in [0, 1]$. Then*

$$\mathrm{TVD}_\infty(2, \alpha) = \alpha(1-\alpha) \tag{97}$$

Proof. In this case we obtain

$$\varphi_{2,\alpha}(t) = \max\left\{0, \min\left(1, \frac{t}{\alpha}\right) + \min\left(1, \frac{1-t}{\alpha}\right) - 1\right\}. \tag{98}$$

Plugging this into the limiting total variation expression yields the one-dimensional integral

$$\mathrm{TVD}_\infty(2, \alpha) = \frac{1}{2}\int_0^1 \left|1 - \frac{\varphi_{2,\alpha}(t)}{1-\alpha}\right| dt. \tag{99}$$

By symmetry, assume $\alpha \le \frac{1}{2}$. Then

$$\varphi_{2,\alpha}(t) = \begin{cases} \frac{t}{\alpha}, & 0 \le t \le \alpha, \\ 1, & \alpha \le t \le 1-\alpha, \\ \frac{1-t}{\alpha}, & 1-\alpha \le t \le 1. \end{cases} \tag{100}$$

Splitting the integral over the three intervals gives

$$\begin{aligned} \mathrm{TVD}_\infty(2,\alpha) &= \frac{1}{2}\left(\int_0^\alpha \left|1 - \frac{t}{\alpha(1-\alpha)}\right| dt + \int_\alpha^{1-\alpha} \frac{\alpha}{1-\alpha}\, dt + \int_{1-\alpha}^1 \left|1 - \frac{1-t}{\alpha(1-\alpha)}\right| dt\right) \\ &= \alpha(1-\alpha). \end{aligned} \tag{101}$$

□

To validate this expression, we compare it to exact finite-N computations for several values of α. Numerically one observes the convergence

- $\alpha = 0.5$: $N = 5000$ gives TVD $\approx 0.2498 \to 0.25 = 0.5 \cdot 0.5$.
- $\alpha = 0.25$: $N = 5000$ gives TVD $\approx 0.1874 \to 0.1875 = 0.25 \cdot 0.75$.
- $\alpha = 0.1$: $N = 5000$ gives TVD $\approx 0.09107 \to 0.09 = 0.1 \cdot 0.9$.

This confirms consistency between the exact discrete model and the asymptotic analysis.

Case $d = 3$ (triangle)

Proposition 18. *For $d = 3$ we parameterize a normalized histogram by $\ell = (x, y, z)$ with $x, y, z \ge 0$ and $x + y + z = 1$. Then*

$$\mathrm{TVD}_\infty(3,\alpha) = \frac{2}{(1-\alpha)^2} \int_{\Delta_2} \left(\varphi_{3,\alpha}(x,y) - (1-\alpha)^2\right)_+ dx\, dy. \tag{102}$$

Proof. Using coordinates (x, y) on the standard triangle $\Delta_2 = \{(x, y) : x \ge 0,\ y \ge 0,\ x + y \le 1\}$ and setting $z = 1 - x - y$, the appearance of the $(\cdot)_+$ operator follows from the identity $1 - \min(1, u) = (1 - u)_+$, so the truncation implemented by the min in the Ehrhart counting formula becomes the positive part in the piecewise-polynomial limit. This allows us to write all terms in inclusion–exclusion in a uniform form

$$\varphi_{3,\alpha}(x,y,z) = \sum_{S \subseteq \{1,2,3\}} (-1)^{|S|} \Big(1 - \tfrac{1}{\alpha} \sum_{j \in S} \ell_j\Big)_+^2. \tag{103}$$

Formally the sum runs over all subsets $S \subseteq \{1, 2, 3\}$, including $S = \{1, 2, 3\}$. The three-element term equals $(1 - \frac{1}{\alpha})_+^2 = 0$ for all $\alpha \in (0, 1)$ because $\frac{1}{\alpha} > 1$, hence it is omitted in the compact form. Explicitly,

$$\begin{aligned} \varphi_{3,\alpha}(x,y,z) &= 1 - \sum_{j=1}^{3} \Big(1 - \frac{\ell_j}{\alpha}\Big)_+^2 + \sum_{1 \le i < j \le 3} \Big(1 - \frac{\ell_i + \ell_j}{\alpha}\Big)_+^2 \\ &= 1 - \Big(1 - \frac{x}{\alpha}\Big)_+^2 - \Big(1 - \frac{y}{\alpha}\Big)_+^2 - \Big(1 - \frac{1-x-y}{\alpha}\Big)_+^2 + \\ &+ \Big(1 - \frac{x+y}{\alpha}\Big)_+^2 + \Big(1 - \frac{1-y}{\alpha}\Big)_+^2 + \Big(1 - \frac{1-x}{\alpha}\Big)_+^2. \end{aligned} \tag{104}$$

Since each $(\cdot)_+$ switches between zero and a linear expression at the thresholds $\ell_j = \alpha$ and $\ell_i + \ell_j = \alpha$, $\varphi_{3,\alpha}$ is a piecewise quadratic function on Δ_2.

The limiting aggregate total variation distance can be expressed as

$$\mathrm{TVD}_\infty(3,\alpha) = \frac{1}{2}\,\mathbb{E}_{\ell\sim\mathrm{Unif}(\Delta_2)}\Big[\big|1-\tfrac{\varphi_{3,\alpha}(\ell)}{(1-\alpha)^2}\big|\Big] = \frac{1}{2}\int_{\Delta_2}\left|1-\frac{\varphi_{3,\alpha}(x,y)}{(1-\alpha)^2}\right| 2\,dx\,dy. \tag{105}$$

Since $\mathbb{E}[\varphi_{3,\alpha}] = (1-\alpha)^2$, we can rewrite the absolute deviation as

$$\left|1-\frac{\varphi_{3,\alpha}}{(1-\alpha)^2}\right| = \frac{\big|(1-\alpha)^2-\varphi_{3,\alpha}\big|}{(1-\alpha)^2}. \tag{106}$$

For any random variable X with $\mathbb{E}[X] = m$ we have the identity

$$\mathbb{E}[|X-m|] = 2\,\mathbb{E}[(X-m)_+], \tag{107}$$

because positive and negative deviations from the mean have equal expectation. Applying this with $X = \varphi_{3,\alpha}$ and $m = (1-\alpha)^2$ gives

$$\mathrm{TVD}_\infty(3,\alpha) = \frac{1}{(1-\alpha)^2}\,\mathbb{E}[(\varphi_{3,\alpha}-(1-\alpha)^2)_+]. \tag{108}$$

Finally, using that the density of $\mathrm{Unif}(\Delta_2)$ is 2 on Δ_2 (whose area is $\frac{1}{2}$), we obtain

$$\mathrm{TVD}_\infty(3,\alpha) = \frac{2}{(1-\alpha)^2}\int_{\Delta_2}\big(\varphi_{3,\alpha}(x,y)-(1-\alpha)^2\big)_+\,dx\,dy. \tag{109}$$

□

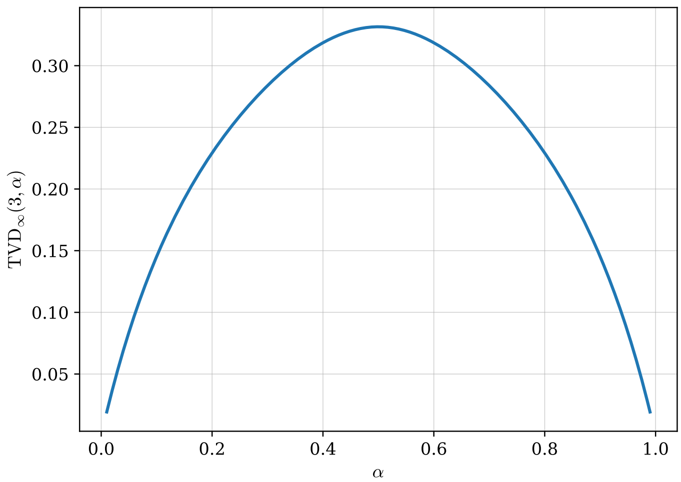

Figure 4: Limiting aggregate total variation distance between $P^{(N)}(\lambda)$ and $P^{(n_1,n_2)}(\lambda)$, i.e. $\mathrm{TVD}_\infty(3,\alpha)$, as a function of α.

The integral appearing in the definition of $\mathrm{TVD}_\infty(3,\alpha)$ does not admit a simple closed-form expression because the number of polyhedral regions grows rapidly. We illustrate this quantity numerically as a function of α. The integral was evaluated using a regular grid with $N = 500$, and the parameter α was sampled at 200 equally spaced values. Numerically one finds, for example,

$$\mathrm{TVD}_\infty\Big(3,\tfrac{1}{2}\Big) \approx 0.332, \qquad \mathrm{TVD}_\infty\Big(3,\tfrac{1}{3}\Big) \approx 0.298, \qquad \mathrm{TVD}_\infty(3,0.1) \approx 0.145.$$

Case $d = 4$ (tetrahedron)

Proposition 19. *For $d = 4$ we write a normalized histogram as $\ell = (x, y, z, w)$ with $x, y, z, w \geq 0$ and $x + y + z + w = 1$. Then*

$$\mathrm{TVD}_\infty(4, \alpha) = \frac{6}{(1-\alpha)^3} \int_{\Delta_3} \left(\varphi_{4,\alpha}(x, y, z) - (1-\alpha)^3\right)_+ dx\, dy\, dz. \tag{110}$$

Proof. Eliminating w by $w = 1 - x - y - z$ we work on the standard 3-simplex (tetrahedron)

$$\Delta_3 = \{(x, y, z) : x \geq 0,\ y \geq 0,\ z \geq 0,\ x + y + z \leq 1\}. \tag{111}$$

By the same inclusion–exclusion argument as before, the truncated-simplex volume admits the uniform description

$$\varphi_{4,\alpha}(\ell) = \sum_{S \subseteq \{1,2,3,4\}} (-1)^{|S|} \Big(1 - \frac{1}{\alpha} \sum_{j \in S} \ell_j\Big)_+^3, \tag{112}$$

i.e. each summand is $(1 - \frac{1}{\alpha} \sum_{j \in S} \ell_j)_+^{d-1}$ with $d - 1 = 3$. The full-set term $S = \{1, 2, 3, 4\}$ equals $(1 - \frac{1}{\alpha})_+^3 = 0$ for every $\alpha \in (0, 1)$ and is therefore omitted in the compact display. Explicitly the expansion contains singletons, pairs, triples and the empty set (the leading 1). Then

$$\varphi_{4,\alpha}(x, y, z, w) = 1 - \sum_{j=1}^{4} \Big(1 - \frac{\ell_j}{\alpha}\Big)_+^3 + \sum_{1 \leq i < j \leq 4} \Big(1 - \frac{\ell_i + \ell_j}{\alpha}\Big)_+^3 - \sum_{1 \leq i < j < k \leq 4} \Big(1 - \frac{\ell_i + \ell_j + \ell_k}{\alpha}\Big)_+^3 \tag{113}$$

The limiting aggregate total variation distance is given by

$$\begin{aligned} \mathrm{TVD}_\infty(4, \alpha) &= \frac{1}{2} \mathbb{E}_{\ell \sim \mathrm{Unif}(\Delta_3)}\Big[\Big|1 - \frac{\varphi_{4,\alpha}(\ell)}{(1-\alpha)^3}\Big|\Big] = \frac{1}{2(1-\alpha)^3} \mathbb{E}_{\ell \sim \mathrm{Unif}(\Delta_3)} \Big[\big|(1-\alpha)^3 - \varphi_{4,\alpha}(\ell)\big|\Big] \\ &= \frac{1}{(1-\alpha)^3} \mathbb{E}_{\ell \sim \mathrm{Unif}(\Delta_3)} \Big[\big(\varphi_{4,\alpha}(\ell) - (1-\alpha)^3\big)_+\Big]. \end{aligned} \tag{114}$$

Equivalently (writing the uniform expectation as an integral over the tetrahedron), since the uniform density on Δ_3 is constant and equals 6 (because $\mathrm{Vol}(\Delta_3) = \frac{1}{6}$), we have

$$\mathrm{TVD}_\infty(4, \alpha) = \frac{1}{2} \int_{\Delta_3} \left|1 - \frac{\varphi_{4,\alpha}(x, y, z)}{(1-\alpha)^3}\right| 6\, dx\, dy\, dz. \tag{115}$$

Replacing the expectation by the integral over T_3 with density 6 gives the computationally useful form

$$\mathrm{TVD}_\infty(4, \alpha) = \frac{6}{(1-\alpha)^3} \int_{\Delta_3} \left(\varphi_{4,\alpha}(x, y, z) - (1-\alpha)^3\right)_+ dx\, dy\, dz. \tag{116}$$

□

Since the integral appearing in the definition of $\mathrm{TVD}_\infty(4, \alpha)$ does not admit a simple closed-form expression, we compute this quantity numerically as a function of α. The integral over the three-dimensional simplex was approximated using a Monte Carlo integration scheme, and the parameter α was sampled at 100 equally spaced values.

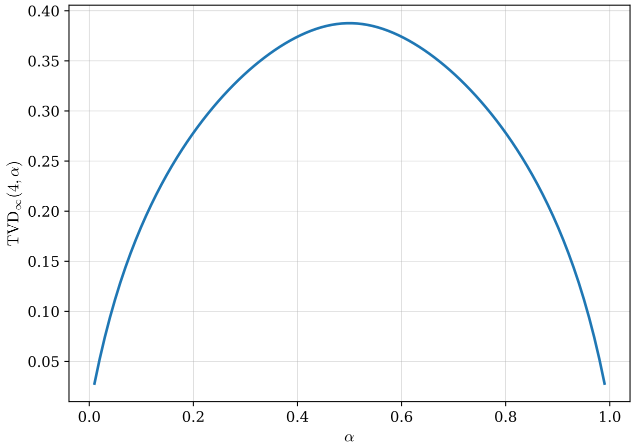

Figure 5: Limiting aggregate total variation distance between $P^{(N)}(\lambda)$ and $P^{(n_1,n_2)}(\lambda)$, i.e. $\mathrm{TVD}_\infty(4,\alpha)$, as a function of α.

Numerically:

$$\mathrm{TVD}_\infty\left(4,\tfrac{1}{2}\right) \approx 0.388, \qquad \mathrm{TVD}_\infty\left(4,\tfrac{1}{3}\right) \approx 0.352, \qquad \mathrm{TVD}_\infty(4,0.1) \approx 0.184.$$

4.3.3 Summary

In summary, for $d = 2,3,4$ we have derived a closed-form expression in the case $d = 2$ and integral representations together with numerical evaluations for $d = 3,4$. The piecewise-polynomial structure of $\varphi_{d,\alpha}$ leads to increasingly complex integrals as d grows. Our numerical results illustrate how $\mathrm{TVD}_\infty(d,\alpha)$ depends on the parameter α. We also observe that the plots of the aggregate total variation distance as a function of α are symmetric.

4.3.4 The symmetric split $\alpha = \frac{1}{2}$

We now focus on the case

$$n_1 = n_2 = \frac{N}{2}, \tag{117}$$

which plays a distinguished role in the fixed-d, large-N regime. In this case the two blocks have equal size, and the limiting aggregate total variation distance exhibits additional symmetry and structural simplifications. Recall that for fixed d the limiting aggregate total variation distance is given by

$$\mathrm{TVD}_\infty(d,\alpha) = \frac{1}{2}\,\mathbb{E}_{\ell\sim\mathrm{Unif}(\Delta_{d-1})}\left|1 - \frac{\varphi_\alpha(\ell)}{(1-\alpha)^{d-1}}\right|. \tag{118}$$

Setting $\alpha = \frac{1}{2}$ yields the universal representation

$$\mathrm{TVD}_\infty\left(d,\tfrac{1}{2}\right) = \frac{1}{2}\,\mathbb{E}_{\ell\sim\mathrm{Unif}(\Delta_{d-1})}\left|1 - 2^{d-1}\,\varphi_{\frac{1}{2}}(\ell)\right|, \tag{119}$$

where

$$\varphi_{\frac{1}{2}}(\ell) = \sum_{S\subseteq[d]}(-1)^{|S|}\Big(1 - \sum_{j\in S}\min\{1,2\ell_j\}\Big)_+^{d-1} = \sum_{S\subseteq[d]}(-1)^{|S|}\Big(1 - 2\sum_{j\in S}\ell_j\Big)_+^{d-1}. \tag{120}$$

Here the second equality follows from the fact that whenever $\ell_j > \frac{1}{2}$ for some $j \in S$, the corresponding term vanishes due to the positive-part truncation. We next observe that the mean value of $\varphi_{\frac{1}{2}}$ can be computed explicitly. Indeed, for general α one has

$$\mathbb{E}_{\ell\sim\mathrm{Unif}(\Delta_{d-1})}[\varphi_\alpha(\ell)] = (1-\alpha)^{d-1}, \tag{121}$$

which follows from the normalization of the truncated simplex volume. Specializing to $\alpha = \frac{1}{2}$ yields

$$\mathbb{E}_{\ell\sim\mathrm{Unif}(\Delta_{d-1})}[\varphi_{\frac{1}{2}}(\ell)] = 2^{-(d-1)}. \tag{122}$$

Using the identity

$$\mathbb{E}[\,|X - \mathbb{E}X|\,] = 2\,\mathbb{E}[(X - \mathbb{E}X)_+], \tag{123}$$

valid for any integrable random variable X, together with $\mathbb{E}[\varphi_{\frac{1}{2}}(\ell)] = 2^{-(d-1)}$, we obtain

$$\begin{aligned} \mathrm{TVD}_\infty\left(d, \tfrac{1}{2}\right) &= \frac{1}{2} 2^{d-1} \mathbb{E}_{\ell\sim\mathrm{Unif}(\Delta_{d-1})}\left[\left|\varphi_{\frac{1}{2}}(\ell) - 2^{-(d-1)}\right|\right] \\ &= \frac{1}{2} 2^{d-1} \mathbb{E}_{\ell\sim\mathrm{Unif}(\Delta_{d-1})}\left[\left|\varphi_{\frac{1}{2}}(\ell) - \mathbb{E}_{\ell\sim\mathrm{Unif}(\Delta_{d-1})}[\varphi_{\frac{1}{2}}(\ell)]\right|\right] = 2^{d-1} \mathbb{E}_{\ell\sim\mathrm{Unif}(\Delta_{d-1})}\left[\left(\varphi_{\frac{1}{2}}(\ell) - 2^{-(d-1)}\right)_+\right]. \end{aligned} \tag{124}$$

The values of $\mathrm{TVD}_\infty(d, \frac{1}{2})$ were computed using Monte Carlo sampling from the Dirichlet distribution. For each dimension d, we generate $n = 10^5$ independent samples. The aggregate total variation distance was then estimated as

$$2^{d-1} \mathbb{E}_{\ell\sim\mathrm{Unif}(\Delta_{d-1})}\left[\left(\varphi_{\frac{1}{2}}(\ell) - 2^{-(d-1)}\right)_+\right]. \tag{125}$$

Error bars represent 95% confidence intervals, computed as twice the standard error of the Monte Carlo estimator,

$$\pm \frac{2\sigma}{\left(\frac{1}{2}\right)^{d-1} \sqrt{n}} = \pm 2^d \frac{\sigma}{\sqrt{n}}, \tag{126}$$

where σ denotes the empirical standard deviation of the sample values.

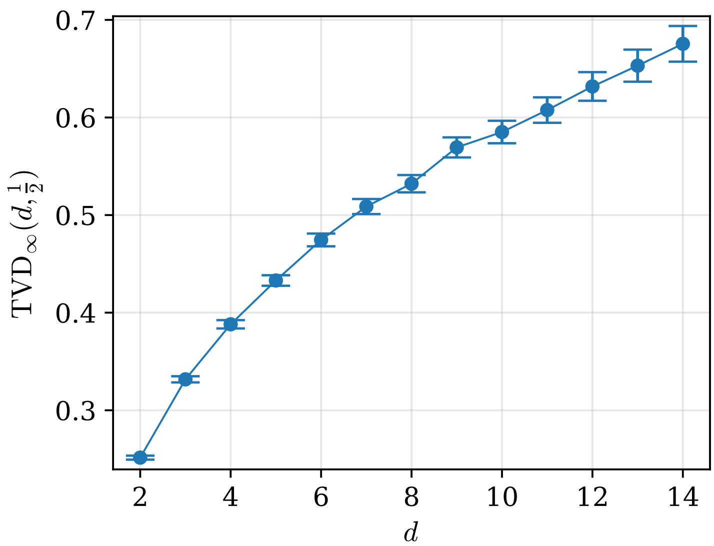

Figure 6: The values of $\mathrm{TVD}_\infty(d, \frac{1}{2})$ with error bars.

4.4 Joint scaling limits for aggregate model

4.4.1 (JS-sparse) $N = o(\sqrt{d})$

In the sparse joint-scaling regime $N = o(\sqrt{d})$, the total number of measurement outcomes grows sufficiently slowly compared to the local dimension d so that collisions of outcomes remain rare. As in the fixed-N, large-d analysis, the dominant contribution to the aggregate total variation distance arises from histograms with at most one collision. This allows us to reduce the problem to a finite stratification of the histogram space and to exploit the combinatorial structure developed in Section 4.2. More precisely, we partition the set of histograms according to their collision structure:

- S_0: no collisions (all counts 0 or 1),
- S_1: exactly one doubleton.

Proposition 20 (Sparse joint scaling)**.** *Let $N = N(d)$ satisfy $N = o(\sqrt{d})$, and let n_1, n_2 be integers with $n_1 + n_2 = N$. Then, as $d \to \infty$, we have,*

$$\mathrm{TVD}\left(P^{(N)}, P^{(n_1,n_2)}\right) = \frac{n_1 n_2}{d} + O\left(\frac{N^4}{d^2}\right). \tag{127}$$

Proof. The argument follows the same collision-based stratification as in Proposition 13. In the regime $N = o(\sqrt{d})$, the contribution of configurations with two or more collisions is negligible, and the aggregate TVD is captured by the strata S_0 and S_1. The leading asymptotics are obtained by combining explicit formulas for $L_\lambda(n_1, n_2)$ with expansions of binomial coefficients and of the normalization factor $\overline{M}$. Full details are provided in Appendix E. □

Remark 21. *In particular, since $N = o(\sqrt{d})$, we have*

$$\frac{N^4}{d^2} = o(1), \tag{128}$$

where the rate of convergence depends on the scaling of N relative to d. If $n_1/N \to \alpha$, we obtain

$$\mathrm{TVD}\left(P^{(N)}, P^{(n_1,n_2)}\right) = \frac{\alpha(1-\alpha)N^2}{d} + o(1). \tag{129}$$

Corollary 22. *Under the assumptions of Proposition 20, we have*

$$\mathrm{TVD}\left(P^{(N)}, P^{(n_1,n_2)}\right) \to 0 \quad \textit{as} \quad d \to \infty \tag{130}$$

This shows that in the sparse joint-scaling regime, the aggregate total variation distance is asymptotically governed by single-collision events, while configurations with multiple collisions contribute only at lower order.

4.4.2 Critical scaling regime $N/\sqrt{d} \to c$

We now investigate the critical joint-scaling regime in which the number of measurement outcomes grows proportionally to the square root of the local Hilbert space dimension, $N/\sqrt{d} \to c$, with a fixed constant $c > 0$. As before, we consider a block decomposition with $n_1/N \to \alpha$ and $n_2/N \to 1 - \alpha$, where $\alpha \in (0, 1)$ is independent of d. Let $\lambda = (\lambda_0, \ldots, \lambda_{d-1})$ be a random histogram of size N, drawn uniformly from all weak compositions of N into d parts. Define the collision count by

$$C(\lambda) := \sum_{i=1}^{d} \binom{\lambda_i}{2}. \tag{131}$$

Proposition 23 (Poisson limit for the collision count)**.** *Let N be an integer sequence such that $N/\sqrt{d} \to c$ for some $c > 0$. Then the collision count $C(\lambda)$ converges in distribution to a Poisson random variable K with mean c^2:*

$$C(\lambda) \xrightarrow[d\to\infty]{\mathrm{d}} K, \qquad K \sim \mathrm{Poisson}(c^2). \tag{132}$$

Equivalently, for each fixed $k \geq 0$,

$$\mathbb{P}(C(\lambda) = k) \longrightarrow e^{-c^2} \frac{(c^2)^k}{k!}. \tag{133}$$

Proof. This type of Poisson approximation is closely related to classical occupancy problems and Chen-Stein techniques, see [24]. The proof is given in Appendix F.1. □

Proposition 24 (Decomposition count for finitely many doubletons)**.** *Fix $k \geq 0$. Let λ be a histogram containing k bins of multiplicity 2 and no higher multiplicities. If $n_1/N \to \alpha \in (0,1)$, then*

$$\frac{L_\lambda(n_1, n_2)}{\binom{N}{n_1}} \longrightarrow (1-\alpha(1-\alpha))^k. \tag{134}$$

Proof. See Appendix F.2. □

Lemma 25 (Critical normalization)**.** *Let $\overline{M} = \frac{M_1 M_2}{M}$. If $N = N(d)$ satisfies $N/\sqrt{d} \to c > 0$ and $n_1/N \to \alpha \in (0,1)$, then*

$$\frac{\binom{N}{n_1}}{\overline{M}} \sim e^{\alpha(1-\alpha)c^2}. \tag{135}$$

Proof. The derivation is given in Appendix F.3. □

Lemma 26 (Critical aggregate domination)**.** *Under the assumptions of Lemma 25, the likelihood ratios*

$$R_d(\lambda) = \frac{L_\lambda(n_1, n_2)}{\overline{M}} \tag{136}$$

are uniformly bounded for all sufficiently large d. Consequently, $\{|1 - R_d(\lambda)|\}_d$ is uniformly integrable under $\lambda \sim P^{(N)}$.

Proof. Every admissible decomposition of a histogram is determined by choosing which n_1 of the N occurrences belong to the first block. Hence $L_\lambda(n_1, n_2) \leq \binom{N}{n_1}$ for every λ. Lemma 25 shows that $\binom{N}{n_1}/\overline{M}$ converges to $e^{\alpha(1-\alpha)c^2}$, and is therefore bounded for large d. □

Intuitively, in the critical regime the histogram consists mostly of singletons together with a Poisson number of double collisions. Each collision modifies the number of admissible histogram decompositions by a factor $(1 - \alpha(1 - \alpha))$, leading to the expectation formula above.

Theorem 27. *Let N satisfy $N/\sqrt{d} \to c > 0$, and let $n_1/N \to \alpha \in (0,1)$ with $n_2 = N - n_1$. Then the aggregate total variation distance converges to*

$$\lim_{d\to\infty} \mathrm{TVD}\left(P^{(N)}, P^{(n_1,n_2)}\right) = \frac{1}{2}\mathbb{E}\left|1 - e^{\alpha(1-\alpha)c^2}(1-\alpha(1-\alpha))^K\right|, \tag{137}$$

where

$$K \sim \mathrm{Poisson}(c^2). \tag{138}$$

Proof. Combining Propositions 23-24 with Lemma 25 reduces the aggregate TVD expression to an expectation over the collision count K. The details are given in Appendix F.4. □

4.5 Block-resolved asymptotics

In this subsection we briefly compare the asymptotic behaviour of the block-resolved total variation distance with the aggregate statistic analyzed above. As shown in Section 3.4, the block-resolved experiment retains additional information about the location of collisions across the two blocks. Consequently, the corresponding TVD is generally larger than the aggregate TVD.

Proposition 28 (Sparse block-resolved asymptotics)**.** *The leading sparse signal is unchanged. Fix integers $N \geq 1$ and $n_1, n_2 \geq 0$ with $n_1 + n_2 = N$. As $d \to \infty$,*

$$\mathrm{TVD}\left(P_{\mathrm{sh}}, P_{\mathrm{ind}}\right) = \frac{n_1 n_2}{d} + O(d^{-2}), \tag{139}$$

with constants depending only on N. More generally, if $N = o(\sqrt{d})$ and n_1, $n_2 = N - n_1$, then, uniformly over the block split,

$$\mathrm{TVD}\left(P_{\mathrm{sh}}, P_{\mathrm{ind}}\right) = \frac{n_1 n_2}{d} + O\left(\frac{N^4}{d^2}\right). \tag{140}$$

The argument follows the same collision-sector decomposition as in the aggregate case. The leading contribution comes from simple cross-block collisions. Full details are provided in Appendix G.1.

Theorem 29. *Let N satisfy $N/\sqrt{d} \to c > 0$, and let $n_1/N \to \alpha \in (0,1)$ with $n_2 = N - n_1$. The block-resolved limit is governed by the number of simple cross-block collisions. With*

$$\beta := \alpha(1-\alpha)c^2, \qquad C \sim \mathrm{Poisson}(\beta), \tag{141}$$

one obtains

$$\mathrm{TVD}\left(P_{\mathrm{sh}}, P_{\mathrm{ind}}\right) \longrightarrow \frac{1}{2}\,\mathbb{E}\left[\left|e^{-\beta}2^C - 1\right|\right]. \tag{142}$$

In the critical regime, the number of simple cross-block collisions converges to a Poisson random variable. Each such collision contributes a factor of two to the likelihood ratio, while all higher-order collision events are asymptotically negligible. Combining the Poisson approximation with the asymptotics of the normalization factor yields the stated limit. Full details are given in Appendix G.2.

Corollary 30. *For fixed $d \geq 2$ and $N \to \infty$, if $\frac{n_1}{N} \to \alpha \in (0,1)$ and $n_2 = N - n_1$, then*

$$\mathrm{TVD}\left(P_{\mathrm{sh}}, P_{\mathrm{ind}}\right) \longrightarrow 1. \tag{143}$$

The shared and independent models converge to mutually singular limiting distributions on the simplex product space. The claim then follows from standard properties of total variation distance under weak convergence. See Appendix G.3 for details. In the block-resolved experiment the two models become asymptotically perfectly distinguishable.

5 Conclusion

In this work we studied distinguishability questions generated by Haar-random measurement channels at two levels of observation. At the channel level, we considered a maximally entangled tester for distinguishing a Haar-random measure-and-prepare channel from the identity channel. This provides a coherence-sensitive benchmark for the random measurement family. At the classical level, we then asked what can be distinguished after the measurement has already produced outcome data. In that setting, the problem becomes a comparison of occupancy statistics and, in particular, of collision structures in random histograms.

For the comparison of the two Haar-averaged measurement models, we obtained exact formulas for the block-unresolved aggregate histogram laws, proved the large-d asymptotic

$$\mathrm{TVD}\left(P^{(N)}, P^{(n_1,n_2)}\right) = \frac{n_1 n_2}{d} + O(d^{-2}) \tag{144}$$

for fixed N, and derived simplex-limit formulas for fixed d and large N. We also showed that the same leading collision-driven asymptotic persists in the sparse joint-scaling regime $N = o(\sqrt{d})$.

We also clarified the operational role of the statistic being observed. If block labels are retained, the ordered pair of block histograms has an exact likelihood ratio governed by $\prod_j \binom{\lambda_j^{(1)}+\lambda_j^{(2)}}{\lambda_j^{(1)}}$, and its TVD dominates the aggregate TVD by data processing. In sparse large-d regimes the first-order asymptotics agree with the aggregate result, but for fixed d and large N the block-resolved TVD tends to one. Thus the nontrivial simplex limits obtained in this manuscript characterize precisely the block-unresolved experiment in which only the total histogram is available.

Our results illustrate how a quantum discrimination problem can be translated into a hierarchy of statistical questions. Coherent testers detect the measurement channel as a departure from the identity, while classical outcome statistics reveal the structure of the underlying Haar randomness only through histogram combinatorics. In the critical scaling regime $N/\sqrt{d} \to c$, we established that the number of collision pairs converges to a Poisson distribution with mean c^2. This leads to an explicit asymptotic formula for the aggregate total variation distance between the collective and block-separated models.

Acknowledgements

MM and ŁP acknowledges support from the National Science Center (NCN), Poland, under Project Opus No. 2024/53/B/ST2/02026. ZP and LM acknowledge support from the National Science Center (NCN), Poland, under Project Opus No. 2022/47/B/ST6/02380

References

[1] Anthony Chefles. "Quantum state discrimination". Contemporary Physics **41**, 401–424 (2000).

[2] Stephen M Barnett and Sarah Croke. "Quantum state discrimination". Advances in Optics and Photonics **1**, 238–278 (2009).

[3] Carl W Helstrom. "Quantum detection and estimation theory". Journal of statistical physics **1**, 231–252 (1969).

[4] Alexander S Holevo. "Statistical decision theory for quantum systems". Journal of multivariate analysis **3**, 337–394 (1973).

[5] Masahito Hayashi. "Discrimination of two channels by adaptive methods and its application to quantum system". IEEE Transactions on Information Theory **55**, 3807–3820 (2009).

[6] Quntao Zhuang and Stefano Pirandola. "Ultimate limits for multiple quantum channel discrimination". Phys. Rev. Lett. **125**, 080505 (2020).

[7] John Watrous. "The theory of quantum information". Cambridge University Press. (2018).

[8] Runyao Duan, Yuan Feng, and Mingsheng Ying. "Entanglement is not necessary for perfect discrimination between unitary operations". Phys. Rev. Lett. **98**, 100503 (2007).

[9] Runyao Duan, Yuan Feng, and Mingsheng Ying. "Perfect distinguishability of quantum operations". Phys. Rev. Lett. **103**, 210501 (2009).

[10] Mark Hillery, Erika Andersson, Stephen M. Barnett, and Daniel Oi. "Decision problems with quantum black boxes". Journal of Modern Optics **57**, 244–252 (2010).

[11] Akihito Soeda, Atsushi Shimbo, and Mio Murao. "Optimal quantum discrimination of single-qubit unitary gates between two candidates". Phys. Rev. A **104**, 022422 (2021).

[12] Yutaka Hashimoto, Akihito Soeda, and Mio Murao. "Comparison of unknown unitary channels with multiple uses" (2022). arXiv:2208.12519.

[13] Zbigniew Puchała, Łukasz Pawela, Aleksandra Krawiec, and Ryszard Kukulski. "Strategies for optimal single-shot discrimination of quantum measurements". Physical Review A **98**, 042103 (2018).

[14] Zbigniew Puchała, Łukasz Pawela, Aleksandra Krawiec, Ryszard Kukulski, and Michał Oszmaniec. "Multiple-shot and unambiguous discrimination of von neumann measurements". Quantum **5**, 425 (2021).

[15] Aleksandra Krawiec, Łukasz Pawela, and Zbigniew Puchała. "Discrimination and certification of unknown quantum measurements". Quantum **8**, 1269 (2024).

[16] Mario Ziman, Teiko Heinosaari, and Michal Sedlák. "Unambiguous comparison of quantum measurements". Physical Review A—Atomic, Molecular, and Optical Physics **80**, 052102 (2009).

[17] Michal Sedlák and Mário Ziman. "Optimal single-shot strategies for discrimination of quantum measurements". Physical Review A **90**, 052312 (2014).

[18] Zbigniew Puchała and Jarosław Adam Miszczak. "Symbolic integration with respect to the haar measure on the unitary group" (2011).

[19] Benoît Collins and Piotr Śniady. "Integration with respect to the haar measure on unitary, orthogonal and symplectic group". Communications in Mathematical Physics **264**, 773–795 (2006). url: https://arxiv.org/abs/math-ph/0402073v1.
[20] Giulio Chiribella, Giacomo Mauro D'Ariano, and Paolo Perinotti. "Theoretical framework for quantum networks". Physical Review A—Atomic, Molecular, and Optical Physics **80**, 022339 (2009).
[21] Teiko Heinosaari, Maria Anastasia Jivulescu, and Ion Nechita. "Random positive operator valued measures". Journal of Mathematical Physics**61** (2020).
[22] Ryszard Kukulski, Ion Nechita, Łukasz Pawela, Zbigniew Puchała, and Karol Życzkowski. "Generating random quantum channels". Journal of Mathematical Physics**62** (2021).
[23] Ion Nechita, Zbigniew Puchała, Łukasz Pawela, and Karol Życzkowski. "Almost all quantum channels are equidistant". Journal of Mathematical Physics **59**, 052201 (2018).
[24] Richard Arratia, Larry Goldstein, and Louis Gordon. "Two moments suffice for poisson approximations: the chen-stein method". The Annals of ProbabilityPages 9–25 (1989). url: https://scispace.com/pdf/two-moments-suffice-for-poisson-approximations-the-chen-2irpxe1jnb.pdf.

A Weingarten calculus

A.1 Evaluation of the integral $I_{n,d}$

We evaluate

$$\frac{1}{d^{2N}} I_{N,d} = \frac{1}{d^{2N}} \sum_{\mathbf{k}} \int |\langle \mathbf{k}| U^{\otimes N} |\mathbf{k}\rangle|^2 \, dU, \tag{145}$$

where the integration is with respect to the Haar measure on $U(d)$. For a fixed multi-index $\mathbf{k} = (k_0, \dots, k_{N-1})$ we have

$$\langle \mathbf{k}| U^{\otimes N} |\mathbf{k}\rangle = \prod_{r=0}^{N-1} U_{k_r k_r}, \qquad |\langle \mathbf{k}| U^{\otimes N} |\mathbf{k}\rangle|^2 = \prod_{r=0}^{N-1} |U_{k_r k_r}|^2. \tag{146}$$

Summing over all $\mathbf{k} \in [d]^N$ yields

$$\sum_{\mathbf{k}} \prod_{r=0}^{N-1} |U_{k_r k_r}|^2 = \Big(\sum_{j=0}^{d-1} |U_{jj}|^2 \Big)^N. \tag{147}$$

Thus we obtain

$$I_{N,d} = \int_{U(d)} \Big(\sum_{j=0}^{d-1} |U_{jj}|^2 \Big)^N dU. \tag{148}$$

To evaluate $I_{N,d}$ we use the general Weingarten formula [18, 19]. For arbitrary indices $(i_r), (j_r), (i'_r), (j'_r)$ of length N one has

$$\int_{U(d)} \prod_{r=0}^{N-1} U_{i_r j_r} \overline{U_{i'_r j'_r}} \, dU = \sum_{\sigma,\tau \in S_N} \Big(\prod_{r=0}^{N-1} \delta_{i_r, i'_{\sigma(r)}} \Big) \Big(\prod_{r=0}^{N-1} \delta_{j_r, j'_{\tau(r)}} \Big) \mathrm{Wg}(\sigma^{-1}\tau, d). \tag{149}$$

The following lemma provides a combinatorial interpretation of the delta constraints appearing in the Weingarten formula and will be used in the proof of the next result.

Lemma 31. *For every pair of permutations $\sigma, \tau \in S_N$ one has*

$$\sum_{\mathbf{k} \in [d]^N} \prod_{r=0}^{N-1} \delta_{k_r, k_{\sigma(r)}} \, \delta_{k_r, k_{\tau(r)}} = d^{c(\sigma \vee \tau)}, \tag{150}$$

where $c(\sigma \vee \tau)$ denotes the number of connected components of the graph obtained by joining $(r, \sigma(r))$ and $(r, \tau(r))$.

Proof. Consider a graph G with vertices $\{0,\dots,N-1\}$, in which for each r we add two edges:

$$r \sim \sigma(r), \qquad r \sim \tau(r). \tag{151}$$

The product

$$\prod_{r=0}^{N-1} \delta_{k_r,k_{\sigma(r)}}\, \delta_{k_r,k_{\tau(r)}} \tag{152}$$

equals 1 if and only if all indices k_r are equal within every connected component of G, and is 0 otherwise. If the graph G has $c(\sigma\vee\tau)$ connected components, then in the tuple $\mathbf{k} = (k_0,\dots,k_{N-1})$ one can freely choose $c(\sigma\vee\tau)$ values (one for each component), while the remaining indices are determined. Since each value can be chosen in d ways, the total number of admissible $\mathbf{k}$ equals $d^{c(\sigma\vee\tau)}$. □

Lemma 32. *Specializing the Weingarten formula (149) to the diagonal case $i_r = j_r = i'_r = j'_r = k_r$ yields*

$$I_{N,d} = \sum_{\sigma,\tau\in S_N} d^{c(\sigma\vee\tau)}\,\mathrm{Wg}(\sigma^{-1}\tau, d), \tag{153}$$

where $c(\sigma\vee\tau)$ denotes the number of blocks in the join of the cycle partitions of σ and τ.

Proof.

$$\begin{aligned} I_{N,d} &= \int_{U(d)} \Big(\sum_{j=0}^{d-1} |U_{jj}|^2\Big)^N dU = \int_{U(d)} \sum_{\mathbf{k}} \prod_{r=0}^{N-1} |U_{k_r k_r}|^2 dU = \int_{U(d)} \sum_{\mathbf{k}} \prod_{r=0}^{N-1} U_{k_r k_r}\overline{U_{k_r k_r}} dU \\ &= \sum_{\mathbf{k}} \sum_{\sigma,\tau\in S_N} \Big(\prod_{r=0}^{N-1} \delta_{k_r,k_{\sigma(r)}}\Big)\Big(\prod_{r=0}^{N-1} \delta_{k_r,k_{\tau(r)}}\Big)\mathrm{Wg}(\sigma^{-1}\tau,d) = \sum_{\sigma,\tau\in S_N} d^{c(\sigma\vee\tau)}\mathrm{Wg}(\sigma^{-1}\tau,d) \end{aligned} \tag{154}$$

□

A.2 Explicit evaluation for small N

Example 33 (Case $N=2$)**.** *For $N=2$ the symmetric group S_2 has two elements: the identity* id *and the transposition* (12)*. Using the permutation-sum representation*

$$I_{2,d} = \sum_{\sigma,\tau\in S_2} d^{c(\sigma\vee\tau)}\mathrm{Wg}(\sigma^{-1}\tau, d), \tag{155}$$

we list the four contributions and their combinatorial factors:

(σ,τ)	$\sigma^{-1}\tau$	$c(\sigma\vee\tau)$
$(\mathrm{id},\mathrm{id})$	id	2
$(\mathrm{id},(12))$	(12)	1
$((12),\mathrm{id})$	(12)	1
$((12),(12))$	id	1

(156)

Hence

$$I_{2,d} = (d^2+d)\,\mathrm{Wg}(\mathrm{id},d) + 2d\,\mathrm{Wg}((12),d). \tag{157}$$

The Weingarten values for $N=2$ are

$$\mathrm{Wg}(\mathrm{id},d) = \frac{1}{d^2-1}, \tag{158}$$

$$\mathrm{Wg}((12),d) = -\frac{1}{d(d^2-1)}. \tag{159}$$

Substituting these into the formula gives, for $d\geq 2$,

$$I_{2,d} = \frac{d+2}{d+1}. \tag{160}$$

Example 34 (Case $N = 3$)**.** *For $N = 3$ we start directly from the general formula*

$$I_{3,d} = \sum_{\sigma,\tau \in S_3} d^{\,c(\sigma \vee \tau)}\, \mathrm{Wg}(\sigma^{-1}\tau, d). \tag{161}$$

It is convenient to group the terms by the permutation $\pi = \sigma^{-1}\tau$. For a fixed π the pairs (σ, τ) with $\tau = \pi\sigma$ are in one-to-one correspondence with $\sigma \in S_3$, so that

$$I_{3,d} = \sum_{\pi \in S_3} \mathrm{Wg}(\pi, d) \left(\sum_{\sigma \in S_3} d^{c(\sigma \vee \sigma\pi)} \right). \tag{162}$$

There are three conjugacy classes in S_3:

- $\pi = \mathrm{id}$,
- π *a transposition,*
- π *a* 3*-cycle.*

We now compute the inner sums

$$S(\pi) := \sum_{\sigma \in S_3} d^{\,c(\sigma \vee \sigma\pi)}. \tag{163}$$

1. *For $\pi = \mathrm{id}$: If $\sigma = \mathrm{id}$ then $c(\sigma \vee \sigma\pi) = 3$ giving d^3. For the three transpositions $c(\sigma \vee \sigma\pi) = 2$, contributing $3d^2$. For the two* 3*-cycles $c(\sigma \vee \sigma\pi) = 1$, contributing $2d$. Thus*

$$S(\mathrm{id}) = d^3 + 3d^2 + 2d. \tag{164}$$

2. *For π a transposition: Among the six choices of $\sigma \in S_3$ we obtain two cases with $c(\sigma \vee \sigma\pi) = 2$ and four cases with $c(\sigma \vee \sigma\pi) = 1$, hence*

$$S(\textit{transposition}) = 2d^2 + 4d. \tag{165}$$

3. *For π a* 3*-cycle: Here all six values of σ yield $c(\sigma \vee \sigma\pi) = 1$, so*

$$S(\textit{3-cycle}) = 6d. \tag{166}$$

Collecting all contributions, we find

$$I_{3,d} = \mathrm{Wg}(\mathrm{id}, d)\,(d^3 + 3d^2 + 2d) + 3\,\mathrm{Wg}(\textit{transposition}, d)\,(2d^2 + 4d) + 2\,\mathrm{Wg}(\textit{3-cycle}, d)\,(6d). \tag{167}$$

Using the standard tabulated Weingarten expressions, for $d \geq 3$, we have

$$I_{3,d} = \frac{d^2 + 6d + 10}{(d+1)(d+2)}. \tag{168}$$

In the case of $d = 2$, we have

$$I_{3,2} = 2. \tag{169}$$

Indeed, for $U \in U(2)$ one has

$$|U_{00}|^2 + |U_{11}|^2 = 2|U_{00}|^2. \tag{170}$$

Since $|U_{00}|^2$ is uniformly distributed on $[0, 1]$, we obtain

$$I_{3,2} = \mathbb{E}[(2|U_{00}|^2)^3] = 8\mathbb{E}[(|U_{00}|^2)^3] = 2. \tag{171}$$

A.3 Asymptotic analysis for fixed N

We now analyze the behavior of $I_{N,d}$ for fixed N and $d \to \infty$. Recall that

$$I_{N,d} = \sum_{\sigma,\tau \in S_N} d^{c(\sigma \vee \tau)} \mathrm{Wg}(\sigma^{-1}\tau, d) = \sum_{\pi \in S_N} \mathrm{Wg}(\pi, d)\, S(\pi), \tag{172}$$

$$S(\pi) := \sum_{\sigma \in S_N} d^{\,c(\sigma \vee \sigma\pi)}. \tag{173}$$

From the general asymptotics of the Weingarten function [19] one has

$$\mathrm{Wg}(\pi, d) = d^{-N-|\pi|}(\mathrm{Moeb}(\pi) + O(d^{-2})), \tag{174}$$

where $|\pi| = N - c(\pi)$ is the length of the permutation, and $\mathrm{Moeb}(\pi)$ denotes the Möbius-type constant depending only on the cycle type of π (see [19] for the precise definition). On the other hand, $S(\pi)$ is a polynomial in d of degree $N - |\pi|$. Hence the leading order of the product is

$$\mathrm{Wg}(\pi, d)\, S(\pi) \;\sim\; d^{-2|\pi|}. \tag{175}$$

Therefore:

- For $\pi = \mathrm{id}$ ($|\pi| = 0$) one has

  $$S(\mathrm{id}) = \sum_{\sigma \in S_N} d^{c(\sigma)} = d^N + \binom{N}{2} d^{N-1} + O(d^{N-2}), \tag{176}$$

  and since

  $$\mathrm{Wg}(\mathrm{id}, d) = d^{-N}(1 + O(d^{-2})), \tag{177}$$

  their product contributes at orders 1, $\frac{1}{d}$, $\frac{1}{d^2}$.
- π a transposition ($|\pi| = 1$) contributes terms of order $\frac{1}{d^2}$,
- all other permutations ($|\pi| \geq 2$) contribute at most $O(d^{-4})$.

Thus the coefficient of d^{-1} can only come from the identity permutation $\pi = \mathrm{id}$; all nonidentity permutations contribute at order $O(d^{-2})$ or smaller.

For $\pi = \mathrm{id}$ we have

$$\begin{aligned} S(\mathrm{id}) &= \sum_{\sigma \in S_N} d^{c(\sigma \vee \sigma)} = \sum_{\sigma \in S_N} d^{c(\sigma)} = \sum_{k=1}^{N} \begin{bmatrix} N \\ k \end{bmatrix} d^k = d(d+1)\cdots(d+N-1) \\ &= d^N + \binom{N}{2} d^{N-1} + O(d^{N-2}). \end{aligned} \tag{178}$$

where $\left[\begin{smallmatrix} N \\ k \end{smallmatrix}\right]$ is the (unsigned) Stirling number of the first kind. These numbers count the permutations of N elements that consist of exactly k disjoint cycles, that is, $\left[\begin{smallmatrix} N \\ k \end{smallmatrix}\right]$ equals the number of $\sigma \in S_N$ such that $c(\sigma) = k$. Multiplying with $\mathrm{Wg}(\mathrm{id}, d) = d^{-N}(1 + O(d^{-2}))$ yields

$$\begin{aligned} \mathrm{Wg}(\mathrm{id}, d)\, S(\mathrm{id}) &= d^{-N}(1 + O(d^{-2})) \left(d^N + \binom{N}{2} d^{N-1} + O(d^{N-2}) \right) \\ &= 1 + \frac{\binom{N}{2}}{d} + O(d^{-2}). \end{aligned} \tag{179}$$

For π a transposition one has $|\pi| = 1$, so $S(\pi)$ is of degree at most $N-1$ and

$$\mathrm{Wg}(\pi, d)\, S(\pi) = O(d^{-2}). \tag{180}$$

Since N is fixed, summing over all nonidentity permutations contributes $O(d^{-2})$ in total. Therefore

$$I_{N,d} = 1 + \frac{\binom{N}{2}}{d} + O(d^{-2}), \tag{181}$$

so

$$I_{N,\infty} = 1. \tag{182}$$

This completes the proof of the asymptotic expansion.

B Haar-averaged histogram distributions

B.1 Proof of Lemma 5

Consider the state $\rho = |00\dots0\rangle\langle 00\dots0|$ and, for simplicity, denote $N = n_1 + n_2$.

$$\overline{Q}^{(N)}(\rho) = \sum_{\mathbf{k}} \mathrm{tr}\left(\overline{E}^{(N)}_{\mathbf{k}} |00\dots0\rangle\langle 00\dots0|\right) |\mathbf{k}\rangle\langle\mathbf{k}| \tag{183}$$

If $\mathbf{k} = \sigma(\mathbf{m})$, then

$$\begin{aligned}\overline{E}^{(N)}_{\mathbf{k}} &= \int U^{\dagger\otimes N} |\mathbf{k}\rangle\langle\mathbf{k}| U^{\otimes N} dU = \int U^{\dagger\otimes N} P_\sigma |\mathbf{m}\rangle\langle\mathbf{m}| P_\sigma^T U^{\otimes N} dU \\ &= \int P_\sigma P_\sigma^T U^{\dagger\otimes N} P_\sigma |\mathbf{m}\rangle\langle\mathbf{m}| P_\sigma^T U^{\otimes N} P_\sigma P_\sigma^T dU = \int P_\sigma U^{\dagger\otimes N} |\mathbf{m}\rangle\langle\mathbf{m}| U^{\otimes N} P_\sigma^T dU = P_\sigma \overline{E}^{(N)}_{\mathbf{m}} P_\sigma^T.\end{aligned} \tag{184}$$

Hence

$$\begin{aligned}&\mathrm{tr}\overline{E}^{(N)}_{\mathbf{k}} |00\dots0\rangle\langle 00\dots0| = \mathrm{tr} P_\sigma \overline{E}^{(N)}_{\mathbf{m}} P_\sigma^T |00\dots0\rangle\langle 00\dots0| = \mathrm{tr}\overline{E}^{(N)}_{\mathbf{m}} P_\sigma^T |00\dots0\rangle\langle 00\dots0| P_\sigma \\ &= \mathrm{tr}\overline{E}^{(N)}_{\mathbf{m}} |00\dots0\rangle\langle 00\dots0| .\end{aligned} \tag{185}$$

That is, the trace depends only on the composition of $\mathbf{k}$, not on its particular ordering. On the other hand,

$$\begin{aligned}&\mathrm{tr}\overline{E}^{(N)}_{\mathbf{k}} |00\dots0\rangle\langle 00\dots0| = \int \langle 00\dots0| U^{\dagger\otimes N} |\mathbf{k}\rangle\langle\mathbf{k}| U^{\otimes N} |00\dots0\rangle\, dU \\ &= \int \langle 0| U^\dagger |k_0\rangle \dots \langle 0| U^\dagger |k_{N-1}\rangle \langle k_0| U |0\rangle \dots \langle k_{N-1}| U |0\rangle\, dU = \int |\langle k_0| U |0\rangle|^2 \dots |\langle k_{N-1}| U |0\rangle|^2 dU \\ &= \int \prod_{r=0}^{N-1} |\langle k_r| U |0\rangle|^2 dU.\end{aligned} \tag{186}$$

B.2 Conditional multinomial distribution

For the block-unresolved experiment, the variable Λ provides a complete description of the observed aggregate statistics. For the random channel $Q_U^{(N)}$, the probability of obtaining a given outcome sequence depends only on its histogram, not on the ordering of outcomes; for the product channel $Q_{U,V}^{(n_1,n_2)}$, aggregate statistics are obtained by summing over all block-wise histograms $\lambda^{(1)}$ and $\lambda^{(2)}$ satisfying $\lambda^{(1)} + \lambda^{(2)} = \lambda$. Denote $W_{k_r} := |\langle k_r| U |0\rangle|^2$. Then

$$\prod_{r=0}^{N-1} |\langle k_r| U |0\rangle|^2 = \prod_{r=0}^{N-1} W_{k_r} = \prod_{j=0}^{d-1} W_j^{\lambda_j}. \tag{187}$$

Hence we obtain the conditional multinomial distribution

$$\mathbb{P}(\Lambda = \lambda | U) = m(\lambda)(W_0)^{\lambda_0}(W_1)^{\lambda_1}\dots(W_{d-1})^{\lambda_{d-1}} = m(\lambda)\prod_{j=0}^{d-1} W_j^{\lambda_j} \tag{188}$$

where $m(\lambda)$ is the multinomial coefficient, i.e.,

$$m(\lambda) = \frac{N!}{\lambda_0!\lambda_1! \dots \lambda_{d-1}!}. \tag{189}$$

B.3 Dirichlet integral and Haar averaging

The density of the Dirichlet distribution $\mathrm{Dir}(\alpha_0, \dots, \alpha_{d-1})$ on the simplex Δ_{d-1} is given by

$$f_{\mathrm{Dir}(\alpha)}(p) = \frac{1}{B(\alpha)} \prod_{j=0}^{d-1} p_j^{\alpha_j - 1}, \tag{190}$$

where $\alpha = (\alpha_0, \dots, \alpha_{d-1})$ is the vector of parameters and

$$B(\alpha) = \frac{\prod_{j=0}^{d-1} \Gamma(\alpha_j)}{\Gamma(\sum_{j=0}^{d-1} \alpha_j)}, \tag{191}$$

where $\Gamma(\cdot)$ denotes the gamma function. In our case $\alpha = (1, \dots, 1)$, hence

$$f_{\mathrm{Dir}(1,\dots,1)}(p) = \frac{1}{B(\mathbf{1})} = \frac{\Gamma(\sum_{j=0}^{d-1} 1)}{\prod_{j=0}^{d-1} \Gamma(1)} = \Gamma(d) = (d-1)! \tag{192}$$

Since for a Haar-distributed unitary matrix U the vector $W = (W_0, \dots, W_{d-1})$ is uniformly distributed on the simplex Δ_{d-1}, we can replace the integration over U by integration over the simplex with the Dirichlet density:

$$I(\lambda) = \int \prod_{j=0}^{d-1} W_j^{\lambda_j} dU = \int_{\Delta_{d-1}} \prod_{j=0}^{d-1} p_j^{\lambda_j} f_{\mathrm{Dir}(1,\dots,1)}(p) dp. \tag{193}$$

By substituting the values calculated above, we have

$$I(\lambda) = (d-1)! \int_{\Delta_{d-1}} \prod_{j=0}^{d-1} p_j^{\lambda_j} dp. \tag{194}$$

Lemma 35. *For any parameters $\alpha_0, \dots, \alpha_{d-1} > 0$, the following identity holds:*

$$\int_{\Delta_{d-1}} \prod_{j=0}^{d-1} p_j^{\alpha_j - 1} dp = \frac{\prod_{j=0}^{d-1} \Gamma(\alpha_j)}{\Gamma\left(\sum_{j=0}^{d-1} \alpha_j\right)}. \tag{195}$$

Proof. Recall the definition of the Gamma function,

$$\Gamma(s) = \int_0^\infty t^{s-1} e^{-t} dt. \tag{196}$$

The product $\prod_{j=0}^{d-1} \Gamma(\alpha_j)$ can therefore be expressed as

$$\prod_{j=0}^{d-1} \Gamma(\alpha_j) = \prod_{j=0}^{d-1} \int_0^\infty x_j^{\alpha_j - 1} e^{-x_j} dx_j = \int_{[0,\infty)^d} \Big(\prod_{j=0}^{d-1} x_j^{\alpha_j - 1} \Big) e^{-\sum_{j=0}^{d-1} x_j} dx_0 \cdots dx_{d-1}. \tag{197}$$

We now perform the change of variables

$$t = \sum_{j=0}^{d-1} x_j, \qquad p_j = \frac{x_j}{t}, \qquad p \in \Delta_{d-1}. \tag{198}$$

The inverse transformation is $x_j = tp_j$, and the Jacobian of this change is $|\det J| = t^{d-1}$. Hence $dx_0 \cdots dx_{d-1} = t^{d-1} dt dp$. Substituting into the integral yields

$$\prod_{j=0}^{d-1} \Gamma(\alpha_j) = \int_{t=0}^{\infty} \int_{p \in \Delta_{d-1}} \Big(\prod_{j=0}^{d-1} (tp_j)^{\alpha_j - 1} \Big) e^{-t} t^{d-1} dp dt = \Big(\int_0^{\infty} t^{\sum_j \alpha_j - 1} e^{-t} dt \Big) \Big(\int_{\Delta_{d-1}} \prod_{j=0}^{d-1} p_j^{\alpha_j - 1} dp \Big)$$
$$= \Gamma\left(\sum_{j=0}^{d-1} \alpha_j \right) \int_{\Delta_{d-1}} \prod_{j=0}^{d-1} p_j^{\alpha_j - 1} dp. \tag{199}$$

Hence

$$\int_{\Delta_{d-1}} \prod_{j=0}^{d-1} p_j^{\alpha_j - 1} dp = \frac{\prod_{j=0}^{d-1} \Gamma(\alpha_j)}{\Gamma\left(\sum_{j=0}^{d-1} \alpha_j \right)}. \tag{200}$$

□

Using Lemma 35 and $\Gamma(d) = (d-1)!$ we obtain

$$I(\lambda) = (d-1)! \int_{\Delta_{d-1}} \prod_{j=0}^{d-1} p_j^{\lambda_j} dp = (d-1)! \frac{\prod_{j=0}^{d-1} \lambda_j!}{\left(\sum_{j=0}^{d-1} (\lambda_j + 1) - 1 \right)!} = (d-1)! \frac{\prod_{j=0}^{d-1} \lambda_j!}{(N+d-1)!} \tag{201}$$

Finally,

$$\mathbb{P}(\Lambda = \lambda) = m_\lambda I(\lambda) = \frac{N!}{\lambda_0! \lambda_1! \ldots \lambda_{d-1}!} (d-1)! \frac{\prod_{j=0}^{d-1} \lambda_j!}{(N+d-1)!} = \frac{(d-1)! N!}{(N+d-1)!}. \tag{202}$$

B.4 Proof of Lemma 9

Assume the measurement operators act on separate blocks, i.e. $E_{U,\mathbf{l}}^{(n_1)}$ acts on the first block subsystems n_1 and $E_{V,\mathbf{m}}^{(n_2)}$ acts on the second block n_2, and the input state factorizes as $\rho = \rho^{(n_1)} \otimes \rho^{(n_2)}$ with $\rho^{(n_1)} = \rho^{(n_2)} = |0 \ldots 0\rangle \langle 0 \ldots 0|$. Then

$$\begin{aligned}
&\operatorname{tr}\Big[(E_{U,\mathbf{l}}^{(n_1)} \otimes E_{V,\mathbf{m}}^{(n_2)}) (\rho^{(n_1)} \otimes \rho^{(n_2)}) \Big] \\
&= \operatorname{tr}\Big(E_{U,\mathbf{l}}^{(n_1)} \rho^{(n_1)} \Big) \operatorname{tr}\Big(E_{V,\mathbf{m}}^{(n_2)} \rho^{(n_2)} \Big) \\
&= \langle 0 \ldots 0 | E_{U,\mathbf{l}}^{(n_1)} | 0 \ldots 0 \rangle \langle 0 \ldots 0 | E_{V,\mathbf{m}}^{(n_2)} | 0 \ldots 0 \rangle \\
&= \int \prod_{s=0}^{n_1 - 1} | \langle l_s | U | 0 \rangle |^2 dU \int \prod_{t=0}^{n_2 - 1} | \langle m_t | V | 0 \rangle |^2 dV
\end{aligned} \tag{203}$$

For $Q_{U,V}^{(n_1, n_2)}$ we introduce the count vectors $\lambda^{(1)}$ and $\lambda^{(2)}$ for the first and the second block, respectively. Computing analogously as for $Q_U^{(N)}$, we obtain

$$\mathbb{P}(\Lambda^{(1)} = \lambda^{(1)}, \Lambda^{(2)} = \lambda^{(2)} | U, V) = m^{(1)}(\lambda^{(1)}) \prod_{j=0}^{d-1} W_j^{\lambda_j^{(1)}} \cdot m^{(2)}(\lambda^{(2)}) \prod_{j=0}^{d-1} Z_j^{\lambda_j^{(2)}} \tag{204}$$

where

$$\begin{aligned}
W_j &= | \langle j | U | 0 \rangle |^2, \\
Z_j &= | \langle j | V | 0 \rangle |^2.
\end{aligned} \tag{205}$$

Using the same reasoning as above

$$\mathbb{P}(\Lambda^{(1)} = \lambda^{(1)}, \Lambda^{(2)} = \lambda^{(2)}) = m^{(1)}(\lambda^{(1)}) \cdot I(\lambda^{(1)}) \cdot m^{(2)}(\lambda^{(2)}) \cdot I(\lambda^{(2)}) =$$

$$\frac{n_1!}{\lambda_0^{(1)}! \dots \lambda_{d-1}^{(1)}!}(d-1)! \frac{\prod_{j=0}^{d-1}(\lambda_j^{(1)})!}{(n_1+d-1)!} \frac{n_2!}{\lambda_0^{(2)}! \dots \lambda_{d-1}^{(2)}!}(d-1)! \frac{\prod_{j=0}^{d-1}(\lambda_j^{(2)})!}{(n_2+d-1)!} = \frac{n_1!n_2!(d-1)!(d-1)!}{(n_1+d-1)!(n_2+d-1)!} \tag{206}$$

This shows that the joint Haar-averaged distribution factorizes into two independent multinomial blocks corresponding to the two subsystems.

B.5 Counting admissible histogram decompositions

We first ignore the upper bounds and consider $k_j \geq 0$. In this case, the number of solutions is given by

$$\#\{k_j \geq 0 : \sum_j k_j = n_1\} = \binom{n_1+d-1}{d-1}. \tag{207}$$

We now subtract the solutions that violate the constraints $k_j \leq \lambda_j$ using the inclusion–exclusion principle. For each index j, let

$$A_j = \{(k_0, \dots, k_{d-1}) : k_j > \lambda_j\} \tag{208}$$

be the set of solutions where the j-th component exceeds its allowed limit. By the inclusion–exclusion principle,

$$L_\lambda(n_1, n_2) = \sum_{S \subseteq \{0,\dots,d-1\}} (-1)^{|S|} |A_S|, \tag{209}$$

where $A_S = \bigcap_{j \in S} A_j$.

For a fixed nonempty subset S, the condition A_S means that for each $j \in S$ we have $k_j \geq \lambda_j + 1$. Introducing new variables

$$k_j' = \begin{cases} k_j - (\lambda_j + 1), & j \in S, \\ k_j, & j \notin S, \end{cases} \tag{210}$$

we obtain

$$n_1 = \sum_{j \in S} k_j' + (\lambda_j + 1) + \sum_{j \notin S} k_j' \tag{211}$$

$$\sum_{j=0}^{d-1} k_j' = n_1 - \sum_{j \in S}(\lambda_j + 1), \tag{212}$$

where all $k_j' \geq 0$. Hence, the number of solutions corresponding to A_S is

$$|A_S| = \binom{n_1 - \sum_{j \in S}(\lambda_j + 1) + d - 1}{d-1}, \tag{213}$$

with the convention that $\binom{a}{b} = 0$ if $a < b$. For $S = \varnothing$ we recover $|A_\varnothing| = \binom{n_1+d-1}{d-1}$. Thus, the final expression is

$$L_\lambda(n_1, n_2) = \sum_{S \subseteq \{0,\dots,d-1\}} (-1)^{|S|} \binom{n_1 - \sum_{j \in S}(\lambda_j + 1) + d - 1}{d-1}, \tag{214}$$

where $\binom{a}{b} = 0$ for $a < b$.

Example 36. *Let $d = 3$ and $\lambda = (2, 1, 1)$, so that $N = 4$. We compute the number of pairs for $n_1 = 3$ and $n_2 = 1$.*

By applying the inclusion-exclusion principle, we have

$$L_\lambda(3, 1) = \sum_{S \subseteq \{0,1,2\}} (-1)^{|S|} \binom{3 - \sum_{j \in S}(\lambda_j + 1) + 2}{2}, \tag{215}$$

where we assume that $\binom{a}{b} = 0$ for $a < b$. Let us compute the contributions for all subsets S:

- *$S = \varnothing$: $\sum_{j \in S}(\lambda_j + 1) = 0$,*

$$(+1)\binom{3+2}{2} = \binom{5}{2} = 10. \tag{216}$$

- *$S = \{0\}$: $\sum_{j \in S}(\lambda_j + 1) = \lambda_0 + 1 = 2 + 1 = 3$,*

$$(-1)\binom{3-3+2}{2} = -\binom{2}{2} = -1. \tag{217}$$

- *$S = \{1\}$: $\lambda_1 + 1 = 1 + 1 = 2$,*

$$(-1)\binom{3-2+2}{2} = -\binom{3}{2} = -3. \tag{218}$$

- *$S = \{2\}$: analogously to $S = \{1\}$,*

$$-\binom{3}{2} = -3 \tag{219}$$

- *$S = \{0, 1\}$: $\sum_{j \in S}(\lambda_j + 1) = 3 + 2 = 5$,*

$$(+1)\binom{3-5+2}{2} = \binom{0}{2} = 0. \tag{220}$$

- *$S = \{0, 2\}$: analogously we have 0.*
- *$S = \{1, 2\}$: $\sum = 2 + 2 = 4$,*

$$(+1)\binom{3-4+2}{2} = \binom{1}{2} = 0. \tag{221}$$

- *$S = \{0, 1, 2\}$: $\sum = 3 + 2 + 2 = 7$,*

$$(-1)\binom{3-7+2}{2} = \binom{-2}{2} = 0 \tag{222}$$

Summing all the contributions, we obtain

$$L_\lambda(3, 1) = 10 - 1 - 3 - 3 + 0 + 0 + 0 - 0 = 3. \tag{223}$$

C Proof of Proposition 13

Recall that

$$S_0 = \{\lambda : \lambda_j \in \{0,1\},\ \sum_j \lambda_j = N\} \tag{224}$$

is the set of collision-free histograms, while

$$S_1 = \{\lambda : \text{exactly one}, \lambda_j = 2,\ \text{all others in } \{0,1\}\} \tag{225}$$

is the set of single-collision histograms. The histogram in S_0 is specified by choosing the N occupied bins among the d available bins. Hence

$$|S_0| = \binom{d}{N}. \tag{226}$$

Similarly, to construct a histogram in S_1, one first chooses the bin carrying the double occupation (in d possible ways), and then chooses the remaining $(N-2)$ singly occupied bins among the remaining $(d-1)$ bins. Therefore

$$|S_1| = d\binom{d-1}{N-2}. \tag{227}$$

Consequently,

$$\frac{|S_1|}{|S_0|} = \frac{N(N-1)}{d} + O(d^{-2}). \tag{228}$$

When λ has no collision ($\lambda \in S_0$), each of the N hits occupies a distinct bin. Assigning exactly n_1 to the first block amounts to choosing n_1 of these N bins, hence

$$L_{S_0}(n_1, n_2) = \binom{N}{n_1}. \tag{229}$$

Using the asymptotic expansion

$$\frac{(d-1)!}{(d+k-1)!} = d^{-k}\Big(1 - \frac{k(k-1)}{2d} + O(d^{-2})\Big), \qquad k = \text{const}, \tag{230}$$

we obtain

$$\begin{aligned} P^{(N)}(\lambda) &= N!\, d^{-N}\Big(1 - \frac{N(N-1)}{2d} + O(d^{-2})\Big), \\ P_{S_0}^{(n_1,n_2)}(\lambda) &= c_{n_1,n_2}\binom{N}{n_1} = N!\, d^{-N}\Big(1 - \frac{n_1(n_1-1)+n_2(n_2-1)}{2d} + O(d^{-2})\Big). \end{aligned} \tag{231}$$

where

$$c_{n_1,n_2} := \frac{n_1!\, n_2!\, (d-1)!^2}{(n_1+d-1)!\,(n_2+d-1)!}. \tag{232}$$

Thus

$$\begin{aligned} |P^{(N)}(\lambda) - P_{S_0}^{(n_1,n_2)}(\lambda)| &= \left| N! d^{-N}\left(1 - \frac{N(N-1)}{2d} - 1 + \frac{n_1(n_1-1)+n_2(n_2-1)}{2d}\right) + O(d^{-N-2})\right| \\ &= \left| -N! d^{-N}\left(\frac{N(N-1) - n_1(n_1-1) - n_2(n_2-1)}{2d}\right) + O(d^{-N-2})\right| \\ &= \left| -N! d^{-N}\frac{N^2 - n_1^2 - n_2^2}{2d} + O(d^{-N-2})\right| = \left| -N! d^{-N}\frac{2n_1n_2}{2d} + O(d^{-N-2})\right| \\ &= N!\, d^{-N}\frac{n_1 n_2}{d} + O(d^{-N-2}), \end{aligned} \tag{233}$$

Summing over all $\lambda \in S_0$ gives

$$\begin{aligned}\mathrm{TVD}_{S_0} &= \frac{1}{2}|S_0|\,N!d^{-N}\frac{n_1n_2}{d} = \frac{1}{2}\binom{d}{N}N!d^{-N}\frac{n_1n_2}{d} = \frac{n_1n_2}{2d}\frac{d!N!}{d^N N!(d-N)!}\\ &= \frac{n_1n_2}{2d}\frac{d(d-1)\cdots(d-N+1)}{d^N} = \frac{n_1n_2}{2d}\prod_{i=0}^{N-1}\Big(1-\frac{i}{d}\Big) = \frac{n_1n_2}{2d}(1+O(d^{-1})) = \frac{n_1n_2}{2d}+O(d^{-2}).\end{aligned} \tag{234}$$

The expansion holds uniformly for every fixed N. We now turn to histograms with exactly one collision. For a histogram with one bin equal to 2 ($\lambda \in S_1$),

$$L_{S_1}(n_1,n_2) = \binom{N-2}{n_1} + \binom{N-2}{n_1-1} + \binom{N-2}{n_1-2} = \binom{N}{n_1} - \binom{N-2}{n_1-1}. \tag{235}$$

Thus

$$\begin{aligned}P^{(n_1,n_2)}_{S_1}(\lambda) &= c_{n_1,n_2}L_{S_1}(n_1,n_2)\\ &= n_1!n_2!\,d^{-N}\left(\binom{N}{n_1} - \binom{N-2}{n_1-1}\right)\Big(1-\frac{n_1(n_1-1)+n_2(n_2-1)}{2d}+O(d^{-2})\Big)\\ &= N!\,d^{-N}\Big(1-\frac{n_1(n_1-1)+n_2(n_2-1)}{2d}\Big) - n_1!n_2!\,d^{-N}\binom{N-2}{n_1-1} + O(d^{-N-2}).\end{aligned} \tag{236}$$

and

$$\begin{aligned}&P^{(N)}(\lambda) - P^{(n_1,n_2)}_{S_1}(\lambda) = N!d^{-N}\Big(\frac{-N(N-1)+n_1(n_1-1)+n_2(n_2-1)}{2d}\Big)+\\ &+ n_1!n_2!\,d^{-N}\binom{N-2}{n_1-1} + O(d^{-N-2}) = -N!d^{-N}\frac{n_1n_2}{d} + n_1!n_2!\,d^{-N}\binom{N-2}{n_1-1} + O(d^{-N-2}).\end{aligned} \tag{237}$$

We now evaluate the contribution of single-collision histograms $\lambda \in S_1$. Using $|S_1| = d\binom{d-1}{N-2}$ we get

$$\begin{aligned}\mathrm{TVD}_{S_1} &= \frac{1}{2}|S_1|\left|N!d^{-N}\frac{n_1n_2}{d} - n_1!n_2!\,d^{-N}\binom{N-2}{n_1-1} + O(d^{-N-2})\right|\\ &= \left|\frac{1}{2}d\binom{d-1}{N-2}N!d^{-N}\frac{n_1n_2}{d} - \frac{1}{2}d\binom{d-1}{N-2}n_1!n_2!\,d^{-N}\binom{N-2}{n_1-1}\right| + O(d^{-2}) =: |T_A - T_B| + O(d^{-2})\end{aligned} \tag{238}$$

The first term becomes

$$\begin{aligned}T_A &= \frac{1}{2}d\binom{d-1}{N-2}N!d^{-N}\frac{n_1n_2}{d} = \frac{n_1n_2}{2d}\frac{d(d-1)\cdots(d-N+2)}{d^{N-1}}\frac{N(N-1)}{d}\\ &= \frac{n_1n_2}{2d}\frac{N(N-1)}{d}(1+O(d^{-1})) = \frac{n_1n_2N(N-1)}{2d^2}(1+O(d^{-1})) = O(d^{-2})\end{aligned} \tag{239}$$

The second term

$$\begin{aligned}T_B &= \frac{1}{2}d\binom{d-1}{N-2}n_1!n_2!\,d^{-N}\binom{N-2}{n_1-1} = \frac{1}{2}\frac{1}{(N-2)!}\frac{1}{d}(1+O(d^{-1}))\frac{(N-2)!n_1!n_2!}{(n_1-1)!(n_2-1)!}\\ &= \frac{n_1n_2}{2d}(1+O(d^{-1})) = \frac{n_1n_2}{2d}+O(d^{-2})\end{aligned} \tag{240}$$

Collecting terms we have

$$\mathrm{TVD}_{S_1} = \left|O\left(d^{-2}\right) - \Big(\frac{n_1n_2}{2d}+O\left(d^{-2}\right)\Big)\right| + O\left(d^{-2}\right) = \frac{n_1n_2}{2d}+O\left(d^{-2}\right) \tag{241}$$

Summing the two leading contributions yields

$$\begin{aligned}\mathrm{TVD}\left(P^{(N)},P^{(n_1,n_2)}\right) &= \mathrm{TVD}_{S_0}+\mathrm{TVD}_{S_1}+O(d^{-2}) = \frac{n_1n_2}{2d}+O(d^{-2})+\frac{n_1n_2}{2d}+O(d^{-2})\\ &= \frac{n_1n_2}{d}+O(d^{-2})\end{aligned} \tag{242}$$

D Proof of Lemma 14

We use the inclusion–exclusion formula

$$L_\lambda(n_1,n_2) = \sum_{S\subseteq\{0,\dots,d-1\}} (-1)^{|S|}\binom{n_1-\sum_{j\in S}(\lambda_j+1)+d-1}{d-1}, \tag{243}$$

with the convention that the binomial coefficient is zero if the upper argument is smaller than $d-1$. For fixed $d\geq 2$ and all integers $a=O(n_1)$, uniformly including the boundary region $a=0$, we have

$$\binom{a+d-1}{d-1}\mathbf{1}_{\{a\geq 0\}} = \frac{a_+^{d-1}}{(d-1)!} + O(n_1^{d-2}). \tag{244}$$

Indeed, for $a\geq 0$ this follows from expanding $(a+1)\cdots(a+d-1)/(d-1)!$ as a polynomial in a, while for $a<0$ both leading terms vanish. Applying this estimate to

$$a_S := n_1 - \sum_{j\in S}(\lambda_j+1) \tag{245}$$

and summing over the finitely many subsets S gives the uniform expansion

$$L_\lambda(n_1,n_2) = \frac{1}{(d-1)!}\sum_{S\subseteq\{0,\dots,d-1\}} (-1)^{|S|}\left(n_1-\sum_{j\in S}(\lambda_j+1)\right)_+^{d-1} O(n_1^{d-2}). \tag{246}$$

Let $\ell_j=\lambda_j/N$ and $\alpha_N=n_1/N$. Since $n_1=\lfloor \alpha N\rfloor$, we have $\alpha_N\to\alpha$ and

$$\frac{\sum_{j\in S}(\lambda_j+1)}{n_1} = \sum_{j\in S}\frac{\ell_j}{\alpha_N} + \frac{|S|}{n_1}. \tag{247}$$

The function $x\mapsto(1-x)_+^{d-1}$ is Lipschitz on every bounded interval, and the quantities above remain uniformly bounded for $\ell\in\Delta_{d-1}$. Hence, uniformly in λ and in S,

$$\left(1-\frac{\sum_{j\in S}(\lambda_j+1)}{n_1}\right)_+^{d-1} = \left(1-\sum_{j\in S}\frac{\ell_j}{\alpha}\right)_+^{d-1} O(n_1^{-1}). \tag{248}$$

After multiplication by n_1^{d-1} and summation over S, this error is $O(n_1^{d-2})$. Finally, replacing ℓ_j/α by $\min\{1,\ell_j/\alpha\}$ does not change any term: if some $\ell_j/\alpha\geq 1$ appears in a subset S, then the corresponding positive part is already zero; otherwise the minimum leaves the term unchanged. Therefore

$$L_\lambda(n_1,n_2) = \frac{n_1^{d-1}}{(d-1)!}\varphi_\alpha(\ell) + O(n_1^{d-2}), \tag{249}$$

uniformly over all histograms λ, where

$$\varphi_\alpha(\ell) = \sum_{S\subseteq\{0,\dots,d-1\}} (-1)^{|S|}\left(1-\sum_{j\in S}\min\left\{1,\frac{\ell_j}{\alpha}\right\}\right)_+^{d-1}. \tag{250}$$

E Proof of Proposition 20

Observe that

$$\binom{k+d-1}{d-1} = \frac{d(d+1)\cdots(d+k-1)}{k!} = \frac{d^k}{k!}\prod_{j=0}^{k-1}\left(1+\frac{j}{d}\right). \tag{251}$$

Taking logarithms, we obtain

$$\log \prod_{j=0}^{k-1} \left(1 + \frac{j}{d}\right) = \sum_{j=0}^{k-1} \log\left(1 + \frac{j}{d}\right). \tag{252}$$

Using $\log(1+x) = x - \frac{x^2}{2} + O(x^3)$, valid for small x, we get

$$\begin{aligned}
&\sum_{j=0}^{k-1} \log\left(1+\frac{j}{d}\right) = \sum_{j=0}^{k-1}\left(\frac{j}{d} - \frac{j^2}{2d^2} + O\left(\frac{j^3}{d^3}\right)\right) = \frac{1}{d}\sum_{j=0}^{k-1} j - \frac{1}{2d^2}\sum_{j=0}^{k-1} j^2 + O\left(\frac{1}{d^3}\sum_{j=0}^{k-1} j^3\right) \\
&= \frac{k(k-1)}{2d} - \frac{k(k-1)(2k-1)}{12d^2} + O\left(\frac{k^4}{d^3}\right) = \frac{k(k-1)}{2d} + O\left(\frac{k^3}{d^2}\right) + O\left(\frac{k^4}{d^3}\right) = \frac{k(k-1)}{2d} + O\left(\frac{k^3}{d^2}\right).
\end{aligned} \tag{253}$$

Therefore

$$\log \prod_{j=0}^{k-1}\left(1+\frac{j}{d}\right) = \frac{k(k-1)}{2d} + O\left(\frac{k^3}{d^2}\right). \tag{254}$$

Exponentiating, and using $e^x = 1 + x + O(x^2)$ for small x, we obtain

$$\begin{aligned}
\prod_{j=0}^{k-1}\left(1+\frac{j}{d}\right) &= \exp\left(\frac{k(k-1)}{2d} + O\left(\frac{k^3}{d^2}\right)\right) = \exp\left(\frac{k(k-1)}{2d}\right)\exp\left(O\left(\frac{k^3}{d^2}\right)\right) \\
&= \left(1 + \frac{k(k-1)}{2d} + O\left(\frac{k^4}{d^2}\right)\right)\left(1 + O\left(\frac{k^3}{d^2}\right)\right) = 1 + \frac{k(k-1)}{2d} + O\left(\frac{k^4}{d^2}\right).
\end{aligned} \tag{255}$$

Here we used that $\left(\frac{k(k-1)}{2d}\right)^2 = O\left(\frac{k^4}{d^2}\right)$ dominates the error term $O\left(\frac{k^3}{d^2}\right)$. Hence

$$\binom{k+d-1}{d-1} = \frac{d^k}{k!}\left(1 + \frac{k(k-1)}{2d} + O\left(\frac{k^4}{d^2}\right)\right). \tag{256}$$

Applying this asymptotic expansion together with

$$\frac{1+a}{1+b} = (1+a)\frac{1}{1+b} = (1+a)(1-b+O(b^2)) = 1 + (a-b) + O(a^2+b^2), \tag{257}$$

for sufficiently small a, b, we obtain the expansion for $\overline{M}$

$$\begin{aligned}
\overline{M} &= \frac{M_1 M_2}{M} = \frac{N!}{n_1! n_2!}\frac{1 + \frac{n_1(n_1-1)+n_2(n_2-1)}{2d}}{1 + \frac{N(N-1)}{2d}} + O\left(\frac{N^4}{d^2}\right) \\
&= \binom{N}{n_1}\left(1 + \frac{1}{d}\left(\frac{n_1(n_1-1)+n_2(n_2-1)}{2} - \frac{N(N-1)}{2}\right) + O(d^{-2})\right) + O\left(\frac{N^4}{d^2}\right) \\
&= \binom{N}{n_1}\left(1 - \frac{n_1 n_2}{d} + O\left(\frac{N^4}{d^2}\right)\right).
\end{aligned} \tag{258}$$

The same stratification by collision structure as in Proposition 13 remains valid uniformly when $N = o(\sqrt{d})$. For collision-free histograms S_0, each outcome appears at most once, and the decomposition of the total histogram into two blocks reduces to choosing which n_1 outcomes belong to the first block. Then

$$L_{S_0}(n_1, n_2) = \binom{N}{n_1}, \tag{259}$$

Hence, using $\frac{1}{1-x} = 1 + x + O(x^2)$ with $x = \frac{n_1 n_2}{d}$,

$$\frac{L_{S_0}(n_1, n_2)}{M} = \frac{1}{1 - \frac{n_1 n_2}{d} + O\left(\frac{N^4}{d^2}\right)} = 1 + \frac{n_1 n_2}{d} + O\left(\frac{N^4}{d^2}\right) \tag{260}$$

For histograms in S_1 with a single double occupation, the combinatorics is slightly modified, since the doubleton may be assigned entirely to one block or split across the two blocks.

$$L_{S_1}(n_1, n_2) = \binom{N-2}{n_1} + \binom{N-2}{n_1 - 1} + \binom{N-2}{n_1 - 2} = \binom{N}{n_1} - \binom{N-2}{n_1 - 1}. \tag{261}$$

The drop is

$$\Delta_{S_1} = \binom{N}{n_1} - L_{S_1}(n_1, n_2) = \binom{N-2}{n_1 - 1} = \binom{N}{n_1} \frac{n_1 n_2}{N(N-1)}, \tag{262}$$

Hence

$$\frac{\Delta_{S_1}}{M} = \frac{\binom{N}{n_1} \frac{n_1 n_2}{N(N-1)}}{\binom{N}{n_1}\left(1 - \frac{n_1 n_2}{d}\right)} + O\left(\frac{N^4}{d^2}\right) = \frac{n_1 n_2}{N(N-1)}\left(1 + \frac{n_1 n_2}{d}\right) + O\left(\frac{N^4}{d^2}\right) \tag{263}$$

Stratum sizes:

$$|S_0| = \binom{d}{N}, \qquad |S_1| = d\binom{d-1}{N-2} = |S_0| \frac{N(N-1)}{d - N + 1}. \tag{264}$$

Moreover, the contribution of histograms with at least two collisions is negligible in the sparse regime. Indeed, such configurations require either two distinct double occupations or one occupation number at least three. Their respective cardinalities satisfy

$$|S_2| = \binom{d}{2}\binom{d-2}{N-4}, \qquad |S_3| = d\binom{d-1}{N-3}, \tag{265}$$

and therefore

$$\mathbb{P}(S_{\geq 2\text{coll.}}) = O\left(\frac{N^4}{d^2}\right), \tag{266}$$

uniformly in the regime $N = o(\sqrt{d})$. Putting this into the uniform-expectation identity gives

$$\begin{aligned} \mathrm{TVD}\left(P^{(N)}, P^{(n_1,n_2)}\right) &= \frac{1}{2}\mathbb{E}_\lambda\left[\left|1 - \frac{L_\lambda(n_1, n_2)}{M}\right|\right] = \frac{1}{2}\left(\mathbb{P}(S_0)\mathbb{E}\left[\left|1 - \frac{L_{S_0}(n_1, n_2)}{M}\right| \middle| S_0\right] + \right. \\ &\left. + \mathbb{P}(S_1)\mathbb{E}\left[\left|1 - \frac{L_{S_1}(n_1, n_2)}{M}\right| \middle| S_1\right] + \mathbb{P}(S_{\geq 2\text{coll.}})\mathbb{E}\left[\left|1 - \frac{L_{\geq 2\text{coll.}}(n_1, n_2)}{M}\right| \middle| S_{\geq 2\text{coll.}}\right]\right) \\ &= \frac{1}{2}\mathbb{P}(S_0)\mathbb{E}\left[\left|1 - \frac{L_{S_0}(n_1, n_2)}{M}\right| \middle| S_0\right] + \frac{1}{2}\mathbb{P}(S_1)\mathbb{E}\left[\left|1 - \frac{L_{S_1}(n_1, n_2)}{M}\right| \middle| S_1\right] + O\left(\frac{N^4}{d^2}\right) \end{aligned} \tag{267}$$

Since the integrand is constant on the strata S_0 and S_1, the corresponding conditional expectations reduce to deterministic values.

$$\mathrm{TVD}\left(P^{(N)}, P^{(n_1,n_2)}\right) = \frac{1}{2}\mathbb{P}(S_0)\left|1 - \frac{L_{S_0}(n_1, n_2)}{M}\right| + \frac{1}{2}\mathbb{P}(S_1)\left|1 - \frac{L_{S_1}(n_1, n_2)}{M}\right| + O\left(\frac{N^4}{d^2}\right) \tag{268}$$

Since histograms are uniformly distributed, the probability of observing zero or one collision is given by

$$\begin{aligned} \mathbb{P}(S_0) &= \frac{|S_0|}{M} = \frac{\binom{d}{N}}{M} = \frac{d^N}{N!}\left(1 - \frac{N(N-1)}{2d} + O\left(\frac{N^4}{d^2}\right)\right) \cdot \frac{N!}{d^N}\left(1 + \frac{N(N-1)}{2d} + O\left(\frac{N^4}{d^2}\right)\right) \\ &= \frac{1 - \frac{N(N-1)}{2d}}{1 + \frac{N(N-1)}{2d}} + O\left(\frac{N^4}{d^2}\right) = 1 - \frac{N(N-1)}{d} + O\left(\frac{N^4}{d^2}\right) \end{aligned} \tag{269}$$

and

$$\begin{aligned}
\mathbb{P}(S_1) &= \frac{|S_1|}{M} = \frac{|S_0|}{M}\frac{N(N-1)}{d-N+1} = \left(1 - \frac{N(N-1)}{d} + O\left(\frac{N^4}{d^2}\right)\right)\frac{N(N-1)}{d-N+1} \\
&= \left(1 - \frac{N(N-1)}{d} + O\left(\frac{N^4}{d^2}\right)\right)\frac{N(N-1)}{d}\left(\frac{1}{1-\frac{N-1}{d}}\right) \\
&= \frac{N(N-1)}{d}\left(1 - \frac{N(N-1)}{d} + O\left(\frac{N^4}{d^2}\right)\right)\left(1 + \frac{N-1}{d} + O\left(\frac{N^2}{d^2}\right)\right) \\
&= \frac{N(N-1)}{d}\left(1 - \frac{N(N-1)}{d} + O\left(\frac{N^4}{d^2}\right)\right)\left(1 + O\left(\frac{N}{d}\right)\right) \\
&= \frac{N(N-1)}{d} + O\left(\frac{N^3}{d^2}\right) + O\left(\frac{N^4}{d^2}\right) = \frac{N(N-1)}{d} + O\left(\frac{N^4}{d^2}\right)
\end{aligned} \tag{270}$$

where $M = \binom{N+d-1}{d-1}$ is the total number of histograms. Then

$$\begin{aligned}
&\mathrm{TVD}\left(P^{(N)}, P^{(n_1,n_2)}\right) = \frac{1}{2}\frac{|S_0|}{M}\frac{n_1n_2}{d} + \frac{1}{2}\frac{|S_1|}{M}\frac{\Delta_{S_1}}{M} + O\left(\frac{N^4}{d^2}\right) \\
&= \frac{1}{2}\left(1 - \frac{N(N-1)}{d}\right)\frac{n_1n_2}{d} + \frac{1}{2}\frac{N(N-1)}{d}\frac{n_1n_2}{N(N-1)}\left(1 + \frac{n_1n_2}{d}\right) + O\left(\frac{N^4}{d^2}\right) \\
&= \frac{n_1n_2}{2d} - \frac{n_1n_2N(N-1)}{2d^2} + \frac{n_1n_2}{2d} + \frac{n_1^2n_2^2}{2d^2} + O\left(\frac{N^4}{d^2}\right) \\
&= \frac{n_1n_2}{2d} + \frac{n_1n_2}{2d} + O\left(\frac{N^4}{d^2}\right) = \frac{n_1n_2}{d} + O\left(\frac{N^4}{d^2}\right).
\end{aligned} \tag{271}$$

F Critical scaling regime

F.1 Proof of Proposition 23

The robust route is to work with the number of doubletons

$$D_2(\lambda) := \#\{i : \lambda_i = 2\}, \tag{272}$$

and to separate higher occupancies through the event

$$A_d := \left\{\max_i \lambda_i \le 2\right\}. \tag{273}$$

On A_d, we have $C(\lambda) = D_2(\lambda)$.

Step 1: higher occupancies are rare. Under the uniform composition law, λ has the Dirichlet-multinomial representation

$$W \sim \mathrm{Dirichlet}(1, \dots, 1), \qquad (Y_1, \dots, Y_N) \mid W \text{ i.i.d. with law } W. \tag{274}$$

Hence

$$\mathbb{P}(Y_1 = Y_2 = Y_3) = \mathbb{E}\left[\sum_i W_i^3\right] = \frac{6}{(d+1)(d+2)}. \tag{275}$$

Therefore

$$\mathbb{P}(A_d^c) \le \mathbb{E}\left[\sum_i \binom{\lambda_i}{3}\right] = \binom{N}{3}\frac{6}{(d+1)(d+2)}. \tag{276}$$

For $N/\sqrt{d} \to c$, this gives

$$\mathbb{P}(A_d^c) = O\left(d^{-\frac{1}{2}}\right) \to 0. \tag{277}$$

Step 2: Poisson limit for $D_2(\lambda)$. For fixed $r \geq 1$, the falling factorial moment is

$$\mathbb{E}[(D_2)_r] = (d)_r \frac{\binom{N-2r+d-r-1}{d-r-1}}{\binom{N+d-1}{d-1}}, \tag{278}$$

where $(d)_j = \prod_{l=0}^{j-1}(d-l)$ is the Pochhammer symbol. This identity is exact: choose r ordered bins to carry occupancy 2, then distribute the remaining $N-2r$ balls among the other $d-r$ bins. With $N/\sqrt{d} \to c$ and fixed r,

$$\mathbb{E}[(D_2)_r] \longrightarrow c^{2r}. \tag{279}$$

Thus $D_2(\lambda) \Rightarrow \mathrm{Poisson}(c^2)$ by the method of factorial moments.

Step 3: transfer from D_2 to C. Since $C(\lambda) = D_2(\lambda)$ on A_d and $\mathbb{P}(A_d^c) \to 0$,

$$\mathbb{P}(C(\lambda) \neq D_2(\lambda)) \to 0. \tag{280}$$

Combining this with $D_2(\lambda) \Rightarrow \mathrm{Poisson}(c^2)$ yields

$$C(\lambda) \Rightarrow \mathrm{Poisson}(c^2), \tag{281}$$

which proves Proposition 23.

F.2 Proof of Proposition 24

Assume λ has exactly k bins of multiplicity 2 and no higher multiplicities. Then

$$L_\lambda(n_1, n_2) = [x^{n_1}](1+x)^{N-2k}(1+x+x^2)^k. \tag{282}$$

Using $1 + x + x^2 = (1+x)^2 - x$,

$$L_\lambda(n_1, n_2) = [x^{n_1}](1+x)^{N-2k} \sum_{j=0}^{k} (-1)^j \binom{k}{j} x^j (1+x)^{2k-2j} = \sum_{j=0}^{k} (-1)^j \binom{k}{j} \binom{N-2j}{n_1-j}. \tag{283}$$

Hence

$$\begin{aligned} \frac{L_\lambda(n_1, n_2)}{\binom{N}{n_1}} &= \sum_{j=0}^{k} (-1)^j \binom{k}{j} \frac{\binom{N-2j}{n_1-j}}{\binom{N}{n_1}} = \sum_{j=0}^{k} (-1)^j \binom{k}{j} \frac{n_1 \cdots (n_1-j+1) n_2 \cdots (n_2-j+1)}{N \cdots (N-2j+1)} \\ &= \sum_{j=0}^{k} (-1)^j \binom{k}{j} \frac{(n_1)_j (n_2)_j}{(N)_{2j}}, \end{aligned} \tag{284}$$

where $(m)_j = \prod_{l=0}^{j-1}(m-l)$. For each fixed j, if $n_1/N \to \alpha$ and $n_2/N \to 1-\alpha$, then

$$\frac{(n_1)_j (n_2)_j}{(N)_{2j}} \longrightarrow \alpha^j (1-\alpha)^j. \tag{285}$$

Therefore, for fixed k,

$$\frac{L_\lambda(n_1, n_2)}{\binom{N}{n_1}} \longrightarrow \sum_{j=0}^{k} (-1)^j \binom{k}{j} \big(\alpha(1-\alpha)\big)^j = \big(1 - \alpha(1-\alpha)\big)^k. \tag{286}$$

This proves Proposition 24.

F.3 Proof of normalization Lemma 25

Write

$$\overline{M} = \frac{M_1 M_2}{M}, \qquad M = \binom{N+d-1}{N}, \quad M_1 = \binom{n_1+d-1}{n_1}, \quad M_2 = \binom{n_2+d-1}{n_2}. \tag{287}$$

For $k = O(\sqrt{d})$,

$$\binom{d+k-1}{k} = \frac{d^k}{k!} \exp\left(\frac{k(k-1)}{2d} + O\left(\frac{k^3}{d^2}\right)\right). \tag{288}$$

Applying this to $k = N, n_1, n_2$ gives

$$M = \frac{d^N}{N!} \exp\left(\frac{N(N-1)}{2d} + O\left(\frac{N^3}{d^2}\right)\right), \tag{289}$$

$$M_i = \frac{d^{n_i}}{n_i!} \exp\left(\frac{n_i(n_i-1)}{2d} + O\left(\frac{N^3}{d^2}\right)\right), \quad i = 1, 2. \tag{290}$$

Hence

$$\overline{M} = \binom{N}{n_1} \exp\left(-\frac{n_1 n_2}{d} + O\left(\frac{N^3}{d^2}\right)\right), \tag{291}$$

so

$$\frac{\binom{N}{n_1}}{\overline{M}} = \exp\left(\frac{n_1 n_2}{d} + O\left(\frac{N^3}{d^2}\right)\right). \tag{292}$$

In the critical scaling $N/\sqrt{d} \to c$ with $n_1/N \to \alpha$,

$$\frac{\binom{N}{n_1}}{\overline{M}} \longrightarrow e^{\alpha(1-\alpha)c^2}. \tag{293}$$

This proves Lemma 25.

F.4 Proof of Theorem 27

Recall

$$\mathrm{TVD}\left(P^{(N)}, P^{(n_1,n_2)}\right) = \frac{1}{2} \mathbb{E}_{\lambda \sim P^{(N)}} \left[\left|1 - \frac{L_\lambda(n_1,n_2)}{\overline{M}}\right|\right]. \tag{294}$$

On the event A_d, a histogram with $D_2(\lambda) = k$ has only singletons and doubletons. For each fixed k, Proposition 24 gives

$$\frac{L_\lambda(n_1,n_2)}{\binom{N}{n_1}} \longrightarrow (1-\alpha(1-\alpha))^k. \tag{295}$$

Fix $\varepsilon > 0$. Choose R large enough so that

$$\mathbb{P}(D_2(\lambda) > R) < \varepsilon \tag{296}$$

for all sufficiently large d, using the convergence $D_2(\lambda) \Rightarrow \mathrm{Poisson}(c^2)$. Since R is fixed, Proposition 24 applies to each $k \leq R$, and taking the maximum over finitely many values of k yields uniform convergence on $\{0, \ldots, R\}$. Hence

$$\sup_{\lambda:\, D_2(\lambda) \leq R} \left| \frac{L_\lambda(n_1,n_2)}{\binom{N}{n_1}} - (1-\alpha(1-\alpha))^{D_2(\lambda)} \right| \longrightarrow 0, \tag{297}$$

where the supremum is taken over histograms satisfying $D_2(\lambda) = k$. Therefore, uniformly on the event $A_d \cap \{D_2(\lambda) \le R\}$,

$$\left| \frac{L_\lambda(n_1, n_2)}{\binom{N}{n_1}} - (1 - \alpha(1-\alpha))^{D_2(\lambda)} \right| \longrightarrow 0. \tag{298}$$

Since $\mathbb{P}(A_d^c) \to 0$ and $\mathbb{P}(D_2(\lambda) > R) < \varepsilon$, it follows that

$$\frac{L_\lambda(n_1, n_2)}{\binom{N}{n_1}} - (1 - \alpha(1-\alpha))^{D_2(\lambda)} \xrightarrow{p} 0. \tag{299}$$

Combining this with Lemma 25,

$$\frac{L_\lambda(n_1, n_2)}{\overline{M}} - e^{\alpha(1-\alpha)c^2}(1 - \alpha(1-\alpha))^{D_2(\lambda)} \xrightarrow{p} 0. \tag{300}$$

Since $D_2(\lambda) \Rightarrow K$ with $K \sim \text{Poisson}(c^2)$,

$$e^{\alpha(1-\alpha)c^2}(1 - \alpha(1-\alpha))^{D_2(\lambda)} \xrightarrow{d} e^{\alpha(1-\alpha)c^2}(1 - \alpha(1-\alpha))^{K}. \tag{301}$$

Since the difference between the two quantities converges to zero in probability, we have

$$\frac{L_\lambda(n_1, n_2)}{\overline{M}} \xrightarrow{d} e^{\alpha(1-\alpha)c^2}(1 - \alpha(1-\alpha))^{K}. \tag{302}$$

By Lemma 26, the corresponding absolute deviations are uniformly integrable. Therefore convergence in distribution upgrades to convergence of the expectation:

$$\frac{1}{2}\,\mathbb{E}\left[\left|1 - \frac{L_\lambda(n_1, n_2)}{\overline{M}}\right|\right] \longrightarrow \frac{1}{2}\,\mathbb{E}\left[\left|1 - e^{\alpha(1-\alpha)c^2}(1 - \alpha(1-\alpha))^{K}\right|\right]. \tag{303}$$

This is exactly Theorem 27.

G Block-resolved asymptotic

G.1 Proof of Proposition 28

We follow the collision decomposition used in Proposition 13 for the aggregate statistic, now applied to the pair histogram (a, b). Recall that the likelihood ratio is

$$\frac{P_{\text{sh}}(a, b)}{P_{\text{ind}}(a, b)} = \frac{M_1 M_2}{M\binom{N}{n_1}} \prod_{j=0}^{d-1} \binom{a_j + b_j}{a_j}. \tag{304}$$

Hence, by (64),

$$\text{TVD}\left(P_{\text{sh}}, P_{\text{ind}}\right) = \frac{1}{2}\mathbb{E}_{\text{ind}}\left[\left|\frac{P_{\text{sh}}(a, b)}{P_{\text{ind}}(a, b)} - 1\right|\right]. \tag{305}$$

We decompose the configuration space according to the collision structure. Define

$$S_0 = \{(a, b) : a_j + b_j \in \{0, 1\} \text{ for all } j\}, \tag{306}$$

$$S_\times = \left\{ (a, b) : \begin{array}{l} \exists!_j \text{ such that } (a_j, b_j) = (1, 1), \\ \text{and all remaining coordinates belong to} \\ \{(1, 0), (0, 1), (0, 0)\} \end{array} \right\}, \tag{307}$$

$$S_{\text{int}} = \left\{ (a, b) : \begin{array}{l} \exists!_j \text{ such that } (a_j, b_j) \in \{(2, 0), (0, 2)\}, \\ \text{and all remaining coordinates belong to} \\ \{(1, 0), (0, 1), (0, 0)\} \end{array} \right\}. \tag{308}$$

and let $S_{\geq 2}$ denote the remaining configurations, that is configurations containing either at least two collision events or a non-simple collision pattern (such as occupancies of size at least three). Accordingly, we write

$$\mathrm{TVD}\left(P_{\mathrm{sh}}, P_{\mathrm{ind}}\right) = \mathrm{TVD}_{S_0} + \mathrm{TVD}_{S_\times} + \mathrm{TVD}_{S_{\mathrm{int}}} + \mathrm{TVD}_{S_{\geq 2}}, \tag{309}$$

where for any sector S,

$$\mathrm{TVD}_S := \frac{1}{2} \sum_{(a,b)\in S} P_{\mathrm{ind}}(a,b) \left| \frac{P_{\mathrm{sh}}(a,b)}{P_{\mathrm{ind}}(a,b)} - 1 \right|. \tag{310}$$

We may therefore decompose the expectation according to the collision sectors as follows:

$$\mathrm{TVD}\left(P_{\mathrm{sh}}, P_{\mathrm{ind}}\right) = \frac{1}{2} \sum_{S\in\{S_0, S_\times, S_{\mathrm{int}}, S_{\geq 2}\}} \mathbb{P}_{\mathrm{ind}}(S)\mathbb{E}_{\mathrm{ind}}\left[\left| \frac{P_{\mathrm{sh}}(a,b)}{P_{\mathrm{ind}}(a,b)} - 1 \right| \,\middle|\, (a,b) \in S \right]. \tag{311}$$

We first estimate the contribution of $S_{\geq 2}$. Exactly as in Section 4.2, every additional collision reduces by one the number of freely chosen occupied bins. Therefore

$$|S_{\geq 2}| = O(d^{N-2}). \tag{312}$$

Moreover, uniformly in (a, b),

$$P_{\mathrm{sh}}(a,b) = O(d^{-N}), \qquad P_{\mathrm{ind}}(a,b) = O(d^{-N}). \tag{313}$$

Hence

$$\mathrm{TVD}_{S_{\geq 2}} = O(d^{-2}). \tag{314}$$

We now analyze configurations containing at most one collision. For $(a, b) \in S_0$, since $(a_j, b_j) \in \{(1, 0), (0, 1), (0, 0)\}$, all local factors satisfy

$$\binom{a_j + b_j}{a_j} = 1. \tag{315}$$

Therefore

$$\frac{P_{\mathrm{sh}}(a,b)}{P_{\mathrm{ind}}(a,b)} = \frac{M_1 M_2}{M\binom{N}{n_1}}. \tag{316}$$

Using the asymptotic expansion already employed in Proposition 13,

$$\binom{k+d-1}{d-1} = \frac{d^k}{k!}\left(1 + \frac{k(k-1)}{2d} + O(d^{-2})\right). \tag{317}$$

we obtain

$$\frac{P_{\mathrm{sh}}(a,b)}{P_{\mathrm{ind}}(a,b)} = 1 - \frac{n_1 n_2}{d} + O(d^{-2}). \tag{318}$$

Consequently

$$\left| \frac{P_{\mathrm{sh}}(a,b)}{P_{\mathrm{ind}}(a,b)} - 1 \right| = \frac{n_1 n_2}{d} + O(d^{-2}). \tag{319}$$

The complement of S_0 consists of configurations containing at least one collision. Under the independent-block model, there are $\binom{N}{2}$ pairs of samples, and each pair-collision event has probability of order d^{-1}. By the union bound, the probability of observing at least one collision is therefore

$$O\left(\frac{N^2}{d}\right). \tag{320}$$

For fixed N, this reduces to $O(d^{-1})$, hence

$$\mathbb{P}_{\text{ind}}(S_0) = 1 - O(d^{-1}). \tag{321}$$

Therefore

$$\text{TVD}_{S_0} = \frac{1}{2}\left(\frac{n_1 n_2}{d} + O(d^{-2})\right)\left(1 - O(d^{-1})\right) = \frac{n_1 n_2}{2d} + O(d^{-2}). \tag{322}$$

Now consider $(a, b) \in S_{\text{int}}$. Then there exists a unique coordinate such that

$$(a_j, b_j) = (2, 0) \qquad \text{or} \qquad (a_j, b_j) = (0, 2). \tag{323}$$

For this coordinate,

$$\binom{2}{2} = \binom{2}{0} = 1, \tag{324}$$

while all remaining factors are also equal to 1. Consequently,

$$\frac{P_{\text{sh}}(a, b)}{P_{\text{ind}}(a, b)} = 1 - \frac{n_1 n_2}{d} + O(d^{-2}), \tag{325}$$

exactly as in the collision-free case. Then

$$\left|\frac{P_{\text{sh}}(a, b)}{P_{\text{ind}}(a, b)} - 1\right| = \frac{n_1 n_2}{d} + O(d^{-2}). \tag{326}$$

This event requires an internal collision within one of the two blocks. In the first block there are $\binom{n_1}{2}$ pairs of samples, while in the second block there are $\binom{n_2}{2}$ such pairs. For two samples drawn from the same Haar-distributed probability vector, the collision probability equals

$$\sum_{j=0}^{d-1} \mathbb{E}[W_j^2] = \frac{2}{d+1}. \tag{327}$$

Therefore

$$\mathbb{P}_{\text{ind}}(S_{\text{int}}) = O\left(\frac{n_1^2 + n_2^2}{d}\right). \tag{328}$$

For fixed $N = n_1 + n_2$, this is $O(d^{-1})$. Consequently,

$$\text{TVD}_{S_{\text{int}}} = \frac{1}{2}\left(\frac{n_1 n_2}{d} + O(d^{-2})\right) O(d^{-1}) = O(d^{-2}). \tag{329}$$

Thus internal collisions do not contribute at leading order. Finally, consider $(a, b) \in S_\times$. Then, for the unique collision coordinate,

$$(a_j, b_j) = (1, 1), \tag{330}$$

and therefore

$$\binom{a_j + b_j}{a_j} = \binom{2}{1} = 2. \tag{331}$$

All remaining factors equal 1, hence

$$\frac{P_{\text{sh}}(a, b)}{P_{\text{ind}}(a, b)} = 2\left(1 - \frac{n_1 n_2}{d} + O(d^{-2})\right) = 2 - \frac{2n_1 n_2}{d} + O(d^{-2}). \tag{332}$$

In this case,

$$\left|\frac{P_{\text{sh}}(a, b)}{P_{\text{ind}}(a, b)} - 1\right| = \left|1 - \frac{2n_1 n_2}{d} + O(d^{-2})\right| = 1 + O(d^{-1}). \tag{333}$$

A configuration in $S_\times$ contains exactly one cross-block collision. There are $n_1 n_2$ pairs consisting of one sample from the first block and one sample from the second block. For each such pair, the probability that both samples occupy the same bin is $\frac{1}{d}$. Hence the expected number of cross-block collisions equals $\frac{n_1 n_2}{d}$. The probability of observing two or more cross-block collisions is $O(d^{-2})$, since this requires at least two distinct pairs of samples to collide simultaneously. Indeed, each additional collision contributes another factor d^{-1}. Therefore

$$\mathbb{P}_{\text{ind}}(S_\times) = \frac{n_1 n_2}{d} + O(d^{-2}). \tag{334}$$

Hence

$$\mathrm{TVD}_{S_\times} = \frac{1}{2}\left(\frac{n_1 n_2}{d} + O(d^{-2})\right)\left(1 + O(d^{-1})\right) = \frac{n_1 n_2}{2d} + O(d^{-2}). \tag{335}$$

Collecting all contributions gives

$$\mathrm{TVD}\left(P_{\text{sh}}, P_{\text{ind}}\right) = \frac{n_1 n_2}{d} + O(d^{-2}). \tag{336}$$

The same argument extends to the sparse regime $N = o(\sqrt{d})$. Indeed, the probability of observing at least two collisions is now

$$O\left(\frac{N^4}{d^2}\right), \tag{337}$$

since specifying two collisions requires choosing two pairs of samples, which can be done in $O(N^4)$ ways, while each collision contributes a factor d^{-1}. Moreover,

$$\mathbb{P}_{\text{ind}}(S_\times) = \frac{n_1 n_2}{d} + O\left(\frac{N^4}{d^2}\right), \tag{338}$$

while all remaining sectors contribute only

$$O\left(\frac{N^4}{d^2}\right). \tag{339}$$

Consequently,

$$\mathrm{TVD}\left(P_{\text{sh}}, P_{\text{ind}}\right) = \frac{n_1 n_2}{d} + O\left(\frac{N^4}{d^2}\right). \tag{340}$$

G.2 Proof of Theorem 29

We again start from the likelihood ratio

$$\frac{P_{\text{sh}}(a,b)}{P_{\text{ind}}(a,b)} = \frac{M_1 M_2}{M\binom{N}{n_1}} \prod_{j=0}^{d-1} \binom{a_j + b_j}{a_j}. \tag{341}$$

Recall that in the critical scaling regime

$$\frac{N}{\sqrt{d}} \to c, \qquad \frac{n_1}{N} \to \alpha, \qquad \frac{n_2}{N} \to 1 - \alpha. \tag{342}$$

We define the number of cross-block collisions by

$$C(a,b) := \#\{j : (a_j, b_j) = (1,1)\}. \tag{343}$$

Thus $C(a,b)$ counts the number of bins simultaneously occupied by exactly one sample from the first block and exactly one sample from the second block. Since

$$n_1 \sim \alpha N, \qquad n_2 \sim (1-\alpha)N, \qquad \frac{N}{\sqrt{d}} \to c, \tag{344}$$

we have

$$\frac{n_1 n_2}{d} \longrightarrow \alpha(1-\alpha)c^2 =: \beta. \tag{345}$$

We now compute the factorial moments of $C(a,b)$ under the independent-block model. For fixed $r \geq 1$ and all sufficiently large d,

$$\mathbb{E}_{\text{ind}}\big[(C)_r\big] = (d)_r \frac{\binom{n_1-r+d-r-1}{d-r-1}}{\binom{n_1+d-1}{d-1}} \frac{\binom{n_2-r+d-r-1}{d-r-1}}{\binom{n_2+d-1}{d-1}}. \tag{346}$$

Indeed, choose the ordered list of r distinct bins with $(a_j, b_j) = (1,1)$ and then distribute the remaining $n_1 - r$ and $n_2 - r$ counts over the remaining $d - r$ bins. Since $n_1, n_2 = O(\sqrt{d})$, the two ratios are asymptotic to $(n_1/d)^r$ and $(n_2/d)^r$, respectively. Hence

$$\mathbb{E}_{\text{ind}}\big[(C)_r\big] \longrightarrow \beta^r. \tag{347}$$

The method of factorial moments gives

$$C(a,b) \xrightarrow{d} \text{Poisson}(\beta). \tag{348}$$

It remains to check that non-simple collision configurations do not affect the likelihood ratio. Let

$$B_d := \{a_j + b_j \leq 2 \text{ for all } j\}. \tag{349}$$

By a union bound over triples of samples, using the Dirichlet moment estimates $\mathbb{E}\sum_j W_j^3 = O(d^{-2})$ and $\mathbb{E}\sum_j W_j^2 Z_j = O(d^{-2})$ for independent $W, Z \sim \text{Dirichlet}(1,\ldots,1)$, we get

$$\mathbb{P}_{\text{ind}}(B_d^c) = O\left(\frac{N^3}{d^2}\right) = o(1). \tag{350}$$

We now analyze the product

$$\prod_{j=0}^{d-1} \binom{a_j + b_j}{a_j}. \tag{351}$$

If a bin contains no collision, then

$$(a_j, b_j) \in \{(1,0), (0,1), (0,0)\}, \tag{352}$$

and therefore

$$\binom{a_j + b_j}{a_j} = 1. \tag{353}$$

If a bin contains an internal collision, i.e.

$$(a_j, b_j) = (2,0) \qquad \text{or} \qquad (a_j, b_j) = (0,2), \tag{354}$$

then again

$$\binom{2}{2} = \binom{2}{0} = 1. \tag{355}$$

Hence internal collisions do not affect the likelihood ratio. The only nontrivial contribution comes from cross-block collisions:

$$(a_j, b_j) = (1,1), \tag{356}$$

for which

$$\binom{a_j + b_j}{a_j} = \binom{2}{1} = 2. \tag{357}$$

Consequently, every cross-block collision contributes a factor 2, and therefore on B_d

$$\prod_{j=0}^{d-1} \binom{a_j + b_j}{a_j} = 2^{C(a,b)}. \tag{358}$$

Recall

$$\overline{M} = \frac{M_1 M_2}{M}. \tag{359}$$

By Lemma 25,

$$\frac{\binom{N}{n_1}}{\overline{M}} \longrightarrow e^{\beta}. \tag{360}$$

Equivalently,

$$\frac{\overline{M}}{\binom{N}{n_1}} = \frac{M_1 M_2}{M\binom{N}{n_1}} \longrightarrow e^{-\beta}. \tag{361}$$

Combining this with the previous step yields

$$\frac{P_{\text{sh}}(a,b)}{P_{\text{ind}}(a,b)} = \frac{M_1 M_2}{M\binom{N}{n_1}} \prod_{j=0}^{d-1} \binom{a_j + b_j}{a_j} \Longrightarrow e^{-\beta} 2^{C}, \tag{362}$$

where

$$C \sim \text{Poisson}(\beta). \tag{363}$$

Let

$$R_d(a,b) := \frac{P_{\text{sh}}(a,b)}{P_{\text{ind}}(a,b)}. \tag{364}$$

The variables R_d are nonnegative and satisfy $\mathbb{E}_{\text{ind}}[R_d] = 1$ exactly. The limiting variable also has mean one, since

$$\mathbb{E}\Big[e^{-\beta} 2^{C}\Big] = e^{-\beta} \exp(\beta(2-1)) = 1. \tag{365}$$

Thus convergence in distribution together with convergence of the first moments implies uniform integrability of $\{R_d\}$, and hence of $\{|R_d - 1|\}$. Using the likelihood-ratio representation

$$\text{TVD}(P_{\text{sh}}, P_{\text{ind}}) = \frac{1}{2}\mathbb{E}_{\text{ind}}\left[|R_d(a,b) - 1|\right], \tag{366}$$

we conclude that

$$\text{TVD}(P_{\text{sh}}, P_{\text{ind}}) \longrightarrow \frac{1}{2}\mathbb{E}\left[\left|e^{-\beta} 2^{C} - 1\right|\right]. \tag{367}$$

This proves the claim.

G.3 Proof of 30

Fix $d \geq 2$ and let $N \to \infty$ with $\frac{n_1}{N} \to \alpha \in (0,1)$. Under the shared-unitary model, conditioned on

$$W \sim \text{Unif}(\Delta_{d-1}), \tag{368}$$

all samples in both blocks are i.i.d. with distribution W. Therefore, by the law of large numbers,

$$\left(\frac{a}{n_1}, \frac{b}{n_2}\right) \xrightarrow{p} (W, W). \tag{369}$$

Under the independent-block model, the two blocks are generated independently from independent Haar-random probability vectors W_1, W_2. Hence

$$\left(\frac{a}{n_1}, \frac{b}{n_2}\right) \xrightarrow{p} (W_1, W_2), \tag{370}$$

where W_1, W_2 are independent and uniformly distributed on Δ_{d-1}. Thus the shared-model limit is supported on the diagonal

$$\mathcal{D} = \{(w, w) : w \in \Delta_{d-1}\}, \tag{371}$$

whereas the independent-model limit is absolutely continuous with respect to Lebesgue measure on

$$\Delta_{d-1} \times \Delta_{d-1}. \tag{372}$$

Since the diagonal $\mathcal{D}$ has codimension $d-1$, it has Lebesgue measure zero. Therefore

$$\mathbb{P}((W_1, W_2) \in \mathcal{D}) = 0. \tag{373}$$

Hence the two limiting measures are mutually singular. To turn this into a total variation statement, for $\delta > 0$ let

$$\mathcal{D}_\delta := \{(x, y) \in \Delta_{d-1} \times \Delta_{d-1} : \|x - y\|_1 \le \delta\}. \tag{374}$$

Under the shared model,

$$\mathbb{P}_{\text{sh}}\left(\left(\frac{a}{n_1}, \frac{b}{n_2}\right) \in \mathcal{D}_\delta\right) \longrightarrow 1 \tag{375}$$

for every fixed $\delta > 0$. Under the independent model, weak convergence gives

$$\limsup_{N \to \infty} \mathbb{P}_{\text{ind}}\left(\left(\frac{a}{n_1}, \frac{b}{n_2}\right) \in \mathcal{D}_\delta\right) \le \mathbb{P}(\|W_1 - W_2\|_1 \le \delta). \tag{376}$$

Because W_1 and W_2 have a joint density and $d \ge 2$, $\mathbb{P}(\|W_1 - W_2\|_1 \le \delta) \to 0$ as $\delta \downarrow 0$. Therefore, for arbitrarily small $\varepsilon > 0$ we may choose δ so that the independent model assigns asymptotic mass at most ε to $\mathcal{D}_\delta$, while the shared model assigns asymptotic mass one. Consequently,

$$\liminf_{N \to \infty} \text{TVD}(P_{\text{sh}}, P_{\text{ind}}) \ge 1 - \varepsilon. \tag{377}$$

Since total variation distance is always at most one and ε is arbitrary, the desired convergence follows.