Two-Sample Testing for Monte Carlo Evaluation in Particle Physics: Challenges and Opportunities

Zeyu Ding [1 2] **Cornelius Grunwald** [3] **Katja Ickstadt** [1 2 4] **Kevin Kröninger** [3 4] **Salvatore La Cagnina** [3]

Abstract

Monte Carlo (MC) simulation is central to modern particle physics, where evaluating the fidelity of simulated samples against reference distributions is a fundamental yet under-standardized task. We present `MCBench`, a modular benchmark suite for assessing MC sampling quality using distributional distance metrics, including the sliced Wasserstein distance and the maximum mean discrepancy (MMD). Applied to the high-dimensional, structured distributions arising in high-energy physics, the choice of test metric and kernel is known to affect sensitivity in high-dimensional settings (Gretton et al., 2012b). Whether modern learned or adaptive testing methods can provide improved sensitivity for the structured distributions arising in particle physics remains an open and practically important question. We describe the architecture of `MCBench`, illustrate key challenges posed by physics data through a controlled power study on gluon jet kinematics from the JetNet dataset, and outline concrete opportunities for distribution-distance-based hypothesis testing methods to advance the evaluation of physics simulators.

1. Introduction

Monte Carlo (MC) simulation lies at the heart of modern experimental particle physics. At facilities such as the Large Hadron Collider (LHC) at CERN, collisions between protons produce complex multi-particle final states whose statistical properties are modeled by MC event generators such as PYTHIA (Sjöstrand et al., 2015) and HERWIG (Bellm et al., 2016). A central challenge in this pipeline is *evaluating the fidelity of simulated samples*: do the generated events faithfully reproduce the distributions observed in real detector data, or in reference simulations? This is, fundamentally, a two-sample testing problem.

Despite its centrality, the problem of quantitatively benchmarking MC sampling quality in physics has received comparatively little systematic attention from the statistical machine learning community. Existing practice in physics largely relies on visual comparison of one-dimensional marginal histograms or domain-specific metrics that are not designed to detect subtle high-dimensional discrepancies (Grossi et al., 2025). At the same time, the hypothesis testing community has developed powerful methods, including kernel-based two-sample tests (Gretton et al., 2012a), learned kernels (Liu et al., 2020), and aggregated testing procedures (Schrab et al., 2023), that remain largely disconnected from physics applications.

This paper aims to bridge this gap. We make the following contributions:

1. We present `MCBench` (Ding et al., 2025), a modular, metric-agnostic benchmark suite for MC sampling quality assessment, implemented as an open-source Julia package.
2. We characterize the specific challenges that high-energy physics data poses for standard two-sample testing methods, including high dimensionality, complex correlation structure, multimodality, and non-IID sample generation.
3. We identify concrete opportunities for learned and adaptive kernel methods to improve sensitivity in this domain, demonstrated concretely through a power study on gluon jet kinematics from the JetNet dataset, and describe the infrastructure we provide for the hypothesis testing community to engage with real physics data.

Related work. The evaluation of generative models and simulators using distributional metrics has a growing literature in machine learning (Gretton et al., 2012a; Schrab et al., 2023; Liu et al., 2020; Zhou et al., 2026). In particle

[1]Department of Statistics, TU Dortmund University, Germany [2]Lamarr Institute for Machine Learning and AI, Germany [3]Department of Physics, TU Dortmund University, Germany [4]TU Dortmund Center for Data Science and Simulation, Germany. Correspondence to: Zeyu Ding <zeyu.ding@tu-dortmund.de>.

ICML 2026 Workshop on Hypothesis Testing, Seoul, South Korea, 2026. Copyright 2026 by the author(s).

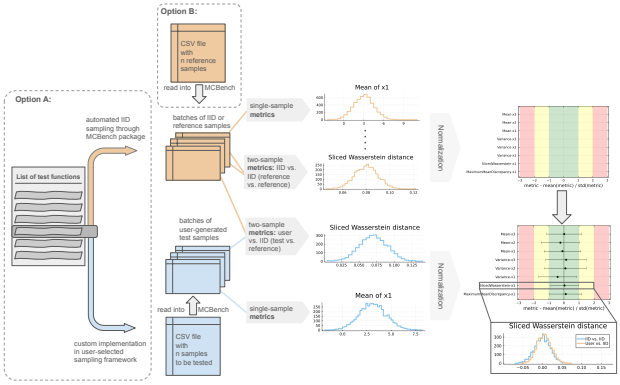

Figure 1. Workflow of the `MCBench` benchmark suite from Ding et al. (2025). Users select a target distribution and generate samples using an external sampler of their choice. The samples are imported and compared against reference IID samples from the same target using a configurable set of metrics, yielding normalized test statistic distributions that enable statistically grounded comparisons between sampling algorithms.

physics, recent work has begun to apply two-sample tests for generator validation (Grossi et al., 2025), but systematic, domain-neutral benchmarking infrastructure has been lacking. MCBENCH is designed to fill this gap and to serve as a shared platform for cross-disciplinary collaboration.

2. MCBench: A Benchmark Suite for Monte Carlo Evaluation

`MCBench` (Ding et al., 2025) is an open-source Julia package (`https://github.com/tudo-physik-e4/MCBench`) designed to provide standardized, quantitative evaluation of MC sampling algorithms. Its core design principle is *modularity*: test functions, sampling algorithms, and distance metrics are fully decoupled, allowing any combination to be evaluated within a unified workflow.

2.1. Workflow

The benchmarking pipeline operates in three stages. First, users select a *target function* from the provided collection or define a custom one. Second, user-generated samples, produced by any external sampler and provided in CSV or HDF5 format, are compared against large reference IID samples from the same target, generated internally by the package. Third, a configurable set of metrics is evaluated across batches of samples, yielding distributions of test statistics that enable statistically grounded comparisons.

2.2. Target Functions

`MCBench` includes a diverse collection of target distributions spanning a wide range of complexities: standard Gaussian distributions in dimensions $d \in \{1, 2, 3, 10, 100\}$, weakly and strongly correlated multivariate normals, Cauchy distributions, multimodal mixture models, and the eight-schools hierarchical model from `posteriorDB` (Magnusson et al., 2024). These targets are chosen to reflect both standard statistical benchmarks and the structural features, high dimensionality, multimodality, strong correlations, that arise in physics applications.

2.3. Metrics

Basic metrics. The suite evaluates standard marginal statistics (mean, variance, chi-square test statistics) across batches, providing intuitive diagnostics for low-dimensional or projected data.

Sliced Wasserstein Distance (SWD). The Wasserstein distance (Panaretos & Zemel, 2019) measures the minimum transport cost between two distributions, capturing geometric structure that marginal statistics miss. To overcome the $O(n^3 \log n)$ computational complexity of exact Wasserstein computation in high dimensions, `MCBench` implements the sliced Wasserstein distance (Rabin et al., 2011) as a two-sample test statistic via a permutation-based procedure, which has recently been shown to achieve the minimax-optimal separation rate $n^{-1/2}$ (Tran & Schreuder, 2025).

We note that while `MCBench` applies SWD uniformly across all target distributions, the theoretical guarantees of SWD require finite first moments and do not extend to heavy-tailed targets such as the Cauchy distribution. Incorporating moment-robust alternatives for such targets is planned for future versions of the benchmark.

Maximum Mean Discrepancy (MMD). The MMD (Gretton et al., 2006) measures the distance between distributions via their mean embeddings in a reproducing kernel Hilbert space (RKHS):

$$\mathrm{MMD}^2[p, q] = \|\mu_p - \mu_q\|^2_{\mathcal{H}}, \tag{1}$$

where $\mu_p = \mathbb{E}_{x\sim p}[k(x, \cdot)]$ is the kernel mean embedding of p. `MCBench` implements both the exact Gaussian kernel MMD ($O((n + m)^2 d)$ complexity) and a Random Fourier Features (RFF) approximation (Rahimi & Recht, 2007) with complexity $O(D(n + m)d)$, where D is the number of random features. The kernel bandwidth is selected via the median heuristic (Gretton et al., 2012b).

2.4. Interface for External Methods

A key design feature of `MCBench` is its open interface: any two-sample test statistic can be registered as a metric and evaluated within the benchmark pipeline. Samples are exchanged via standard file formats (CSV, HDF5) or directly as in-memory arrays, making integration of external meth-

ods, including learned kernel tests, straightforward. This is the primary entry point through which the hypothesis testing community can contribute methods and evaluate them on physics-relevant test cases.

3. Challenges in Particle Physics: A Testing Perspective

At experiments such as ATLAS at the Large Hadron Collider, proton–proton collisions produce complex multi-particle final states whose statistical properties are modeled by MC event generators (Sjöstrand et al., 2015; Bellm et al., 2016). A central task in this pipeline is to determine whether the simulated samples faithfully reproduce the distributions observed in real detector data. What makes this problem particularly challenging is not the bulk of the distributions, where simulated and observed events typically agree well, but rather the tails. Deviations from the Standard Model, if they exist, are expected to be most noticeable there, either as localized excess events at a specific energy scale (resonant structures arising from the production of new, heavy particles), or as a broad and relatively flat increase in event counts across a wide kinematic range. Critically, even when a discrepancy is detected in the tail, it remains unclear whether it reflects genuine new physics or simply a mismodelling of known Standard Model processes by the MC generator. Disentangling these two possibilities demands two-sample tests with well-calibrated Type-I error control and provable sensitivity in the tail regions.

Figure 2 illustrates this challenge concretely using gluon jets from the JetNet dataset (Kansal et al., 2023), a publicly available benchmark representative of the kinematic distributions encountered in hadron collider experiments. The three jet-level features, namely transverse momentum p_T, pseudorapidity η, and jet mass m_j, exhibit qualitatively different distributional structures: p_T and m_j are strongly right-skewed with heavy tails spanning several orders of magnitude, while η is approximately symmetric and near-Gaussian. We overlay three representative deformations following Grossi et al. (2025): a mean shift (μ-deformation), a variance scaling (Σ_{ii}-deformation), and an additive Gaussian smearing ($\mathcal{N}$-deformation), each with $\epsilon = 0.5$. As the figure shows, deformations that are nearly invisible in the bulk become apparent only in the tail regions, where event counts are low and standard tests with fixed kernels or uniform projection weights lose sensitivity. This is precisely the regime where the choice of test statistic, kernel, and bandwidth matters most, and where the hypothesis testing community has the most to contribute.

To quantitatively assess the sensitivity of SWD and MMD to physically motivated distributional deformations, we conduct a controlled power study on the JetNet gluon jet dataset (Kansal et al., 2023). We use the three jet-level

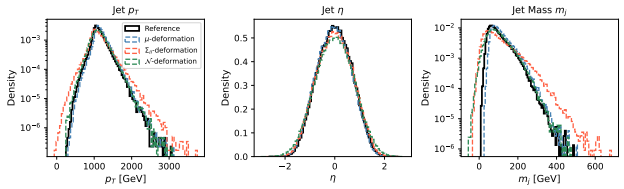

Figure 2. Marginal distributions of jet-level kinematic features (p_T, η, m_j) for gluon jets from the JetNet dataset (Kansal et al., 2023), shown for the reference distribution and under three representative deformations following Grossi et al. (2025): a mean shift (μ-deformation), a variance scaling (Σ_{ii}-deformation), and an additive Gaussian smearing ($\mathcal{N}$-deformation), each with $\epsilon = 0.5$. The heavy-tailed structure of p_T and m_j (log scale) and the near-Gaussian η illustrate the distributional complexity typical of high-energy physics data. Discrepancies introduced by imperfect generators may be subtle in the bulk but pronounced in the tail regions, precisely where standard two-sample tests with fixed kernels lose sensitivity.

kinematic features (p_T, η, m_j), standardized to zero mean and unit variance prior to metric computation. Following Grossi et al. (2025), we consider three deformation types applied to the standardized data: a mean shift (μ-deformation, $X \mapsto X + \varepsilon\mathbf{1}$), a variance scaling ($\Sigma_{ii}$-deformation, $X \mapsto (1+\varepsilon)(X - \bar{X}) + \bar{X}$), and additive Gaussian noise ($\mathcal{N}$-deformation, $X \mapsto X + \mathcal{N}(0, \varepsilon^2 I)$), for $\varepsilon \in \{0.05, 0.10, \ldots, 0.50\}$.

For each deformation type and strength ε, we draw $n = 1{,}000$ reference samples and $n = 1{,}000$ deformed samples from the dataset, compute the SWD (via random projections with 200 directions) and MMD (Gaussian kernel with median-heuristic bandwidth), and repeat this procedure 500 times. The rejection threshold is determined at level $\alpha = 0.05$ from a null distribution estimated as follows: we draw two independent batches of $n = 1{,}000$ samples from the undeformed JetNet reference data, compute the test statistic on each pair, and repeat this procedure 500 times. The resulting 500 null statistic values are used to estimate the $(1-\alpha)$-quantile, which serves as the rejection threshold. Empirical power is then estimated as the fraction of repetitions—each using one reference batch and one deformed batch—in which the test statistic exceeds this threshold.

To illustrate these challenges quantitatively, we conduct a controlled power study on the JetNet gluon jet dataset using SWD and MMD under the three deformation types described above; full experimental details are provided in Appendix A. Table 1 summarizes the results for selected values of ε. Both metrics detect μ- and Σ_{ii}-deformations reliably at moderate deformation strength ($\varepsilon \geq 0.2$), consistent with the visible bulk shifts observed in Figure 2. In contrast, the $\mathcal{N}$-deformation remains largely undetected at small scales: empirical power stays near the nominal level $\alpha = 0.05$ for $\varepsilon \leq 0.2$, and only approaches 1.000 at

Table 1. Empirical power curves of SWD and MMD under three deformations on JetNet gluon jets ($n = 1{,}000$, $\alpha = 0.05$, 500 repetitions). Selected values of ε; full results in Table 2.

Deformation	ε	SWD (mean)	SWD power	MMD (mean)	MMD power
μ	0.1	0.1052	0.908	0.0031	0.920
μ	0.2	0.1850	1.000	0.0123	1.000
μ	0.3	0.2682	1.000	0.0268	1.000
μ	0.5	0.4380	1.000	0.0727	1.000
Σ_{ii}	0.1	0.0954	0.868	0.0027	0.938
Σ_{ii}	0.2	0.1622	1.000	0.0096	1.000
Σ_{ii}	0.3	0.2370	1.000	0.0200	1.000
Σ_{ii}	0.5	0.3907	1.000	0.0471	1.000
$\mathcal{N}$	0.1	0.0580	0.060	-0.0000	0.064
$\mathcal{N}$	0.2	0.0608	0.086	0.0001	0.106
$\mathcal{N}$	0.3	0.0716	0.286	0.0009	0.430
$\mathcal{N}$	0.5	0.1176	1.000	0.0051	1.000

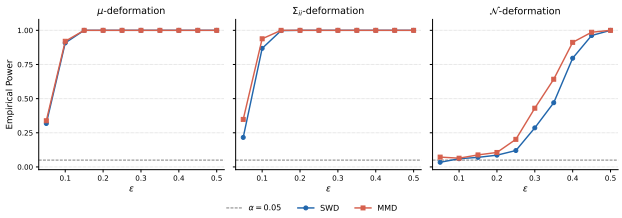

Figure 3. Empirical power of SWD (blue circles) and MMD (red squares) as a function of deformation strength ε, for three deformation types applied to JetNet gluon jets ($n = 1{,}000$, $\alpha = 0.05$, 500 repetitions). The dashed horizontal line marks the nominal level $\alpha = 0.05$. Both metrics detect μ- and Σ_{ii}-deformations reliably at moderate ε, while $\mathcal{N}$-deformation remains largely undetected below $\varepsilon = 0.3$, highlighting the limited sensitivity of fixed-kernel methods to additive Gaussian smearing at small scales.

$\varepsilon = 0.5$. This asymmetry directly reflects the distributional challenge described above—additive Gaussian smearing affects primarily the tails of heavy-tailed distributions such as p_T and m_j, precisely where fixed-kernel methods with median-heuristic bandwidth lose sensitivity.

4. Open Challenges and Opportunities

The application of modern hypothesis testing methods to MC evaluation in particle physics raises a number of open challenges that we believe represent compelling research directions for both the physics and the statistics communities.

Kernel selection for structured physics distributions. `MCBench` currently relies on a fixed Gaussian kernel with median-heuristic bandwidth selection. While this provides a reasonable default, it is known to be suboptimal for distributions with complex correlation structure or multiple scales (Gretton et al., 2012b). Methods for learned or adaptive kernel selection, such as those based on deep kernels (Liu et al., 2020) or aggregated tests with diverse kernels (Zhou et al., 2026), are natural candidates to improve sensitivity. Classical alternatives include the finite-dimensional witness function test (Jitkrittum et al., 2016), the energy-based classifier two-sample test (Lopez-Paz & Oquab, 2016), and bandwidth selection procedures for large-scale MMD (Sutherland & Schneider, 2015). A key open question is: *which kernel families are most sensitive to physically meaningful distributional differences in high-energy physics data?*

Sensitivity under heavy-tailed distributions. Many physically relevant distributions in high-energy physics, including jet transverse momentum spectra and cross-section distributions, exhibit heavy-tailed behavior. Standard metrics such as SWD and MMD with fixed kernels may exhibit reduced sensitivity to discrepancies concentrated in the tail regions of such distributions, where rare but physically relevant events reside. Developing two-sample tests with provable sensitivity guarantees under heavy-tailed alternatives, and understanding the role of kernel or projection choice in capturing tail discrepancies, represents an important open problem for physics applications.

Testing with non-IID samples. MC samples generated by Markov Chain Monte Carlo (MCMC) algorithms are not independently and identically distributed: they exhibit autocorrelation that deflates the effective sample size. Standard two-sample tests assume IID observations and may be miscalibrated in this setting. Developing tests with valid type-I error control under dependent sampling is an important open problem. A concrete instance arises in Bayesian parameter estimation in particle physics: as shown in the `MCBench` walkthrough example (Ding et al., 2025), Metropolis-Hastings samples exhibit autocorrelation that leads to a systematic overestimation of the effective sample size, resulting in miscalibrated SWD and MMD statistics relative to their IID baselines.

Calibrated p-values and decision criteria. In practice, physicists require a single, interpretable decision criterion, a calibrated p-value or an explicit accept/reject decision, rather than a visual comparison of metric distributions. This practical requirement motivates the development of well-calibrated hypothesis tests with explicit Type-I error control for the physics setting. The connection between modern testing frameworks, including anytime-valid inference and e-values (Ramdas et al., 2023), and the needs of physics analyses represents a largely unexplored bridge between the two communities.

High-dimensional scalability. Real detector data from experiments such as ATLAS can involve $\mathcal{O}(10^2)$-dimensional observations with millions of events. Scalable two-sample testing methods that maintain statistical power in this regime, while remaining computationally tractable, are essential for practical physics applications.

Finally, the current implementation of MCBENCH focuses on kernel-based metrics (SWD and MMD); classifier two-sample tests (Lopez-Paz & Oquab, 2016), which use the classification accuracy of a trained discriminator as a test statistic, represent a natural and practically important extension that we plan to incorporate in future versions of the benchmark.

An open invitation. `MCBench` provides open-source infrastructure, a growing library of physics-motivated test cases, and a pathway to evaluate methods on selected real or collaboration-approved detector datasets. We invite the hypothesis testing community to integrate their methods into this framework, evaluate them on physics-relevant benchmarks, and engage with the open problems described above. Code and data access are available at `https://github.com/tudo-physik-e4/MCBench`.

Acknowledgements

The authors are supported by the German Federal Ministry of Education and Research (BMBF) in the ErUM-Data action plan via the KISS consortium (Verbundprojekt 05D2022).

Impact Statement

This paper presents work whose goal is to advance the field of Machine Learning, specifically in the context of statistical hypothesis testing for scientific simulation validation. The primary application domain is particle physics, where improved evaluation methods for Monte Carlo simulators may benefit the broader scientific community. There are no potential societal consequences which we feel must be specifically highlighted here.

References

Bellm, J., Gieseke, S., Grellscheid, D., Plätzer, S., Rauch, M., Reuschle, C., Richardson, P., Schichtel, P., Seymour, M. H., Siódmok, A., et al. Herwig 7.0/herwig++ 3.0 release note. *The European Physical Journal C*, 76(4): 196, 2016.

Ding, Z., Grunwald, C., Ickstadt, K., Kröninger, K., and La Cagnina, S. Mcbench: A benchmark suite for monte carlo sampling algorithms. *arXiv preprint arXiv:2501.03138*, 2025.

Gretton, A., Borgwardt, K., Rasch, M., Schölkopf, B., and Smola, A. A kernel method for the two-sample-problem. *Advances in neural information processing systems*, 19, 2006.

Gretton, A., Borgwardt, K. M., Rasch, M. J., Schölkopf, B., and Smola, A. A kernel two-sample test. *The journal of machine learning research*, 13(1):723–773, 2012a.

Gretton, A., Sejdinovic, D., Strathmann, H., Balakrishnan, S., Pontil, M., Fukumizu, K., and Sriperumbudur, B. K. Optimal kernel choice for large-scale two-sample tests. *Advances in neural information processing systems*, 25, 2012b.

Grossi, S., Letizia, M., and Torre, R. Refereeing the referees: evaluating two-sample tests for validating generators in precision sciences. *Machine Learning: Science and Technology*, 6(1):015052, 2025.

Jitkrittum, W., Szabó, Z., Chwialkowski, K. P., and Gretton, A. Interpretable distribution features with maximum testing power. *Advances in Neural Information Processing Systems*, 29, 2016.

Kansal, R., Pareja, C., Hao, Z., and Duarte, J. Jetnet: A python package for accessing open datasets and benchmarking machine learning methods in high energy physics. *Journal of Open Source Software*, 8(90):5789, 2023.

Liu, F., Xu, W., Lu, J., Zhang, G., Gretton, A., and Sutherland, D. J. Learning deep kernels for non-parametric two-sample tests. In *International conference on machine learning*, pp. 6316–6326. PMLR, 2020.

Lopez-Paz, D. and Oquab, M. Revisiting classifier two-sample tests. *arXiv preprint arXiv:1610.06545*, 2016.

Magnusson, M., Torgander, J., Bürkner, P.-C., Zhang, L., Carpenter, B., and Vehtari, A. posteriordb: Testing, benchmarking and developing bayesian inference algorithms. *arXiv preprint arXiv:2407.04967*, 2024.

Panaretos, V. M. and Zemel, Y. Statistical aspects of wasserstein distances. *Annual review of statistics and its application*, 6(1):405–431, 2019.

Rabin, J., Peyré, G., Delon, J., and Bernot, M. Wasserstein barycenter and its application to texture mixing. In *International conference on scale space and variational methods in computer vision*, pp. 435–446. Springer, 2011.

Rahimi, A. and Recht, B. Random features for large-scale kernel machines. *Advances in neural information processing systems*, 20, 2007.

Ramdas, A., Grünwald, P., Vovk, V., and Shafer, G. Game-theoretic statistics and safe anytime-valid inference. *Statistical Science*, 38(4):576–601, 2023.

Schrab, A., Kim, I., Albert, M., Laurent, B., Guedj, B., and Gretton, A. Mmd aggregated two-sample test. *Journal of Machine Learning Research*, 24(194):1–81, 2023.

Sjöstrand, T., Ask, S., Christiansen, J. R., Corke, R., Desai, N., Ilten, P., Mrenna, S., Prestel, S., Rasmussen, C. O., and Skands, P. Z. An introduction to pythia 8.2. *Computer physics communications*, 191:159–177, 2015.

Sutherland, D. J. and Schneider, J. On the error of random fourier features. *arXiv preprint arXiv:1506.02785*, 2015.

Tran, B. T. and Schreuder, N. Minimax-optimal two-sample test with sliced wasserstein. *arXiv preprint arXiv:2510.27498*, 2025.

Zhou, Z., Tian, X., Peng, L., Lei, C., Schrab, A., Sutherland, D. J., and Liu, F. Dual: Learning diverse kernels for aggregated two-sample and independence testing. *Advances in Neural Information Processing Systems*, 38: 130661–130695, 2026.

A. Empirical Power of SWD and MMD on JetNet Gluon Jets

Table 2. Full empirical power results for all ε values. SWD and MMD evaluated on JetNet gluon jets ($n = 1{,}000$, $\alpha = 0.05$, 500 repetitions).

Deformation	ε	SWD (mean)	SWD power	MMD (mean)	MMD power
μ	0.05	0.0719	0.318	0.0007	0.338
μ	0.10	0.1052	0.908	0.0031	0.920
μ	0.15	0.1426	1.000	0.0068	1.000
μ	0.20	0.1850	1.000	0.0123	1.000
μ	0.25	0.2243	1.000	0.0186	1.000
μ	0.30	0.2682	1.000	0.0268	1.000
μ	0.35	0.3090	1.000	0.0361	1.000
μ	0.40	0.3517	1.000	0.0469	1.000
μ	0.45	0.3943	1.000	0.0589	1.000
μ	0.50	0.4380	1.000	0.0727	1.000
Σ_{ii}	0.05	0.0683	0.216	0.0007	0.348
Σ_{ii}	0.10	0.0954	0.868	0.0027	0.938
Σ_{ii}	0.15	0.1273	0.998	0.0058	1.000
Σ_{ii}	0.20	0.1622	1.000	0.0096	1.000
Σ_{ii}	0.25	0.2012	1.000	0.0148	1.000
Σ_{ii}	0.30	0.2370	1.000	0.0200	1.000
Σ_{ii}	0.35	0.2766	1.000	0.0264	1.000
Σ_{ii}	0.40	0.3146	1.000	0.0330	1.000
Σ_{ii}	0.45	0.3509	1.000	0.0396	1.000
Σ_{ii}	0.50	0.3907	1.000	0.0471	1.000
$\mathcal{N}$	0.05	0.0570	0.034	-0.0001	0.072
$\mathcal{N}$	0.10	0.0580	0.060	-0.0000	0.064
$\mathcal{N}$	0.15	0.0588	0.070	0.0000	0.088
$\mathcal{N}$	0.20	0.0608	0.086	0.0001	0.106
$\mathcal{N}$	0.25	0.0647	0.120	0.0004	0.202
$\mathcal{N}$	0.30	0.0716	0.286	0.0009	0.430
$\mathcal{N}$	0.35	0.0782	0.470	0.0014	0.642
$\mathcal{N}$	0.40	0.0896	0.796	0.0024	0.912
$\mathcal{N}$	0.45	0.1016	0.962	0.0035	0.986
$\mathcal{N}$	0.50	0.1176	1.000	0.0051	1.000