Consistency Tests of Kerr Black Hole Parameters Across Independent Observational Sectors Using a Covariance-Based Statistic

Aronno Mirdha* and Malte Mühlenfeld
Gymnasium Heissen, Mülheim an der Ruhr, Germany
(Dated: May 2, 2026)

We propose a statistical consistency test for Kerr black hole parameters inferred from independent observational sectors. In General Relativity, a stationary black hole spacetime is fully characterized by its mass and spin (M, a), so different measurements of the same object should yield consistent parameter estimates within their uncertainties. We combine sectoral estimates into a single parameter vector and define a residual with respect to a common best-fit value. Using the associated covariance matrix, we construct a quadratic form that measures the level of disagreement between sectors. Under standard Gaussian assumptions, this statistic follows a χ^2 distribution, which allows us to quantify the significance of any observed inconsistency. This framework provides a simple and general way to test whether independent observations are compatible with a single Kerr spacetime. We discuss the underlying assumptions and outline possible extensions, including applications to simulated or observational data.

I. INTRODUCTION

Black holes provide one of the most direct tests of General Relativity in the strong-field regime [1, 2]. In the case of stationary black holes, the Kerr solution predicts that the spacetime is fully characterized by only two parameters, mass M and spin a [3].

Different observational methods probe these parameters through distinct physical processes. Gravitational wave signals constrain the dynamical evolution of compact objects [4], while imaging and orbital measurements probe the geometry of the spacetime in complementary regimes [5, 6]. Although these approaches rely on different data and modeling assumptions, they are expected to describe the same underlying spacetime.

In practice, independent measurements may differ due to statistical uncertainty, systematic effects, or limitations in the modeling of the data. This raises the question of how to assess whether such differences are consistent with a single Kerr spacetime.

We introduce a covariance-weighted statistical test that quantifies the consistency of Kerr parameter estimates across independent observational sectors. The method combines sectoral estimates into a unified parameter vector and evaluates their agreement using a quadratic form defined by the covariance matrix.

Under standard Gaussian assumptions, the resulting statistic follows a chi-square distribution, allowing the significance of any observed disagreement to be quantified in a well-defined way. More generally, the framework provides a simple approach to comparing parameter estimates obtained from different observational probes and can be applied to simulated data and, in principle, to future multi-messenger observations [7–9].

* aronnomirdha@outlook.com

II. FRAMEWORK

A. Parameter Estimates and Sectors

We consider a set of K observational sectors, each providing an estimate of the Kerr parameters based on distinct physical probes [2]. For each sector $k = 1, \ldots, K$, we define the parameter vector

$$\theta_k = \begin{pmatrix} M_k \\ a_k \end{pmatrix} \in \mathbb{R}^2,$$

where M_k and a_k denote the inferred mass and spin.

Each θ_k is treated as a random vector representing the outcome of an inference procedure, characterized by an associated covariance matrix Σ_k. This statistical description reflects the standard treatment of parameter estimation in gravitational-wave and astrophysical inference [2, 4]. Although the estimates arise from different observational methods, they are assumed to correspond to the same underlying spacetime under the null hypothesis of a common Kerr solution.

To analyze the sectoral estimates jointly, we define the stacked parameter vector

$$\Theta = \begin{pmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta_K \end{pmatrix} \in \mathbb{R}^{2K},$$

which embeds all sectoral estimates into a single parameter space.

This representation enables a unified statistical treatment of the agreement between independent sectoral estimates, forming the basis for the consistency test introduced in this work.

B. Common Parameter and Residuals

Under the assumption that all sectors probe the same underlying spacetime, we introduce a common parameter vector

$$\bar{\theta} = \begin{pmatrix} \bar{M} \\ \bar{a} \end{pmatrix} \in \mathbb{R}^2,$$

which represents a parameter value shared across all sectors under the null hypothesis of a common Kerr spacetime.

To compare the sectoral estimates to this common parameter, we define the vector

$$\bar{\Theta} = \begin{pmatrix} \bar{\theta} \\ \bar{\theta} \\ \vdots \\ \bar{\theta} \end{pmatrix} \in \mathbb{R}^{2K},$$

obtained by repeating $\bar{\theta}$ for each sector.

The residual vector is then defined as

$$r = \Theta - \bar{\Theta},$$

which measures the deviation of the observed sectoral estimates from the common parameter. By construction, $r = 0$ corresponds to perfect agreement between all sectors.

In this formulation, the common parameter $\bar{\theta}$ can be interpreted as the maximum-likelihood estimate under the assumption of a shared underlying spacetime. In the case of independent sectors with Gaussian uncertainties, this reduces to a covariance-weighted average of the individual estimates [2].

C. Covariance Structure

The statistical uncertainties associated with the stacked parameter vector $\Theta \in \mathbb{R}^{2K}$ are described by the covariance matrix

$$C = \mathrm{Cov}(\Theta) \in \mathbb{R}^{2K \times 2K},$$

which encodes the joint uncertainty of all sectoral parameter estimates.

In the absence of correlations between observational sectors, the covariance matrix takes a block-diagonal form,

$$C = \begin{pmatrix} \Sigma_1 & 0 & \cdots & 0 \\ 0 & \Sigma_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \Sigma_K \end{pmatrix},$$

where each block $\Sigma_k \in \mathbb{R}^{2\times 2}$ denotes the covariance matrix associated with sector k.

Each block is given by

$$\Sigma_k = \begin{pmatrix} \sigma_{M,k}^2 & \mathrm{Cov}(M_k, a_k) \\ \mathrm{Cov}(a_k, M_k) & \sigma_{a,k}^2 \end{pmatrix},$$

which encodes the variances of the mass and spin estimates as well as their correlation within the sector. This representation corresponds to the standard Gaussian approximation of parameter uncertainties in astrophysical inference [2].

This block-diagonal structure reflects the assumption that the observational sectors are statistically independent. More generally, correlations between sectors can be incorporated by allowing for non-zero off-diagonal blocks in C, resulting in a fully populated covariance matrix.

In all cases, C is symmetric and positive definite, ensuring that the inverse C^{-1} exists and that the consistency statistic is well-defined.

D. Consistency Statistic

To quantify the level of agreement between sectors, we define the quadratic form

$$T^2 = r^T C^{-1} r,$$

where $r \in \mathbb{R}^{2K}$ is the residual vector and $C \in \mathbb{R}^{2K\times 2K}$ is the covariance matrix of the stacked parameter vector.

This statistic measures the deviation of the sectoral estimates from the common parameter, weighted by their covariance. In particular, it defines a quadratic form induced by C^{-1}, which acts as a metric on the parameter space.

In this sense, T^2 corresponds to a covariance-weighted distance (Mahalanobis distance) between the sectoral estimates and the common parameter, accounting for both uncertainties and correlations [2]. Larger values of T^2 indicate stronger disagreement between sectors relative to their expected statistical variation.

Under the assumption that the residual vector is approximately Gaussian distributed, this statistic forms the basis for a chi-square consistency test, as discussed in the following section.

III. STATISTICAL INTERPRETATION

The statistic

$$T^2 = r^T C^{-1} r$$

provides a quantitative measure of the disagreement between sectoral parameter estimates. It corresponds to a covariance-weighted squared distance, commonly known as the Mahalanobis distance [10], evaluated with respect to the common parameter.

Its interpretation follows from the statistical properties of the residual vector r. Under the null hypothesis that

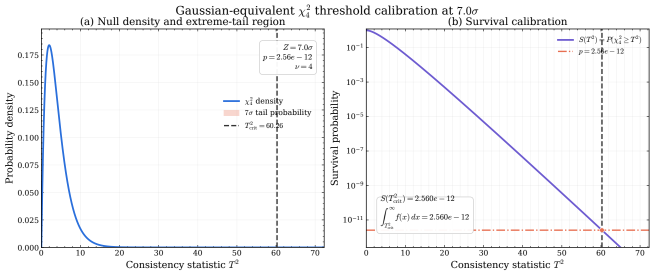

FIG. 1. Null distribution and survival function of the statistic T^2 for $\nu = 4$. Panel (a) shows the probability density, while panel (b) shows the corresponding survival function on a logarithmic scale. The vertical dashed line indicates the critical value corresponding to a chosen significance level, and the shaded region highlights the corresponding tail region.

all sectors are consistent with a single underlying parameter, and assuming that the inferred parameters can be approximated as Gaussian distributed (Laplace approximation) [11], the residual vector is approximately distributed as a multivariate normal distribution centered at zero. In this case, the quadratic form T^2 follows a chi-square distribution,

$$T^2 \sim \chi^2_\nu,$$

where ν denotes the number of degrees of freedom.

For K sectors with p parameters each, the number of degrees of freedom is given by

$$\nu = (K-1)p,$$

corresponding to the dimension of the residual subspace. In the case of three sectors and two parameters, this yields $\nu = 4$.

The chi-square distribution provides a natural way to assess the statistical significance of an observed value of T^2. Small values of T^2 indicate that the sectoral estimates are consistent with a single common parameter within their uncertainties, while large values correspond to statistically significant deviations.

As shown in Fig. 1, the interpretation of T^2 is governed by the tail of the distribution. The survival function $P(\chi^2_\nu \geq T^2)$ directly quantifies the probability of observing a value at least as large as T^2 under the null hypothesis. Large values of T^2 correspond to the extreme tail of the distribution and therefore indicate that the observed deviations are unlikely to arise from statistical fluctuations alone.

Given a chosen significance level, the observed value of T^2 can be compared to the corresponding critical value of the chi-square distribution. If T^2 exceeds this threshold, the null hypothesis of consistency is disfavored at the chosen significance level.

It is important to note that this interpretation relies on the Gaussian approximation of the underlying parameter distributions. In cases where this approximation breaks down, the chi-square behavior may no longer hold exactly, and the resulting p-values should be interpreted with caution.

Finally, the statistic does not identify the origin of any inconsistency. A large value of T^2 may arise from statistical fluctuations, systematic effects, or deviations from the underlying model assumptions. Additional analysis is therefore required to determine the source of any observed discrepancy.

IV. SIMULATION STUDY

To illustrate the behavior of the statistic, we perform a Monte Carlo simulation using synthetic sectoral estimates. The goal is to verify that the statistic follows the expected distribution under the null hypothesis and responds quantitatively to controlled inconsistencies [2].

We choose a fiducial Kerr parameter vector

$$\theta_* = \begin{pmatrix} M_* \\ a_* \end{pmatrix} = \begin{pmatrix} 10 \\ 0.6 \end{pmatrix}.$$

For each sector k, synthetic estimates are generated according to

$$\theta_k = \theta_* + \epsilon_k, \qquad \epsilon_k \sim \mathcal{N}(0, \Sigma_k),$$

corresponding to the standard Gaussian (Laplace) approximation of parameter uncertainties used in astrophysical inference [2, 4].

For simplicity, we adopt diagonal covariance matrices,

$$\begin{aligned} \Sigma_{\rm insp} &= \begin{pmatrix} 0.30^2 & 0 \\ 0 & 0.05^2 \end{pmatrix}, \\ \Sigma_{\rm ring} &= \begin{pmatrix} 0.40^2 & 0 \\ 0 & 0.06^2 \end{pmatrix}, \\ \Sigma_{\rm img} &= \begin{pmatrix} 0.50^2 & 0 \\ 0 & 0.07^2 \end{pmatrix}. \end{aligned} \tag{1}$$

For each realization, we compute the covariance-weighted estimate $\bar{\theta}$, the residual vector r, and the consistency statistic

$$T^2 = r^T C^{-1} r.$$

We generate $N = 2 \times 10^5$ realizations under the null hypothesis. In this case, the statistic is expected to follow a chi-square distribution with

$$\nu = (K-1)p = 4,$$

as discussed in Sec. III. The simulated distribution has mean

$$\langle T^2 \rangle \approx 3.99,$$

and the fraction of realizations exceeding the 95% reference value

$$\chi^2_{4,0.95} \approx 9.49$$

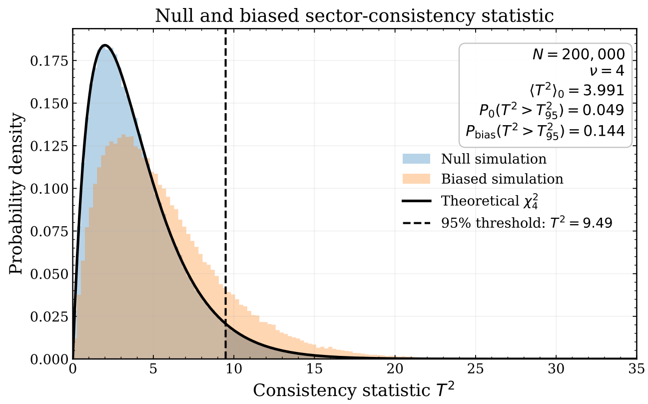

FIG. 2. Distribution of the consistency statistic T^2 obtained from Monte Carlo simulations with $N = 2 \times 10^5$ realizations. The blue histogram shows the null case, while the orange histogram corresponds to a biased simulation in which one sector is shifted by approximately one standard deviation. The solid black curve represents the theoretical χ^2_4 distribution, and the dashed line indicates the 95% reference value $T^2 \approx 9.49$. Under the null hypothesis, the simulated distribution is consistent with the expected χ^2 behavior, while the biased case exhibits a clear shift toward larger values.

is

$$P(T^2 > 9.49) \approx 0.049.$$

To probe the sensitivity of the statistic, we introduce a controlled bias in the imaging sector,

$$\theta_{\text{img}} = \theta_* + \Delta + \epsilon_{\text{img}}, \qquad \Delta = \begin{pmatrix} 0.50 \\ 0.07 \end{pmatrix}.$$

This corresponds approximately to a one-sigma shift in both parameters. In this biased case, the distribution of T^2 shifts toward larger values, and the exceedance fraction increases to

$$P(T^2 > 9.49) \approx 0.144,$$

as illustrated in Fig. 2.

These results confirm the expected behavior of the statistic: it reproduces the predicted chi-square distribution under the null hypothesis and provides a quantitative measure of inconsistency when sectoral biases are present.

V. DISCUSSION

Previous work has tested General Relativity by analyzing individual observational sectors or by comparing specific predictions of the Kerr solution [2, 6]. The present framework differs in that it provides a direct statistical test of the joint consistency of independent parameter estimates within a unified covariance-based approach.

The proposed framework provides a simple and general way to assess the consistency of Kerr parameter estimates across independent observational sectors. However, its interpretation relies on several assumptions that should be considered when applying the method.

A key assumption is that the inferred parameters can be approximated as Gaussian distributed. This corresponds to a Laplace approximation of the posterior near its maximum, in which the uncertainty regions are described by ellipses and fully characterized by their covariance matrices. Under this assumption, the statistic T^2 follows a chi-square distribution. In cases where the parameter distributions are strongly non-Gaussian, for example due to parameter degeneracies or low signal-to-noise ratios, this approximation may break down, and the resulting p-values should be interpreted with caution. In such situations, simulation-based calibration or methods based on full posterior samples may provide a more accurate characterization.

The framework also assumes that the covariance matrix is accurately known. In practice, uncertainties in the covariance estimation can affect the weighting of deviations and therefore the resulting value of T^2. In particular, underestimating uncertainties may lead to artificially large values of the statistic, while overestimating them may reduce sensitivity to genuine inconsistencies.

In the simplest formulation, the covariance matrix is taken to be block diagonal, corresponding to independent observational sectors. However, correlations between sectors may arise in realistic settings, for example through shared data or modeling assumptions. Such correlations can be incorporated by extending the covariance matrix to include off-diagonal terms.

Another limitation is that the statistic quantifies the presence of inconsistency but does not identify its origin. A large value of T^2 may result from statistical fluctuations, systematic effects, or deviations from the underlying model. Further analysis is therefore required to distinguish between these possibilities.

At present, no single astrophysical black hole has been measured with high precision across all observational sectors considered here. The framework is therefore primarily prospective. However, upcoming observations combining gravitational wave measurements (e.g. LISA, Einstein Telescope), imaging (ngEHT), and dynamical constraints may enable such consistency tests in the near future.

Despite these limitations, the framework provides a flexible and model-independent way to compare parameter estimates across different observational probes. More generally, it illustrates how consistency of spacetime geometry can be formulated as a quantitative statistical test, and may be extended to applications involving simulated or observational data.

VI. CONCLUSION

We have introduced a statistical framework to assess the consistency of Kerr black hole parameter estimates inferred from independent observational sectors. By combining sectoral measurements into a unified representation and evaluating their agreement through a covariance-weighted quadratic form, the method provides a simple way to quantify deviations relative to their uncertainties.

Under standard assumptions, the resulting statistic admits a well-defined interpretation in terms of a chi-square distribution, allowing for a direct assessment of statistical significance. The framework is general and can be applied to any set of parameter estimates with associated covariance information.

While the method does not identify the origin of inconsistencies, it provides a clear diagnostic for detecting them. Future work may include applications to simulated data, investigation of non-Gaussian effects, and exploration of connections to observational datasets.

Appendix A: Covariance-weighted estimate of the common parameter

The common parameter $\bar{\theta}$ can be obtained by minimizing the quadratic form

$$T^2 = (\Theta - \bar{\Theta})^T C^{-1} (\Theta - \bar{\Theta}),$$

with respect to $\bar{\theta}$. This corresponds to a weighted least-squares or maximum-likelihood estimate under the assumption of Gaussian uncertainties [2].

In the case of independent sectors with block-diagonal covariance, this minimization reduces to a covariance-weighted combination of the sectoral estimates. The solution is given by

$$\bar{\theta} = \left(\sum_{k=1}^{K} \Sigma_k^{-1} \right)^{-1} \left(\sum_{k=1}^{K} \Sigma_k^{-1} \theta_k \right),$$

where Σ_k denotes the covariance matrix of sector k.

This expression shows that each sector contributes to the common parameter with a weight determined by its inverse covariance, such that more precise measurements have a stronger influence on the combined estimate. In this sense, $\bar{\theta}$ represents the optimal combined estimate under the Gaussian approximation.

Appendix B: Explicit form of the covariance matrix

For three sectors, the covariance matrix is given by

$$C = \mathrm{Cov}(\Theta) \in \mathbb{R}^{6\times 6},$$

where $\Theta \in \mathbb{R}^6$ denotes the stacked parameter vector.

In the case of statistically independent sectors, C takes a block-diagonal form,

$$C = \begin{pmatrix} \Sigma_{\mathrm{insp}} & 0 & 0 \\ 0 & \Sigma_{\mathrm{ring}} & 0 \\ 0 & 0 & \Sigma_{\mathrm{img}} \end{pmatrix},$$

with each block $\Sigma_k \in \mathbb{R}^{2\times 2}$ representing the covariance matrix associated with sector k.

Each block is explicitly given by

$$\Sigma_k = \begin{pmatrix} \sigma_{M,k}^2 & \mathrm{Cov}(M_k, a_k) \\ \mathrm{Cov}(a_k, M_k) & \sigma_{a,k}^2 \end{pmatrix},$$

which encodes the second moments of the parameter estimates within each sector.

The block-diagonal structure follows from the assumption that the sectoral estimators are statistically independent, such that

$$\mathrm{Cov}(\theta_i, \theta_j) = 0 \quad \text{for } i \neq j.$$

More generally, correlations between sectors can be incorporated by allowing non-zero off-diagonal blocks,

$$C = \begin{pmatrix} \Sigma_1 & \Sigma_{12} & \cdots \\ \Sigma_{21} & \Sigma_2 & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix},$$

where $\Sigma_{ij} = \mathrm{Cov}(\theta_i, \theta_j)$.

In all cases, the covariance matrix C is symmetric and positive definite, ensuring the existence of C^{-1} and the well-definedness of the quadratic consistency statistic.

[1] C. M. Will, The confrontation between general relativity and experiment, Living Reviews in Relativity **17**, 4 (2014).

[2] E. B. et al., Testing general relativity with present and future astrophysical observations, Classical and Quantum Gravity **32**, 243001 (2015).

[3] R. P. Kerr, Gravitational field of a spinning mass as an example of algebraically special metrics, Physical Review Letters **11**, 237 (1963).

[4] B. P. A. et al., Observation of gravitational waves from a binary black hole merger, Physical Review Letters **116**, 061102 (2016).

[5] E. H. T. Collaboration, First m87 event horizon telescope results. i. the shadow of the supermassive black hole, The Astrophysical Journal Letters **875**, L1 (2019).

[6] D. P. et al., Gravitational test beyond the first post-newtonian order with the shadow of the m87 black hole, Physical Review Letters **125**, 141104 (2020).

[7] P. A.-S. et al., Laser interferometer space antenna, arXiv preprint (2017), 1702.00786.

[8] M. E. et al., A horizon study for cosmic explorer: Science, observatories, and community, arXiv preprint (2021), 2109.09882.

[9] M. D. J. et al., The next generation event horizon telescope, arXiv preprint (2023), 2304.03402.

[10] P. C. Mahalanobis, On the generalised distance in statistics, Proceedings of the National Institute of Sciences of India **2**, 49 (1936).

[11] C. Cutler and Éanna E. Flanagan, Gravitational waves from merging compact binaries: How accurately can one extract the binary's parameters from the inspiral waveform?, Physical Review D **49**, 2658 (1994).