

A Falsifiable Statistical Framework for Constraint-Sensitive Residual Structure in Gravitational-Wave Ringdown

Geno Marcovici

Continuum Innovative, LLC Mesa, Arizona, USA

*E-mail: dr.marcovici@continuuminnovative.com

(Received 14 August 2026, Accepted 27 September 2026, Published 09 Oct 2026)

Abstract

The interpretation of post-merger gravitational-wave residuals remains constrained by detector noise, waveform uncertainty, and the lack of a general statistical framework to assess whether residual organization exhibits reproducible structure beyond conventional stochastic expectations. Existing analyses provide powerful tools for waveform reconstruction and model comparison, yet offer no model-independent methodology for characterizing the statistical organization of residual behavior itself.

We hypothesize that if post-merger residuals exhibit reproducible organization beyond stochastic expectations, this behavior may reflect sensitivity to underlying consistency constraints that are not fully captured by local semiclassical descriptions. This hypothesis is neither assumed nor established by the present work; rather, it motivates the development of a statistically explicit framework capable of empirical evaluation.

To that end, we introduce a partition-function-based formalism that constructs an ensemble of statistically admissible residual configurations and derives an ensemble response quantity—the Support Sensitivity (χ)—together with an operational diagnostic, the Ringdown Structure Score (SRD). The resulting framework provides a model-independent statistical protocol for identifying candidate residual organization while remaining agnostic regarding its physical origin.

The principal contribution of this work is methodological rather than ontological. Rather than attributing residual structure to new gravitational physics, the proposed framework establishes a falsifiable statistical architecture for determining whether post-merger residuals warrant explanation beyond conventional stochastic models.

Keywords: Gravitational waves; ringdown; statistical inference; residual analysis; black-hole spectroscopy; model-independent methods

1. Introduction

The direct observation of gravitational waves has transformed strong-field gravity from a predominantly theoretical discipline into one increasingly subject to empirical investigation [1-3]. As detector sensitivity continues to improve, attention has progressively shifted from the primary inspiral-merger signal toward the increasingly subtle information potentially contained within post-merger residuals. Ringdown analyses and black-hole spectroscopy have become increasingly important tools for probing remnant compact objects and testing general relativity in the strong-field regime [4-9]. These residuals are generally interpreted in terms of detector noise, waveform inaccuracies, calibration uncertainty, or instrumental systematics, and substantial methodological advances have been made in modeling and discriminating among these conventional explanations [2, 10-13].

Despite these advances, an important methodological question remains unresolved. Existing analyses are principally designed to compare observed signals with specific waveform models or competing physical hypotheses; they provide no general statistical framework for evaluating whether residual organization itself exhibits reproducible structure beyond conventional stochastic expectations, independent of any particular microscopic theory of gravity. Importantly, the absence of such a framework should not be interpreted as a limitation of existing methodologies, which are optimized for waveform reconstruction, parameter estimation, and model comparison. Rather, it reflects a distinct statistical question that has received comparatively little attention.

The present work is motivated by the hypothesis that if statistically significant organization were shown to persist within post-merger residuals after conventional explanations have been rigorously excluded, such behavior might be consistent with sensitivity to underlying consistency constraints not fully represented by local semiclassical descriptions. This hypothesis is intentionally speculative and is neither assumed nor established by the analysis presented here. Rather, it serves to motivate the development of a statistical framework capable of determining whether such organization can be identified in a rigorous and falsifiable manner.

Our objective is therefore methodological rather than explanatory. We do not attempt to attribute residual structure to new gravitational physics, nor do we advocate any particular microscopic description of spacetime. Instead, we seek to construct a model-independent statistical formalism for characterizing residual organization should it exist. By separating the statistical detection of organized residual behavior from subsequent questions regarding its physical interpretation, the framework preserves a clear distinction between empirical observation and theoretical explanation.

To accomplish this, we replace the observed post-merger residual with an ensemble of statistically admissible residual configurations consistent with observational constraints. This ensemble construction naturally gives rise to an effective partition function from which ensemble response quantities may be derived. We introduce the Support Sensitivity (χ) as a measure of the responsiveness of the admissible ensemble to controlled perturbations of the underlying statistical

model and define the Ringdown Structure Score (SRD) as an operational diagnostic for identifying candidate residual organization within gravitational-wave observations.

The resulting framework is explicitly agnostic regarding the physical origin of any detected structure. Elevated values of SRD do not, by themselves, imply departures from general relativity or constitute evidence for new gravitational phenomena. Rather, they identify candidate observations warranting further investigation using established analytical techniques and competing physical hypotheses. In this sense, the proposed formalism functions as a statistical screening methodology whose conclusions remain subject to conventional scientific validation.

The principal contribution of this work is therefore the introduction of a mathematically explicit and empirically falsifiable statistical architecture for investigating organized structure within post- merger residuals. Whether such organization ultimately reflects conventional sources, previously unrecognized systematic effects, or genuinely novel physical phenomena lies beyond the scope of the present work and must be resolved through subsequent empirical investigation.

2. Statistical Framework

2.1 Conceptual Motivation

Conventional gravitational-wave analyses evaluate a single observed residual realization relative to an assumed waveform model and detector-noise description. This approach is both appropriate and highly successful for waveform reconstruction, parameter estimation, and model comparison. However, it does not naturally provide a statistical description of the broader space of residual configurations that remain compatible with the observational constraints.

The present framework adopts a complementary perspective. Rather than treating the observed residual as a unique statistical object, we consider the ensemble of residual configurations that remain statistically admissible under the adopted observational model. This construction does not imply that multiple physical residuals are simultaneously realized. Instead, the ensemble serves as an effective statistical representation of the admissible residual space, permitting ensemble-level response quantities to be defined independently of any specific microscopic interpretation. This shift in perspective—from a single observed realization to an ensemble of statistically admissible configurations—constitutes the central mathematical construction of the present framework.

2.2 Admissible Residual Ensemble

Let $d(t)$ denote the observed detector strain and $h_{\text{best}}(t)$ the maximum-likelihood waveform inferred under the adopted waveform model. The residual strain is defined as

$$r(t) = d(t) - h_{\text{best}}(t) \tag{1}$$

Within conventional analyses, the residual is generally interpreted in terms of detector noise, waveform uncertainty, calibration effects, and other instrumental contributions. These explanations constitute the baseline hypotheses throughout the present work.

Rather than treating the observed residual as the sole object of inference, we define an ensemble of statistically admissible residual configurations,

$$E = \{r_i(t)\} \quad (2)$$

where each configuration satisfies the observational constraints imposed by the adopted statistical model.

Importantly, this ensemble is not interpreted as a collection of physically realized signals. Rather, it is an effective statistical construction that characterizes the admissible organization of residual configurations consistent with the available observations.

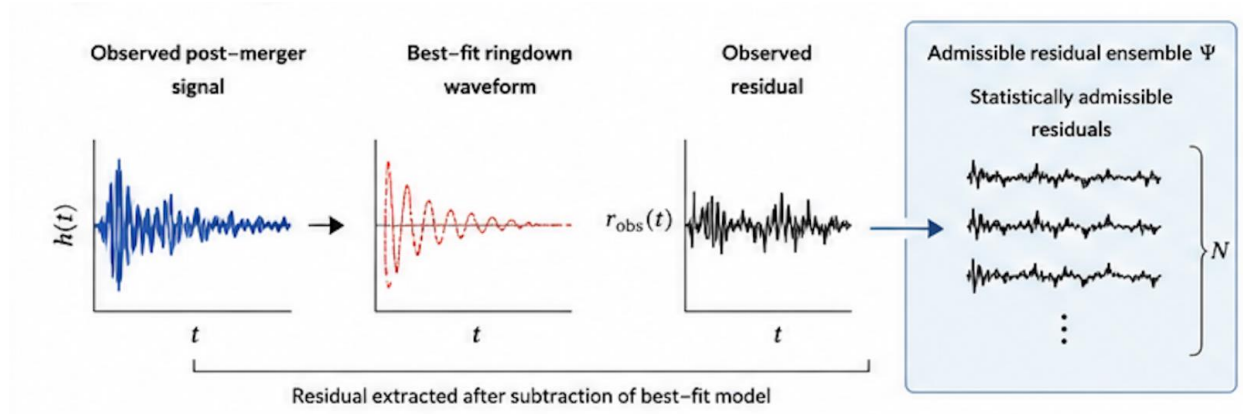


Figure 1. Conceptual construction of the admissible residual ensemble.

A single observed post-merger residual is embedded within an ensemble of statistically admissible residual configurations constructed under the adopted statistical model. The ensemble forms the statistical foundation for the subsequent ensemble-response analysis.

2.3 Effective Partition Function

The admissible ensemble naturally admits an effective statistical weighting through the partition function

$$Z = \int \exp \left[-\frac{S_{\text{eff}}}{\sigma^2} \right] dr \quad (3)$$

where S_{eff} denotes an effective statistical action, and σ^2 characterizes the variance scale of the adopted statistical model.

The partition function is introduced solely as an ensemble-generating construction. It is not intended to represent a microscopic description of gravitational dynamics or to imply any particular theory of spacetime. Rather, it provides the statistical foundation from which ensemble response quantities may be derived.

2.4 Support Sensitivity

The effective free energy is defined conventionally as

$$F = -\sigma^2 \ln Z \quad (4)$$

We define the Support Sensitivity,

$$\chi = -\frac{\partial^2 F}{\partial(\sigma^2)^2} \quad (5)$$

which characterizes the responsiveness of the admissible ensemble to controlled perturbations of the adopted statistical model. Within the present framework, χ is interpreted as an ensemble-level response quantity rather than a property of any individual residual realization.

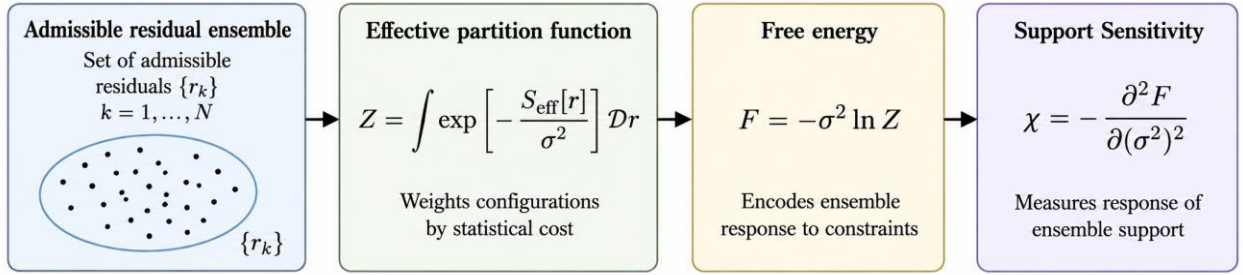


Figure 2. Derivation of the Support Sensitivity (χ) from the ensemble partition function.

The admissible residual ensemble gives rise to an effective partition function, from which the free energy and the Support Sensitivity χ are derived as ensemble-level statistical quantities.

2.5 Operational Motivation

Since the admissible ensemble is not directly observable, an operational statistic is required to evaluate organized residual behavior in observational data. This requirement motivates the introduction of the Ringdown Structure Score (SRD), developed in the following section.

3. Operational Construction of the Ringdown Structure Score

3.1 Motivation

The Support Sensitivity (χ) provides an ensemble-level response quantity describing the statistical behavior of admissible residual configurations. However, because the admissible ensemble itself is not directly observable, an operational statistic is required to evaluate individual gravitational-wave observations.

The Ringdown Structure Score (SRD) is introduced to bridge this gap. Rather than serving as a direct physical observable, it provides a statistical measure of the extent to which an observed post-merger residual exhibits organization relative to the admissible ensemble defined by the statistical framework developed in the preceding section.

Importantly, SRD is intended as a diagnostic statistic rather than a test of any specific theory of gravity. Elevated values identify observations whose residual organization may warrant further investigation, but they do not, by themselves, imply departures from general relativity or establish the presence of new gravitational phenomena.

3.2 Definition

Let $r(t)$ denote the observed post-merger residual after subtraction of the adopted best-fit waveform, and let $P(r)$ denote its probability density under the admissible ensemble generated by the partition function.

The Ringdown Structure Score is defined as

$$\text{SRD} = -\ln P(r) \quad (6)$$

or, more generally,

$$\text{SRD} = \Phi[r] \quad (7)$$

where Φ denotes any statistically calibrated functional that measures the degree of organization exhibited by the observed residual relative to the admissible ensemble.

The first expression represents the simplest realization of the statistic, whereas the second emphasizes that the framework accommodates alternative operational implementations, provided they preserve the underlying statistical interpretation.

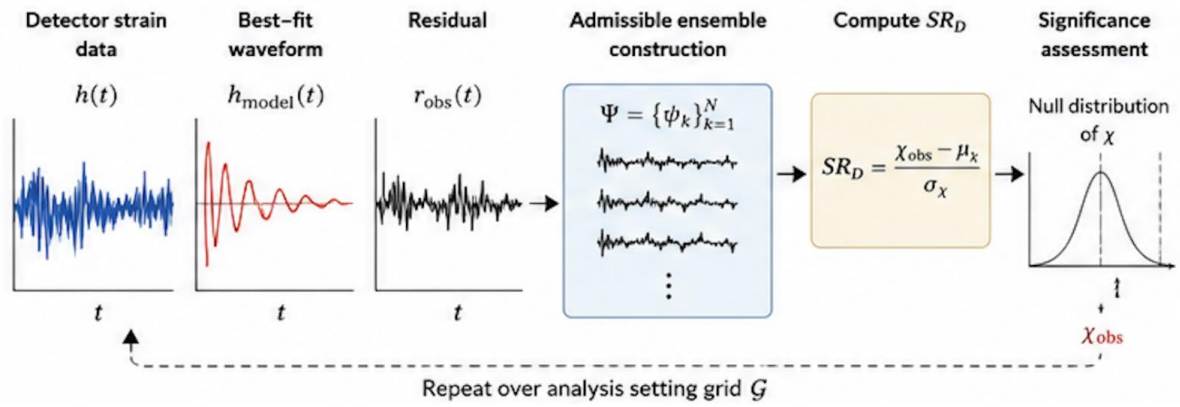


Figure 3. Workflow for computing the Ringdown Structure Score (SRD).

Operational workflow illustrating construction of the admissible residual ensemble, computation of the Support Sensitivity statistic, and calibration of the observed Ringdown Structure Score against the null distribution.

3.3 Interpretation

Within the proposed framework, small values of SRD indicate residual behavior that is statistically consistent with the admissible ensemble, whereas larger values correspond to increasingly uncommon residual organization under the adopted statistical model. Crucially, this interpretation is probabilistic rather than deterministic. An elevated Ringdown

Structure Score does not demonstrate the existence of new gravitational physics. Rather, it identifies observations that warrant closer examination using established waveform analyses, instrumental diagnostics, and competing physical explanations.

3.4 Statistical Calibration

In practical applications, the statistical significance of SRD is determined by comparison with synthetic residual ensembles generated under the null hypothesis of conventional stochastic behavior. Monte Carlo realizations, detector-noise simulations, surrogate waveform models, or bootstrap procedures may all be employed to estimate the expected distribution of SRD.

This calibration permits empirical significance levels to be assigned without requiring assumptions regarding the microscopic origin of any observed organization.

3.5 Decision Framework

The proposed methodology naturally leads to the following sequence of statistical inference:

1. Construct the admissible residual ensemble.
2. Calibrate the ensemble using the adopted statistical model.
3. Evaluate the Ringdown Structure Score.
4. Compare the observed score with the calibrated null distribution.
5. Investigate statistically significant outliers using conventional waveform analyses and competing physical hypotheses.

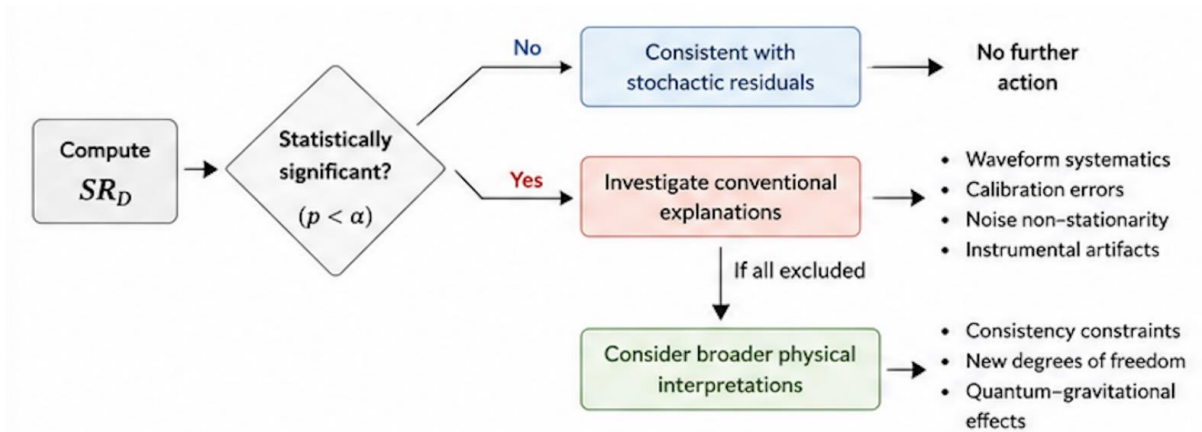


Figure 4. Decision framework for interpretation of residual structure.

Decision framework illustrating the separation of statistical detection from physical interpretation. Conventional waveform and instrumental explanations are evaluated before broader physical interpretations are considered.

This sequence deliberately separates statistical detection from physical interpretation. The framework therefore functions as a statistical screening methodology that identifies observations requiring additional investigation while remaining agnostic regarding the ultimate physical origin of any detected residual organization.

4. Discussion

The statistical framework developed in this work is designed to separate the detection of organized residual behavior from questions regarding its physical interpretation. This distinction is central to the present work: the proposed methodology neither assumes nor attempts to demonstrate the existence of physics beyond general relativity. Instead, it provides a model-independent statistical architecture for determining whether post-merger residuals exhibit reproducible organization warranting further investigation.

Recent advances in gravitational-wave data analysis have produced sophisticated approaches for waveform inference, coherent signal reconstruction, residual consistency testing, and black-hole spectroscopy [11-13, 4, 6-9]. The framework presented here is intended to complement these methodologies rather than replace them. Whereas existing approaches primarily evaluate waveform compatibility or residual consistency with specified models, the present work introduces an ensemble-based statistical framework for quantifying organized residual structure under controlled perturbation. To the best of our review, we have not identified an existing methodology combining this objective with the partition-function formalism, Support Sensitivity statistic, and Ringdown Structure Score developed here.

Should future analyses identify statistically significant residual organization using the Ringdown Structure Score, such observations would not constitute evidence for new gravitational physics. Rather, they would indicate that conventional explanations—including detector noise, waveform systematics, calibration effects, and other instrumental or methodological factors—require careful re-examination. Only after these explanations have been rigorously evaluated and found insufficient would broader physical interpretations warrant serious consideration.

Conversely, if the proposed framework consistently fails to identify statistically significant residual organization across observational datasets, that outcome would also be scientifically informative. Such a result would strengthen the interpretation that post-merger residual behavior is adequately described by existing stochastic models within the sensitivity limits of current observations. The framework is therefore explicitly falsifiable: both positive and negative outcomes contribute meaningfully to our understanding of gravitational-wave residuals.

An important feature of the proposed methodology is its independence from any specific microscopic theory of gravity. Neither the Support Sensitivity nor the Ringdown Structure Score presupposes a particular physical mechanism responsible for residual organization. Consequently, the framework remains applicable regardless of whether future investigations ultimately attribute organized residual behavior to conventional instrumental effects, improved waveform modeling, previously unrecognized statistical phenomena, or genuinely novel gravitational physics.

The present framework also has important limitations. Most notably, it has not yet been validated using observational gravitational-wave datasets, and its practical performance will depend upon the construction of appropriate null ensembles together with robust statistical calibration procedures. Future work will therefore focus on implementation, benchmarking against simulated and observational datasets, and evaluating the sensitivity and specificity of the proposed diagnostics under realistic detector conditions.

The principal contribution of this work is therefore methodological. By introducing an ensemble- based statistical architecture together with operational diagnostics calibrated against conventional null models, the framework provides a systematic and empirically testable approach for evaluating organized post-merger residual behavior. Whether future applications ultimately reveal statistically significant departures from stochastic expectations remains an empirical question to be resolved through empirical investigation rather than assumption.

5. Conclusions

We have introduced a model-independent statistical framework for evaluating organized behavior within post-merger gravitational-wave residuals. Rather than treating the observed residual as a single statistical realization, the proposed methodology constructs an ensemble of statistically admissible residual configurations from which ensemble response quantities and operational diagnostics may be derived.

Within this framework, the Support Sensitivity provides an ensemble-level measure of statistical responsiveness, while the Ringdown Structure Score serves as an operational diagnostic for evaluating individual observations relative to the admissible ensemble. Together, these quantities establish a statistically explicit methodology for identifying candidate residual organization without presupposing any particular microscopic description of gravity.

The proposed framework is intentionally agnostic regarding the physical origin of any detected structure. Elevated values of the Ringdown Structure Score do not constitute evidence for departures from general relativity or for new gravitational phenomena. Rather, they identify observations that warrant more detailed investigation using established analytical methods and competing physical hypotheses.

Since the methodology is explicitly falsifiable, both positive and negative empirical outcomes are scientifically informative. Agreement with conventional stochastic expectations would reinforce current interpretations of post-merger residual behavior, whereas reproducible departures from those expectations would motivate further empirical investigation.

The principal contribution of this work is not the assertion of new gravitational physics, but the introduction of a falsifiable statistical framework occupying a distinct methodological niche within gravitational-wave data analysis. Whether future application ultimately supports or rejects the presence of organized residual structure, the framework remains scientifically valuable because it converts that question into one that is quantitatively testable.

Acknowledgments: The author gratefully acknowledges the many investigators whose foundational contributions to gravitational-wave astronomy have made methodological studies such as the present work possible.

References

- [1] Abbott, B.P.; et al. (LIGO Scientific Collaboration and Virgo Collaboration). Observation of gravitational waves from a binary black hole merger. *Phys. Rev. Lett.* 2016, 116, 061102.
- [2] Abbott, B.P.; et al. (LIGO Scientific Collaboration and Virgo Collaboration). Tests of general relativity with GW150914. *Phys. Rev. Lett.* 2016, 116, 221101.
- [3] Abbott, R.; et al. (LIGO Scientific Collaboration, Virgo Collaboration, and KAGRA Collaboration). GWTC-3: Compact binary coalescences observed by LIGO and Virgo during the second part of the third observing run. *Phys. Rev. X* 2023, 13, 041039.
- [4] Dreyer, O.; Kelly, B.; Krishnan, B.; Finn, L.S.; Garrison, D.; Lopez-Aleman, R. Black-hole spectroscopy: Testing general relativity through gravitational-wave observations. *Class. Quantum Grav.* 2004, 21, 787–804.
- [5] Berti, E.; Cardoso, V.; Will, C.M. Gravitational-wave spectroscopy of massive black holes with the space interferometer LISA. *Phys. Rev. D* 2006, 73, 064030.
- [6] Berti, E.; Cardoso, V.; Starinets, A.O. Quasinormal modes of black holes and black branes. *Class. Quantum Grav.* 2009, 26, 163001.
- [7] Giesler, M.; Isi, M.; Scheel, M.A.; Teukolsky, S.A. Black hole ringdown: The importance of overtones. *Phys. Rev. X* 2019, 9, 041060.
- [8] Isi, M.; Giesler, M.; Farr, W.M.; Scheel, M.A.; Teukolsky, S.A. Testing the no-hair theorem with GW150914. *Phys. Rev. Lett.* 2019, 123, 111102.
- [9] Carullo, G.; van der Schaaf, L.; London, L.; Pang, P.T.H.; Tsang, K.W.; Hannuksela, O.A.; Meidam, J.; Agathos, M.; Samajdar, A.; Ghosh, A.; et al. Empirical tests of the black hole no-hair conjecture using gravitational-wave observations. *Phys. Rev. D* **2018**, 98, 104020.
- [10] Yunes, N.; Siemens, X. Gravitational-wave tests of general relativity with ground-based detectors and pulsar timing arrays. *Living Rev. Relativ.* 2013, 16, 9.
- [11] Cornish, N.J.; Littenberg, T.B. BayesWave: Bayesian inference for gravitational wave bursts and instrument glitches. *Class. Quantum Grav.* 2015, 32, 135012.
- [12] Veitch, J.; et al. Parameter estimation for compact binaries with ground-based gravitational-wave observations using the LAL Inference software library. *Phys. Rev. D* 2015, 91, 042003.
- [13] Thrane, E.; Talbot, C. An introduction to Bayesian inference in gravitational-wave astronomy. *Publ. Astron. Soc. Aust.* 2019, 36, e010.