nature portfolio

Peer Review File

GW250114 reveals signatures of post-merger black hole horizon

Corresponding Author: Dr Sizheng Ma

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Referee #1

(Remarks to the Author)
The article extracts a new signature of the black hole horizon from high-quality data recently obtained by the LIGO detectors. The study is original and marks an important milestone in the understanding of black hole binaries and gravity in its strongest and most dynamical regime. Most importantly, it adds a remarkable new piece of observational evidence to the hypothesis that the sources of LIGO signals are indeed Kerr black holes as predicted by General Relativity.

The data quality and methodology (both statistical and theoretical) are robust. The conclusions are similarly robust and for the vast majority justified (see below).

For these reasons, I believe the study can confidently meet Nature criteria of publication. However, before unconditionally recommending publication, there are a few clarifications that the authors should provide. I attach below an extensive set of comments and requests, split between major and minor. I would appreciate if the authors could consider them in detail.

Major comments

1. Fig.2: Filtering will depend on the remnant parameters (to obtain the frequencies), but the figure seems to show a unique result for the filtered strain. Why isn't there an uncertainty propagated in the strain from the uncertainty in the remnant parameters?

2. Eq.(4): To facilitate immediate reproducibility, the authors should release the exact routine used to compute omega_G, and its numerical output.

3. Fig.4: The NR results lie quite far away from the horizon mode (50% shift in the damping rate). This is in contrast to Fig.3, where the agreement seems to be much stronger. Even for the real data inference, the prediction at best lies at the tail of the posterior.

I understand this difference likely arises from the features observed in Fig.8 and from the inevitable "averaging" of the inference procedure adopted. However, I believe a stronger quantitative justification is needed here to convince the reader that the extracted feature indeed corresponds to the direct waves.

4. Several statements in the text rely on identifying a certain time before/after merger. But the merger time is not known for observational data. What is used in place of it, and how does its inference/uncertainty impact the conclusions?

5. A conceptual point remains to be addressed: the values of the remnant mass and spin of the model are fixed to a constant value, hence they correspond to the asymptotic ones. At the early times considered in this study, considerable emission of GWs is still present, which endows mass and spin with time dependence. The authors should quantify this effect and justify why it does not affect their results.

6. Pg.5 "We perform the analysis using the same time-domain Bayesian framework"

I would need to see more details of the inference process to be convinced of the robustness of the analysis. I ask the authors to provide in the response to the referee:

a. The waveform reconstruction (with error bands) for both the agnostic and informed analyses;

b. The posterior distributions on the parameters.

7. Page 9: "with marginal support for the (l = m = 2, n = 2) mode."

This sentence appears unjustified, as [41] states: "The presence of a second overtone 222 is not significantly preferred at any time." How can this be interpreted as marginal support?

How do the results change when the 222 is not removed? Since [41] reports no evidence for this mode, this technical choice appears questionable.

Minor comments

1. Abstract: "the quadrupole nature" --> "the dominant quadrupole nature"

2. Page 1: "Existing electromagnetic probes"

It is unclear what exactly the authors meant here, and this sentence would likely benefit from slight rephrasing. For example, measuring a light ring is not a direct measurement of an horizon (Ref.[36]).

3. Page 1: "it consists of a series" --> "it mainly consists of a series"

4. Page 2: "the first law of black hole thermodynamics."

Add a reference.

5. Page 2: "and informing waveform modeling [51–55]"

Other notable references that have precisely quantified this agreement are: arXiv:2112.12265, arXiv:2506.02189

6. Page 2: Was Fig.1 AI generated? Does this require a declaration/statement?

I would state explicitly this is a sketch.

7. Page 2: "These observable signals are referred to as"

Besides the contribution identified by the authors, there are additional recent references which have identified additional components. It would be fair to acknowledge them:

arXiv:2506.21668, arXiv:2510.17954, arXiv:2510.18956

8. Page 2: "effectively erasing the memory of the earlier inspiral."

This sentence needs to be justified.

9. Page 2: "until the companion finally crosses the horizon"

So far, the description has been kept at the linear level. At this level, the background is fixed, and the object isn't seen crossing the horizon by asymptotic observers (unlike what stated in this sentence), but only approaching it asymptotically. The authors should either explain what I am missing, or make the sentence more precise.

10. Page 3: "as the final signature of the infalling particle"

I do not think it is appropriate to describe direct waves as the final signature of the infalling particle.

The post-ringdown tails also depend on the particle motion (arXiv:2407.04682, arXiv:2406.17018). The orbital configuration is also known to be imprinted in the QNM amplitudes (e.g. arXiv:2503.09678). Both contributions appear later than direct waves.

11. Page 3: "with the merger time aligned to"

What is the definition of "merger time"? Afterwards it is referred to as peak of the strain. The precise definition should be given in this first instance.

12. Page 4: "follow common settings in quasinormal mode analyses"

A reference should be given, and why the same choice is sensible here should be motivated. I believe the authors checked that a longer segment length does not impact the result (supp. mat.). Should be stated here briefly.

13. Footnote 1:

a) So far, nonlinear effects have been shown to be subdominant (since the direct wave covers most of the merger signal). Why should they become important far away from the merger?

b) The near horizon approximation assumed to derive Eq.(4) was never mentioned before in the main text. Given that these caveats are important to interpret the measurement, this approximation should be mentioned around Eq.(4).

14. Page 5: "The minor dip and surrounding features near t = 0 are artifacts of the whitening process"

The meaning of this sentence is unclear. What in the whitening process is causing the "artifact"?

15. Page 5: "These results support the interpretation of direct waves as the “final sound”"

This sentence is misleading, for the same reasons I highlighted in Comment 10.

16. Eq.(5): this equation should be given above, when this is first mentioned

17. Page 6: "Finally, a systematic search for direct waves in existing gravitational-wave events would be highly valuable"

Why? What features can be observed that cannot already be observed from this loudest event?

18. Page 9: "Since the filter is applied in the frequency domain,"

Since the authors mentioned that the filter is roughly equivalent to a maximum likelihood subtraction when assuming white noise, it should be mentioned explicitly if (after transforming to the Fourier domain) whitening is applied before acting with the filter.

(Remarks on code availability)

Referee #2

(Remarks to the Author)
The manuscript investigates the possibility of interpreting a part of the merger signal from a binary black hole merger as a

"direct wave", and applies this interpretation to the binary merger event GW250114. The interpretation is interesting and may perhaps even be viewed as a compelling model of the merger. The paper is well written as appropriate for a broad scientific audience and the references are appropriate.

I have some comments below and I would ask the authors to address these before I can definitively recommend the paper.

1. The notions of direct waves are, in my view, the clearest for extreme mass ration systems, i.e. for the motion of a small particle around a massive central black hole. When the applied to systems of comparable mass systems such as GW250114 one needs to rely on the EOB formalism to recover such an interpretation. This is reasonable since this formalism has proved to be accurate and is clearly useful in modeling binary systems. Nevertheless, the results presented in this manuscript are framed as fundamental tests of General Relativity. Moreover they are applied to a very high SNR event where the requirements on waveform accuracy are correspondingly very stringent. It should therefore be asked whether these depend on the assumptions built into the EoB formalism or phrased alternatively, how far can the results presented here be viewed as a test of GR rather than a test of the EOB framework within GR?

2. The second set of questions related to the Ringdown analysis. As I understand the method use here, one follows the results of Ref. [41] and removes the 220, 221, and 222 QNMs. The QNMs specifically take the maximum likelihood (ML) values of the mass and spin reported in [41] (and presumably also the start time). How sensitive are the later results on the direct waves on these choices? Ideally, one might prefer a combined model with the direct waves and the QNMs all included. However as a second best approach, rather than taking the ML values, it might be useful to sample from the QNM posteriors found in [41]. In any case, can the authors quantify the effect of these uncertainties in the QNM modeling and in particular if these would affect the extraction of the direct wave in the next step of the analysis? Ideally, I would want to know the impact of these choices on the statistic used in the direct wave detection. As a related question, how important are the windowing (Gibbs phenomena) effects related to the start of the QNM, and could these affect the direct wave analysis?

(Remarks on code availability)
I have not reviewed the software used in this analysis, but for the most part the methods employed here are the standard ones used in the literature and I do not expect this to be of much concern.

Referee #3

(Remarks to the Author)
Summary of the key results

This manuscript reports the identification of a “direct wave” component of the gravitational-wave signal produced from the binary black-hole merger GW250114. The direct wave, emitted approximately simultaneously with the ringdown quasinormal modes, carries information about the horizon of the remnant black hole; namely, the horizon rotation frequency (which determines the frequency of the direct wave) and the surface gravity (which determines its decay timescale). It is expected that the direct wave frequency and decay rate will evolve over time due to screening from the surrounding potential barrier, ultimately tending towards the values of the "horizon mode”. This work performs a measurement of the (average) direct wave frequency and decay rate, demonstrating a broad consistency with the predicted horizon mode values.

Below I number comments which the authors may wish to respond to.

Originality and significance: if not novel, please include reference

This is the first time, to my knowledge, such a measurement has been attempted. I believe such a measurement will be of immediate interest to those who work in gravitational-wave astrophysics, with an emphasis on ringdown tests of general relativity and black-hole perturbation theory.

1. However, whilst the identification of the direct wave is certainly interesting in its own right, this work does not demonstrate how this feature can be used to test general relativity (despite the claim that this work “opens exciting avenues to test general relativity”) and I would be interested in the author's thoughts on this. For example, do the authors envisage a parametrised version of their analytic model used in Fig. 3 which would allow you to measure horizon properties? A more detailed proposal for how a test of general relativity could be performed could improve the manuscript and its impact.

My concern is that the direct wave may not open up tests of general relativity as the authors claim. Figure 4 in the manuscript helps to explain my reasoning: there is a straightforward prediction of the horizon mode frequency and damping rate, but this is not what is observed due to a screening from the potential barrier. So, any discrepancy between what is measured and what is predicted can be attributed to this screening. Unfortunately, it seems this screening (at least for binary mergers with comparable masses) is not sufficiently well understood for any predictions to be tested with data. This can be compared, for example, with tests of general relativity involving quasinormal modes: in the same figure the quasinormal mode frequencies and damping rates are also indicated. In contrast to the direct wave and horizon mode, the quasinormal mode frequencies/damping rates are exactly what is measured by an observer, thus allowing quantitative statements about the consistency of observations with theory. The author’s thoughts on what quantitative statements the direct wave and horizon mode could offer would help to see its potential impact.

Data & methodology: validity of approach, quality of data, quality of presentation

I found the manuscript well written and well presented. I believe the quality of the gravitational-wave data is of no concern; it corresponds to a high-SNR event which has been well studied (see Refs. [40,41] in the manuscript). The core of the method relies on an analysis method known as a quasinormal mode rational filter. This approach is well tested in terms of performing ringdown analysis and used in multiple works (including works by the LIGO Scientific, Virgo, and KAGRA Collaborations). I do, however, have some questions/concerns regarding the use of this method for the direct wave analysis:

2. Firstly, I have a concern that not all of the quasinormal modes in the data are being removed. Ref. [82] in the manuscript demonstrates that as many as nine overtones are excited early in the ringdown, but throughout this work only the first two overtones have been removed. It is true that the SNR of GW250114 only allows the first overtone amplitude to be measured confidently (with some tentative evidence for the second overtone), but I would argue that the higher overtones are still present in the data (even if not individually resolveable). Would it not make sense to remove contributions from all overtones before searching for the direct wave?

This is also true for NRSur7dq4 - in Fig. 3 only the first two overtones were removed, but presumably SNR is not an issue here and so higher overtones can certainly also be removed? In Fig. 3 we are presumably seeing some contribution from the higher overtones in the filtered data, so I am concerned the direct wave model is instead fitting to them. The author's comments on this issue would be appreciated.

3. Secondly, I was curious about the impact of the filter on the pre-ringdown data. As explained in Ref. [60] in the manuscript, applying the filter introduces a “flipped ringdown” to the data immediately preceding the ringdown. I was surprised to see this was not mentioned in the manuscript, especially since the work involves fitting data slightly at or slightly before the time of peak strain. Was it checked if an artificial flipped ringdown was being introduced to the data, and if this interfered with the fitting of the direct wave?

Regarding reproducibility:

4. I don’t believe the priors are provided on the frequency and damping rate used in the main direct-wave analysis (configuration files for the specific analyses carried out would help with reproducibility).

5. The authors should provide more details on the procedure for obtaining the detection threshold. For example, what priors are used when applying the analysis to noise? How many trials are performed?

6. It is also not clear how the analytic model for the direct wave is fitted to the NRSur7dq4 model (I believe this is referred to as the “benchmarked waveform” in the Methods). Is there an amplitude and phase being fitted for? The same applies to the waveform without screening shown in Fig. 9.

7. It is not explained how the sky-location marginalization is performed for the result in Fig. 5.

Finally:

8. The manuscript also makes use of the waveform model NRSur7dq4 to illustrate the presence of the direct wave. However, using a numerical relativity simulation would be more accurate, and since just one set of parameters are used for NRSur7dq4 I see no reason why a numerical relativity simulation couldn’t be used for the results in Figs. 3 and 9 (it may not make a significant difference, but since NRSur7dq4 is known to suffer from some loss of accuracy in the ringdown I believe it would at least be worth checking). The GW250114 companion paper, Ref. [41] in the manuscript, used the simulation SXS:BBH:3617.

Appropriate use of statistics and treatment of uncertainties

9. This analysis relies upon removing the quasinormal-mode content from the data before searching for the direct wave. This removal of quasinormal modes has some inherent uncertainty associated with it, which is not propagated throughout this analysis. Specifically, the filters used to remove the quasinormal modes are constructed using the remnant black hole mass and spin, which has some uncertainty. One could perhaps consider jointly searching for the direct wave and filtering out the quasinormal modes (drawing values for the remnant mass and spin from the posterior) - did the authors consider this? It may also be true that the uncertainty on the remnant mass and spin is sufficiently small to not impact the direct wave measurement shown in Fig. 4, but this is not stated in the manuscript.

This above also applies to the matched-filtering analysis used to obtain the direct wave SNRs.

10. I am also curious about what uncertainties are present in the construction of the analytic model used in Fig. 3. Certain parameters have to be chosen (such as the energy and angular momentum of the orbit), but it is not clear to me how much flexibility there is with picking them. The agreement between the analytic model and the filtered waveform is one of the key

pieces of evidence for the direct wave, so it is important to know how robust that agreement is to choices of parameters.

I believe the detection statistic is a suitable statistical test (provided some more details on the exact procedure are provided). A couple of other comments:

11. It would be more accurate to label the red line in Fig. 2 “Maximum-likelihood waveform reconstruction”, since the reconstruction would have some uncertainty associated with it.

12. It is not stated what level the contours correspond to in Fig. 5.

Conclusions: robustness, validity, reliability

Overall, provided the concerns above are addressed, I believe the results in Fig. 3 and Fig. 4 provide compelling evidence for the detection of a direct wave component (and I consider this to be the main argument for publication in Nature). The authors have provided some checks of the robustness of the result, but as explained above I believe additional checks should be performed. Finally, as explained in point 1, the impact of this result on tests of general relativity is not clear to me (especially when compared to the usual quasinormal-mode analysis) and could use more explanation in the manuscript.

Other comments

13. There are a couple of places in the manuscript where the time of peak strain is referred to as the merger time, but I would avoid saying this since the merger time is not very well defined — simply saying t=0 is set to the time of peak strain is sufficient.

14. I think Fig. 2 could be improved by providing the time axis in units of the (detector frame) maximum-likelihood remnant black hole mass (perhaps with a second axis at the top of the figure). This would aid with comparing Fig. 2 to Fig. 3.

15. It is stated that the inferred direct wave frequency does not fully asymptote to omega_H in part due to “the limited signal-to-noise ratio at late times”. However, we also see that the fits to NRSur7dq4 (plus symbols in Fig. 4) also don’t asymptote to omega_H, despite the SNR presumably not being an issue here. Clarification on this would be appreciated.

(Remarks on code availability)

Version 1:

Reviewer comments:

Referee #1

(Remarks to the Author)
Regarding the sky position marginalisation: the authors' reply is correct only under the assumption that the uncertainty in the absolute time of arrival (e.g. at the geocenter) is negligible, as its effect is distinct wrt time delays, now marginalised over via the sky position. This statement needs to be supported by a quantitative argument, or this uncertainty included as well.

Regarding 5.: This is an acceptable statement, but the fact that the correct inclusion of this effect is unknown, and the consequent assumption of ignoring it, should be explicitly made in the manuscript. The reader should be cautioned more explicitly about this ignorance, and that these statements are based on empirical numerical observations rather than a well-understood model.

6. I also request to add the waveform reconstruction to the manuscript, as it is very informative.

Once these points are incorporated, in my opinion the manuscript can proceed to publication.

(Remarks on code availability)
If the authors will release the scripts and data as promised in the report, I believe that this would allow for full reproducibility.

Referee #3

(Remarks to the Author)
I thank the authors for the detailed responses to my comments and the changes made to the paper, which address my main concerns. I am happy to recommend the paper for publication in Nature. Below I give additional comments on the paper.

The paper provides compelling evidence for the detection of a direct-wave in the GW250114 signal, and establishes its observational relevance. I believe the identification of the direct wave will have implications for waveform modeling, fitting ringdown quasinormal modes, and will open new avenues for testing general relativity. However, the paper does not perform any such tests. It is my understanding that to perform robust tests of general relativity more sophisticated calculations for the evolving direct-wave frequency omega_G are needed (I believe this is at least partly what the sentence “Such tests could be enabled … once theoretical and observational systematics are carefully accounted for” on page 6 is referring to). To reiterate, my concern is that the omega_G currently employed in the paper is calculated in the test-mass limit, and any discrepancy with general relativity could be attributed to applying a test-mass limit calculation to an equal-mass system (I apologise for incorrectly claiming the offset between the final horizon frequency and the measured frequency was due to the screening - my main point was that a robust prediction for the evolving frequency is required to test general relativity).

I believe the analysis method is robust and the conclusions are sound. I would just note (as also mentioned in the paper) that ideally one would fit for the ringdown quasinormal modes and direct wave simultaneously in a self-consistent inference. The current analysis, by effectively pre-processing the data to remove quasinormal modes, misses any correlations that exist between the quasinormal modes and direct wave. Related to my previous point, any high-precision analysis and test of general relativity should perform the fitting simultaneously. But, for the sake of establishing observational relevance, the current analysis is sufficient.

Finally, I thank the authors for the additions made to Fig. 5, which helps to make the conclusions robust. I just request clarification on one point: when the analysis is performed with different sky locations, is the geocenter peak time being fixed? In the main text below Fig. 5 it is stated the geocentric signal start time is varied, but this isn’t stated in the Fig. 5 caption. Instead in the caption it is stated that varying the sky location effectively propagates uncertainty in the peak time, but since the geocenter peak time will have its own uncertainty associated with it I am not sure if this is true.

(Remarks on code availability)

Version 2:

Reviewer comments:

Referee #1

(Remarks to the Author)
All changes look satisfactory, ready for publication.

(Remarks on code availability)

We thank the referee for the careful review and helpful comments. We have addressed all points in the revised manuscript (highlighted in blue) and summarize our responses below. In addition, we have refined parts of the analysis to adopt more physically motivated choices, e.g., using the maximum-a-posteriori values of the remnant mass and spin in the main analysis, which results in minor numerical updates reflected in the latest version of the manuscript.

Referee expertise:
Referee #1: LVK merger searches and waveform interpretation
Referee #2: GW from BH mergers, GR tests, numerical relativity
Referee #3: ringdown signals from BH mergers

Referees' comments:
==

Referee #1 (Remarks to the Author):

The article extracts a new signature of the black hole horizon from high-quality data recently obtained by the LIGO detectors. The study is original and marks an important milestone in the understanding of black hole binaries and gravity in its strongest and most dynamical regime. Most importantly, it adds a remarkable new piece of observational evidence to the hypothesis that the sources of LIGO signals are indeed Kerr black holes as predicted by General Relativity.

The data quality and methodology (both statistical and theoretical) are robust. The conclusions are similarly robust and for the vast majority justified (see below).

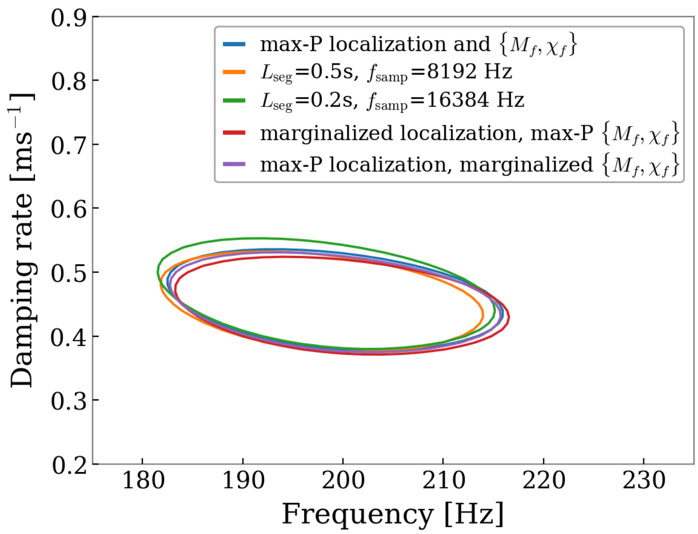
For these reasons, I believe the study can confidently meet Nature criteria of publication. However, before unconditionally recommending publication, there are a few clarifications that the authors should provide. I attach below an extensive set of comments and requests, split between major and minor. I would appreciate if the authors could consider them in detail.

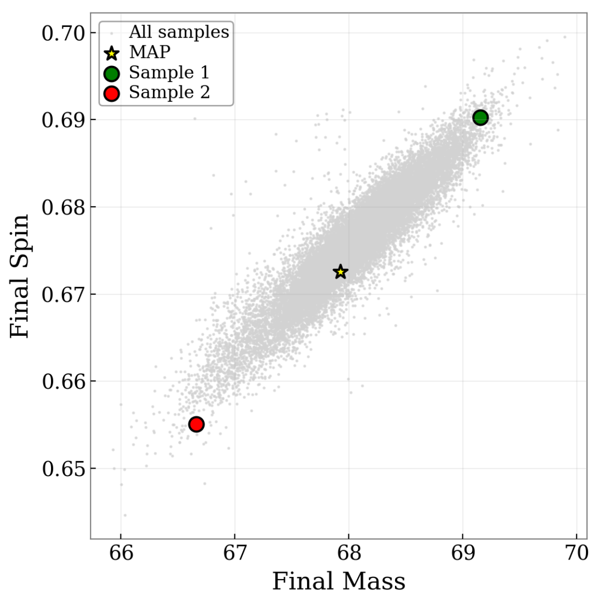
Major comments

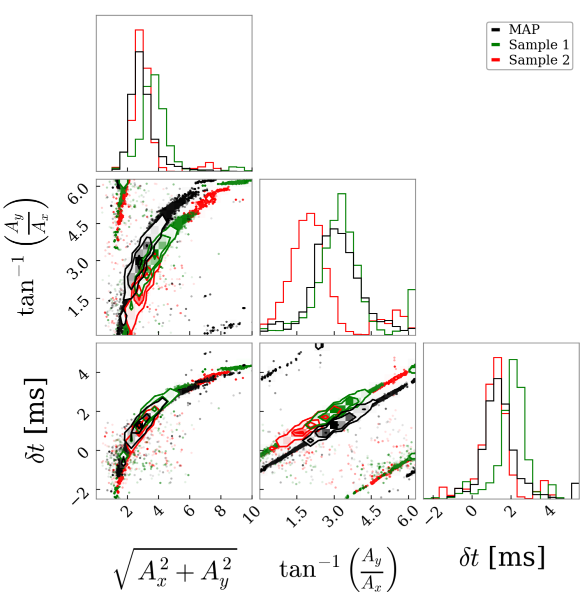
1. Fig.2: Filtering will depend on the remnant parameters (to obtain the frequencies), but the figure seems to show a unique result for the filtered strain. Why isn't there an uncertainty propagated in the strain from the uncertainty in the remnant parameters?

While the filtering depends on the remnant parameters, we have verified that propagating their uncertainty does not change the inferred results. Specifically, we have updated Fig. 5 in the manuscript (also shown below) to include results obtained by sampling the remnant mass and spin from the IMR posterior distribution. The inferred frequency and damping time in the agnostic search remains stable, demonstrating that uncertainty in the remnant parameters has a negligible impact on the filtered strain.

For the matched-filtering analysis, fully propagating the complete posterior distribution of the remnant mass and spin is computationally demanding. Instead, we provide a comparison plot showing the posterior distribution of the direct-wave parameters obtained using the maximum-a-posteriori (MAP) remnant properties, together with results derived from two representative samples drawn from the tails of the mass-spin posterior. The inferred parameters are consistent across these cases (see attached plots below).







2. Eq.(4): To facilitate immediate reproducibility, the authors should release the exact routine used to compute omega_G, and its numerical output.

We have attached the script used to compute \omega_G in Eq. (5) [previously Eq. (4)], along with the corresponding numerical output. Details are explained in Supplemental Material C. We are also in the process of preparing a public release of all relevant scripts and data, which will accompany the final version of the manuscript.

3. Fig.4: The NR results lie quite far away from the horizon mode (50% shift in the damping rate). This is in contrast to Fig.3, where the agreement seems to be much stronger. Even for the real data inference, the prediction at best lies at the tail of the posterior.

I understand this difference likely arises from the features observed in Fig.8 and from the inevitable "averaging" of the inference procedure adopted. However, I believe a stronger quantitative justification is needed here to convince the reader that the extracted feature indeed corresponds to the direct waves.

Indeed, in Fig. 4, the inferred parameters represent averages over finite time segments of the data. But since the signal amplitude decreases rapidly with time, these averages are mainly dominated by the earlier portion of each segment. As illustrated explicitly in Fig. 9 (Fig. 8 in the original version since we added a new figure), within the observational window considered in Fig. 4, the instantaneous frequency and damping rate still evolves dynamically and are therefore not expected to coincide with the horizon mode. The instantaneous values only asymptote to the horizon mode at much later times (t~20M), and notice that it is the regime when the screening effect starts to play a role (see Fig. 10) and makes the signal unobservable.

Within the observational window probed in Fig. 4, omega_G still continues to evolve, and the offset is therefore as expected (rather than an inconsistency of the model). The picture is further supported by the trend observed in Fig. 4: as the fitting window is slid to later times, the inferred decay rate increases. The behavior is fully consistent with the theoretical expectation in Fig. 9, where the increasing radial acceleration leads to a growing decay rate over time.

In the manuscript, Fig. 4 is intended as a largely model-agnostic, data-driven illustration of qualitative features of the direct wave, such as the systematic increase of the decay rate and the clustering of inferred parameters around the horizon mode. By construction, this approach does not aim to provide a precise quantitative match. *We have added clarifications and comparisons to Fig. 9 in the third last paragraph before Conclusion.*

To establish a stronger quantitative connection, we adopt a complementary matched-filtering analysis based on the physically motivated analytic template in Eq. (5) [previously Eq. (4)]. In this context, Fig. 3 plays a crucial role by demonstrating that the analytic model accurately reproduces the direct-wave component of NR waveforms once QNMs are filtered out.

Taken together, the qualitative analysis in Fig. 4 and the quantitative matched-filtering results form a complementary pair. Their consistency supports the conclusion that the extracted features indeed correspond to the direct waves.

4. Several statements in the text rely on identifying a certain time before/after merger. But the merger time is not known for observational data. What is used in place of it, and how does its inference/uncertainty impact the conclusions?

Throughout the analysis we do not require the "merger time" in the sense of a precisely defined dynamical moment; instead, we use a *reference time* defined operationally as the time at which the reconstructed strain amplitude reaches its peak, and we align this reference time to t = 0. This definition is now stated consistently throughout the manuscript.

For observational data, the peak time is inferred from parameter estimation and therefore carries uncertainty (dominated by sky localisation and the corresponding inter-detector time delays). We propagate this uncertainty by marginalizing over sky location in the inference, which effectively marginalizes over the reference time t = 0. As shown in Fig. 5, the resulting marginalized contour is consistent with our main results, indicating that our conclusions are not sensitive to the uncertainty in the reference time. We have clarified this point in the text surrounding Fig. 5 and in the caption.

5. A conceptual point remains to be addressed: the values of the remnant mass and spin of the model are fixed to a constant value, hence they correspond to the asymptotic ones. At the early times considered in this study, considerable emission of GWs is still present, which endows mass and spin with time dependence. The authors should quantify this effect and justify why it does not affect their results.

While there have been studies discussing possible effects of time-dependent mass and spin on gravitational waves, these analyses have so far been formulated only in simplified settings. To our knowledge, there is currently no well-established theoretical or phenomenological framework that quantitatively captures how such effects manifest in the fully dynamical binary-merger context. In particular, no clear "smoking-gun" waveform features attributable specifically to time-dependent remnant properties have been identified. We note that the gravitational-wave memory effect is related to changes in mass and momentum, and is already well understood through the supermomentum balance law. Here we are commenting on something beyond that.

Our approach in this work is **result-driven**. By directly comparing the analytic linear direct-wave model with numerical-relativity waveforms in Fig. 3, we find that the model provides a good description of the late-time portion of the direct-wave signal. This empirical agreement indicates that incorporating time-dependent mass and spin is not necessary to model the last few observable cycles of the direct wave at the level of accuracy probed here.

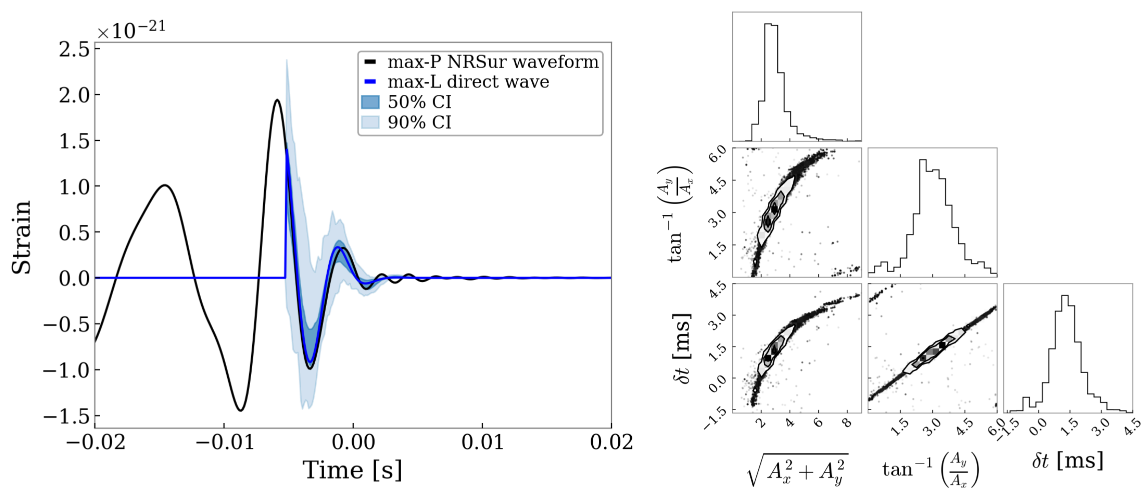
We do observe that the model performs less well at earlier times, suggesting that this regime may be influenced by the dynamical evolution of the remnant's properties (or possible other approximations that enter the modeling).

The consistency is not unexpected. Similar approximations (using fixed mass and spin) appear in other successful modeling approaches: for example, self-force calculations do not consider changes in the central object's mass and spin, yet still show good agreement with numerical-relativity waveforms. Likewise, the close-limit approximation and the Lazarus project assume fixed background parameters while nevertheless capturing key features of the post-merger signal.

One possible interpretation is that, although energy and angular momentum are being absorbed by the black hole during the dynamical regime, the **effective spacetime experienced by outgoing gravitational waves**, and thus by distant observers, can be approximated by one with constant mass and spin over the short time interval relevant for the late-time direct-wave cycles.

In summary, while time-dependent remnant properties may play a role at earlier times, our direct comparison with numerical relativity demonstrates that fixing the remnant mass and spin to their asymptotic values is a valid and sufficient approximation for the late-time direct-wave regime analyzed in this work.

6. Pg.5 "We perform the analysis using the same time-domain Bayesian framework"
I would need to see more details of the inference process to be convinced of the robustness of the analysis. I ask the authors to provide in the response to the referee:
a. The waveform reconstruction (with error bands) for both the agnostic and informed analyses;
b. The posterior distributions on the parameters.



(a) We have attached the waveform reconstruction (left) with associated uncertainty bands for the *informed* analysis. Here, “max-P” refers to the NRSur waveform reconstructed using the maximum-a-posteriori parameters from the IMR analysis. The NRSur waveform is preprocessed in the same manner as the observational data and has the 22{0-2} QNMs removed. “max-L” refers to the direct-wave waveform reconstructed using the maximum-likelihood values of the free-parameters in the direct-wave template, obtained by fitting directly to the observational data; this reconstruction shows good agreement with the NRSur waveform. The shaded regions correspond to the 50% and 90% credible intervals of the reconstructed direct-wave signal, derived from the posterior uncertainties of the direct-wave parameters (right). Note: In the left plot, the small residual wiggles in black (NRSur) around t = 0.005s arise from the subdominant 440 QNM. The cycles covered by the blue band do not correspond to this QNM, since our agnostic fit in Fig. 4 shows that the extracted frequency lies around the horizon mode, which is much lower than the 440 QNM frequency. Since the direct wave decay more rapidly, the 440 mode become visible at late times. As demonstrated in the plot attached in response to Comment 7, this residual 440 mode has negligible SNR and does not affect the fitting to the observational data.

For the *agnostic* analysis shown in Fig. 4, the method directly infers the instantaneous frequency and damping rate at discrete times, rather than a parametric waveform. As a result, it does not yield a unique time-domain waveform reconstruction, and an error-banded waveform representation is therefore not well defined for this case.

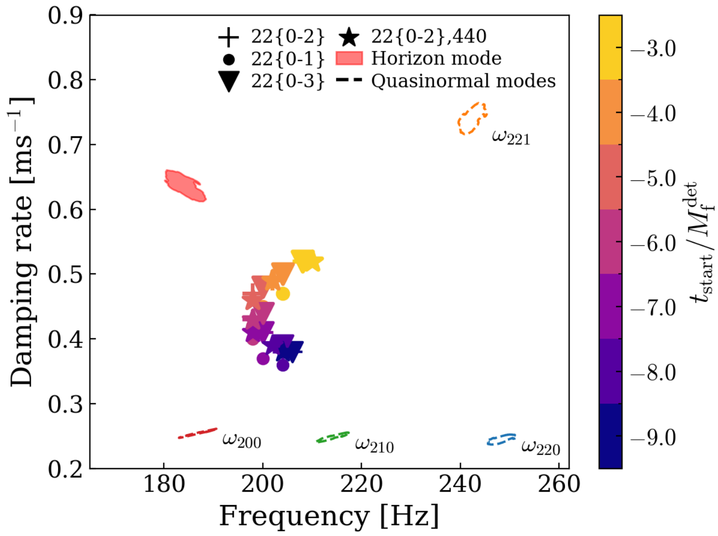
(b) We have added the posterior distributions of the inferred parameters to the Methods section of the revised manuscript (new Fig. 6; also attached above on the right).

7. Page 9: "with marginal support for the (l = m = 2, n = 2) mode."
This sentence appears unjustified, as [41] states: "The presence of a second overtone 222 is not significantly preferred at any time." How can this be interpreted as marginal support?
Since [41] reports no evidence for this mode, this technical choice appears questionable.

We agree that Ref. [41] does not report a statistically significant detection of the 222 mode. However, Ref. [41] also finds that the 220+221+222 model is weakly preferred over the 220+221 model for early start times (approximately earlier than 5M), even though the 222 amplitude is never confidently measured away from zero. Since our analysis of direct waves focuses on this early-time regime, it is therefore reasonable to exclude any potential impact from the 222 mode. We have revised the text to make this clearer when summarizing the results from [41] in Methods C: "..., with weak early-time preference for the presence of the (l=m=2,n=2) mode at $t \lesssim 5\Mfdet$."

To assess the sensitivity of our results to this technical choice, we repeated the analysis using several alternative subtractions. In addition to our default removal of the 220+221+222 modes, we also tested removing (i) 220+221 only, (ii) 220+221+222+223, and additionally (iii) 220+221+222+440 using the NRSur waveform (attached below). In all cases, the inferred frequencies and damping rates remain qualitatively unchanged across the range of start times shown in Fig. 4. This robustness arises because the dominant residual lies at the direct-wave frequency and damping rate, which are well separated from the QNM spectrum.

Since NRSur serves as a controlled, noise-free benchmark, the absence of any meaningful change at the waveform level demonstrates that our analysis is insensitive to the specific choice of subtracted modes. Because the observational data contain additional statistical uncertainties, any corresponding effect would be subdominant and would not alter our conclusions.



Minor comments

1. Abstract: "the quadrupole nature" --> "the dominant quadrupole nature"
Updated.

2. Page 1: "Existing electromagnetic probes"
It is unclear what exactly the authors meant here, and this sentence would likely benefit from slight rephrasing. For example, measuring a light ring is not a direct measurement of an horizon (Ref.[36]).

We have slightly rephrased this sentence to make it clear that the listed electromagnetic observations of black holes (not horizons) are indirect.

3. Page 1: "it consists of a series" --> "it mainly consists of a series"
Updated.

4. Page 2: "the first law of black hole thermodynamics."
Add a reference.
We have added the citation to J. M. Bardeen, B. Carter, and S. W. Hawking, The four laws of black hole mechanics, Communications in Mathematical Physics 31, 161 (1973).

5. Page 2: "and informing waveform modeling [51–55]"
Other notable references that have precisely quantified this agreement are: arXiv:2112.12265, arXiv:2506.02189
Citations added.

6. Page 2: Was Fig.1 AI generated? Does this require a declaration/statement?
I would state explicitly this is a sketch.
It was a manual sketch, not AI generated. We have explicitly stated that in the figure caption.

7. Page 2: "These observable signals are referred to as"
Besides the contribution identified by the authors, there are additional recent references which have identified additional components. It would be fair to acknowledge them:
arXiv:2506.21668, arXiv:2510.17954, arXiv:2510.18956

We have added a footnote to acknowledge them: “See also Refs.~\cite{deamicis2025,kuntz2025,arnaudo2025} for related discussions of other dynamical components.”

8. Page 2: "effectively erasing the memory of the earlier inspiral."
This sentence needs to be justified.

When the remnant spin is high (such as ~0.7 for typical merging binaries), the late-time orbital motion becomes insensitive to the detailed inspiral history, e.g., how the two black holes approached each other prior to merger. Instead, the dynamics in the near-horizon region are dominated by intrinsic properties of the remnant horizon (like frame-dragging). This is the physical reason why horizon signatures can be imprinted on the direct wave and measured in GW250114.

By contrast, for remnants with low spin, frame dragging is weaker and cannot compete with the orbital angular momentum inherited from the inspiral (this can be seen from the geodesics equations listed in Sec. C of the Supplementary Material). In this case, the late-time motion retains a stronger dependence on the early inspiral initial conditions, and the resulting direct-wave signal carries fewer horizon signatures. For example, the oscillation frequency and decay rate of the direct wave are not strongly correlated with the horizon rotation frequency and surface gravity (see Fig. 4 in arXiv: 2509.09165).

To avoid confusion, we have rephrased the sentence as follows: “In particular, strong frame dragging within the ergosphere governs the orbital dynamics, making it insensitive to the history of the earlier inspiral evolution.”

9. Page 2: "until the companion finally crosses the horizon"
So far, the description has been kept at the linear level. At this level, the background is fixed, and the object isn't seen crossing the horizon by asymptotic observers (unlike what stated in this sentence), but only approaching it asymptotically. The authors should either explain what I am missing, or make the sentence more precise.

We thank the referee for pointing out this subtlety. We have revised the sentence to describe the physical picture more precisely: “... The near-horizon redshift effectively suppresses their amplitude, leading to an asymptotic shut-off of outgoing radiation as seen by distant observers.”

10. Page 3: "as the final signature of the infalling particle"
I do not think it is appropriate to describe direct waves as the final signature of the infalling particle. The post-ringdown tails also depend on the particle motion (arXiv:2407.04682, arXiv:2406.17018). The orbital configuration is also known to be imprinted in the QNM amplitudes (e.g. arXiv:2503.09678). Both contributions appear later than direct waves.

We thank the referee for pointing this out. Our intended meaning was that direct waves represent the last **direct emission** sourced by the infalling itself, in contrast to QNMs and late-time tails, which arise from the interaction of the perturbation with the light ring and the long-range potential (in this sense, they can be viewed as **indirect signatures**). We acknowledge that both tails and the amplitudes of QNMs still retain information about the particles's orbital configuration and appear at later times, as noted by the referee.

To avoid ambiguity, we have rephrased the original sentence to be the "final direct emission from the infalling particle".

11. Page 3: "with the merger time aligned to"
What is the definition of "merger time"? Afterwards it is referred to as peak of the strain. The precise definition should be given in this first instance.
In this content, indeed we refer to the time corresponding to the inferred strain peak. We have defined this reference time in the first instance as "...with the reference time defined as the inferred peak of the strain amplitude and set to t=0", and modified other places that may cause confusion.

12. Page 4: "follow common settings in quasinormal mode analyses"
A reference should be given, and why the same choice is sensible here should be motivated. I believe the authors checked that a longer segment length does not impact the result (supp. mat.). Should be stated here briefly.
We have revised the sentence to make it more specific to the rational filter method with a reference: "...follow common settings in quasinormal-mode analyses using rational filters [69]." The following sentence "We have verified that our results are robust against variations in this choice; see Methods for details" has clarified the different choices of segment length do not change the results and pointed the readers to the Methods part.

13. Footnote 1:
a) So far, nonlinear effects have been shown to be subdominant (since the direct wave covers most of the merger signal). Why should they become important far away from the merger?

b) The near horizon approximation assumed to derive Eq.(4) was never mentioned before in the main text. Given that these caveats are important to interpret the measurement, this approximation should be mentioned around Eq.(4).

Note: Now this footnote becomes Footnote 2. And Eq. (4) becomes Eq. (5)

Although the near-horizon approximation is a simplifying assumption entering Eq. (5), Fig. 3 demonstrates that it is sufficient for our current detection purposes. We have now explicitly introduced the "near-horizon approximation" when presenting the model in Eq. (5) and added clarifying remarks afterward: "Relaxing the near-horizon approximation or including nonlinear effects in the analytical model may extend the agreement to earlier times and refine the description of the onset of direct-wave emission, which we defer to future work."

Regarding the role of nonlinear effects at earlier times, we emphasize that identifying the "merger" time via the peak of the strain or gravitational-wave luminosity is largely **phenomenological**. It is therefore not a priori obvious whether extending the direct-wave description to earlier times necessarily corresponds to moving closer to, or farther from, the most dynamical regime of spacetime. In fact, the common (remnant) horizon typically forms a few M ***before*** the strain peak (the time in the bulk is synchronized with that at future null infinity through aligning orbital and GW phases).

Within this context, we note that nonlinear effects may become relevant in modeling the early-time direct-wave regime. Whether such effects are dominant remains an open question, which we defer to future investigation.

We also revised Footnote 2 to "The onset of direct-wave emission is not sharply defined. Relaxing the near-horizon approximation or including nonlinear effects may extend the validity of Eq. (5) to earlier times; see discussion around Eq. (5)."

14. Page 5: "The minor dip and surrounding features near t = 0 are artifacts of the whitening process"
The meaning of this sentence is unclear. What in the whitening process is causing the "artifact"?

Whitening involves frequency-dependent rescaling of the data, which can lead to localized distortions in the time domain visually. We need to use whitening to show the waveform in Fig 2., without which the signal would not be visible.

We have clarified this in the revised sentence: "The minor dip and surrounding features near t = 0 are **visual** artifacts of the whitening procedure **used in the plot, arising from the frequency-dependent rescaling of the data, which exaggerates localized distortions in time domain;**" The following sentence further clarify the only residual component after subtracting the direct wave waveform is the peak earlier than -7M: "the full direct-wave trough has been effectively removed by the template, leaving only a prominent peak at t \lesssim -7 M_f^det (see also Fig. 3)."

15. Page 5: "These results support the interpretation of direct waves as the "final sound""
This sentence is misleading, for the same reasons I highlighted in Comment 10.
We have clarified that this is the "final sound" directly emitted.

16. Eq.(5): this equation should be given above, when this is first mentioned
We now introduce this equation in the second paragraph on page 3 – it becomes Eq. (4).

17. Page 6: "Finally, a systematic search for direct waves in existing gravitational-wave events would be highly valuable"
Why? What features can be observed that cannot already be observed from this loudest event?
While the loudest event provides a clear proof of principle, a systematic search across multiple events would enable population-level tests of the universality, robustness, and parameter dependence of direct waves. This is similar to QNMs: although a QNM was first detected in GW150914, extensive efforts continue to search for QNMs in later events.

We have clarified this motivation by revising the sentence to read: “Finally, a systematic search for direct waves across existing gravitational-wave events would enable population-level tests of their universality and parameter dependence, and allow additional features to be identified through cross-event consistency checks.”

18. Page 9: "Since the filter is applied in the frequency domain,"
Since the authors mentioned that the filter is roughly equivalent to a maximum likelihood subtraction when assuming white noise, it should be mentioned explicitly if (after transforming to the Fourier domain) whitening is applied before acting with the filter.

We thank the referee for this comment. We clarify that no whitening is applied as part of the filtering procedure. The reference to white noise was intended solely to highlight the conceptual connection (under the idealized assumption of white noise) between the filter, which is agnostic to mode amplitudes, and the maximum-likelihood estimator for mode amplitudes used in standard time-domain analyses, as discussed in detail in Refs. [26,27]. It does not describe an operational step in our analysis.

==
========

Referee #3 (Remarks to the Author):

The manuscript investigates the possibility of interpreting a part of the merger signal from a binary black hole merger as a "direct wave", and applies this interpretation to the binary merger event GW250114. The interpretation is interesting and may perhaps even be viewed as a compelling model of the merger. The paper is well written as appropriate for a broad scientific audience and the references are appropriate.

I have some comments below and I would ask the authors to address these before I can definitively recommend the paper.

1. The notions of direct waves are, in my view, the clearest for extreme mass ration systems, i.e. for the motion of a small particle around a massive central black hole. When the applied to systems of comparable mass systems such as GW250114 one needs to rely on the EOB formalism to recover such an interpretation. This is reasonable since this formalism has proved to be accurate and is clearly useful in modeling binary systems. Nevertheless, the results presented in this manuscript are framed as fundamental tests of General Relativity. Moreover they are applied to a very high SNR event where the requirements on waveform accuracy are correspondingly very stringent. It should therefore be asked whether these depend on the assumptions built into the EoB formalism or phrased alternatively, how far can the results presented here be viewed as a test of GR rather than a test of the EOB framework within GR?

We thank the referee for this important comment. The notion and modelling of direct waves is indeed clearest in the extreme mass-ratio limit. For comparable-mass systems such as GW250114, we only rely on EOB formalism to **conceptually** motivate the interpretation of the signal. Neither our waveform template in Eq. (5) nor our measurement strategy (such as the agnostic frequency search in Fig. 4) depends on specific modeling assumptions of any particular EOB implementation.

The only **conceptual** input we adopt from EOB is a viewpoint that the merger phase of a comparable-mass binary can be described as an effective source moving on the spacetime of the remnant black hole. This perspective provides the rationale for employing black hole perturbation theory to model the merger signal.

The validity of this treatment is established in two steps. First, in Fig. 3 (together with the previous theoretical paper arXiv:2509.09165), we demonstrate that the direct-wave template constructed within this framework reproduces the merger portion of **numerical-relativity** waveforms with good accuracy after filtering out quasinormal modes. This comparison is performed against numerical relativity, **not EOB waveforms**. The agreement indicates that the direct-wave template faithfully captures the relevant physics predicted by general relativity.

Second, we apply both an agnostic search (Fig. 4) and a matched-filtering analysis (page 5 and Fig. 2) to GW250114. The results not only qualitatively align with the direct-wave interpretation, but are also quantitatively consistent with the outcome of applying the same analysis pipeline to the maximum-a-posteriori numerical-relativity (surrogate) waveform. For example, the crosses in Fig. 4, obtained by analyzing the numerical-relativity waveform, lie within the measured contours. This provides a three-way cross-check between (1) the theoretical modeling of direct waves, (2) numerical relativity, and (3) GW250114.

We therefore emphasize that no detailed implementation assumption specific to EOB enters the comparison between theory, numerical relativity, and data. In fact, the idea of relating a comparable-mass system to an extreme mass ratio inspiral is not unique to EOB. For instance, results from the self-force community have shown that waveforms constructed in that framework can extend surprisingly well into the comparable-mass regime and resemble numerical-relativity waveforms. Our findings provide one more evidence along this line.

Finally, we note that current EOB waveform models do not include the direct-wave components. Instead, the merger is typically modelled phenomenologically by matching inspiral and ringdown stages. In this sense, our results point to a potential avenue for improving EOB models: incorporating physically motivated direct-wave contributions could lead to a more faithful description of the merger-ringdown signal within the EOB framework.

2. The second set of questions related to the Ringdown analysis. As I understand the method use here, one follows the results of Ref. [41] and removes the 220, 221, and 222 QNMs. The QNMs specifically take the maximum likelihood (ML) values of the mass and spin reported in [41] (and presumably also the start time). How sensitive are the later results on the direct waves on these choices? Ideally, one might prefer a combined model with the direct waves and the QNMs all included. However as a second best approach, rather than taking the ML values, it might be useful to sample from the QNM posteriors found in [41]. In any case, can the authors quantify the effect of these uncertainties in the QNM modeling and in particular if these would affect the extraction of the direct wave in the next step of the analysis? Ideally, I would

want to know the impact of these choices on the statistic used in the direct wave detection. As a related question, how important are the windowing (Gibbs phenomena) effects related to the start of the QNM, and could these affect the direct wave analysis?

We thank the referee for these detailed questions regarding the analysis. We agree that an ideal treatment would involve a joint, self-consistent model including both direct-wave and QNM contributions. Developing such a combined framework requires a flexible, parameterized model of direct-wave emission that can be incorporated alongside QNMs in a unified inference, which we view as an important but substantial theoretical and data analysis task beyond the scope of this first study.

In the present work, to assess the sensitivity of our results to this procedure, we have explicitly tested the impact of uncertainties in the QNM subtraction. We have updated Fig. 5 to include results obtained by sampling the remnant mass and spin from the IMR posterior distribution. The inferred frequency and damping time in the agnostic search remain stable under this variation, demonstrating that uncertainty in the remnant parameters has a negligible impact on the filtered strain.

We have also updated the analysis to use the maximum a posteriori (MAP) values of the remnant mass, spin, and start time, which are more naturally motivated within a Bayesian framework. This update does not qualitatively affect the results.

For the matched-filtering analysis, fully propagating the complete posterior distribution of the remnant mass and spin is computationally demanding. Instead, we provide a comparison plot showing the posterior distribution of the direct-wave parameters obtained using the MAP remnant properties, together with results derived from two representative samples drawn from the tails of the mass-spin posterior. The inferred parameters are consistent across these cases (see plots attached in the response to Comment 1 from Referee #1).

Windowing and Gibbs effects are well-known considerations in ringdown analyses (see arXiv:2107.05609), which shows that performing the inference in the time domain avoids Gibbs artifacts associated with sharp frequency-domain truncation. This approach is now standard in most ringdown pipelines, including the rational filter we use in our study. Although the rational filter applies the QNM subtraction in the frequency domain, it operates on the full IMR time-domain signal using a Fourier transform over the entire waveform (4 seconds) with a standard windowing function. After filtering and transforming back to the time domain, the subsequent ringdown or direct-wave analysis uses the standard time-domain procedure and thus is not impacted by the Gibbs effects. We have also verified that these windowing choices do not bias the extracted direct-wave signal or inferred parameters.

We have clarified these points in the revised manuscript and view the development of a fully joint QNM-direct-wave model as a natural and important direction for future work.

Referee #3 (Remarks on code availability):

I have not reviewed the software used in this analysis, but for the most part the methods employed here are the standard ones used in the literature and I do not expect this to be of much concern.

The QNM rational filter software is publicly available. And we are in the process of releasing all additional analysis scripts and outputs.

==

Referee #4 (Remarks to the Author):

Summary of the key results

This manuscript reports the identification of a “direct wave” component of the gravitational-wave signal produced from the binary black-hole merger GW250114. The direct wave, emitted approximately simultaneously with the ringdown quasinormal modes, carries information about the horizon of the remnant black hole; namely, the horizon rotation frequency (which determines the frequency of the direct wave) and the surface gravity (which determines its decay timescale). It is expected that the direct wave frequency and decay rate will evolve over time due to screening from the surrounding potential barrier, ultimately tending towards the values of the “horizon mode”. This work performs a measurement of the (average) direct wave frequency and decay rate, demonstrating a broad consistency with the predicted horizon mode values.

Below I number comments which the authors may wish to respond to.

Originality and significance: if not novel, please include reference

This is the first time, to my knowledge, such a measurement has been attempted. I believe such a measurement will be of immediate interest to those who work in gravitational-wave astrophysics, with an emphasis on ringdown tests of general relativity and black-hole perturbation theory.

1. However, whilst the identification of the direct wave is certainly interesting in its own right, this work does not demonstrate how this feature can be used to test general relativity (despite the claim that this work “opens exciting avenues to test general relativity”) and I would be interested in the author’s thoughts on this. For example, do the authors envisage a parametrised version of their analytic model used in Fig. 3 which would allow you to measure horizon properties? A more detailed proposal for how a test of general relativity could be performed could improve the manuscript and its impact.

The direct-wave component probes two different aspects of general relativity (GR): as it is sourced in the near-horizon region and subsequently propagating to infinity through the black hole spacetime, it encodes near-horizon orbital dynamics and wave transmissivity from the horizon to infinity (i.e., the gray-body factor). These two ingredients provide natural entry points for parametrized tests of GR.

First, the particle’s motion in the near-horizon region is expected to closely follow Kerr geodesics during the final plunge, where self-force effects are negligible. Deviations from GR could modify this motion in several ways: (i) changes in the underlying metric would lead to altered geodesics, (ii) environmental effects such as a dark-matter halo could introduce additional forces, and (iii) quantum or stochastic effects could induce Brownian-like perturbations. Such effects would imprint themselves on the time evolution of the instantaneous frequency $\omega_G(t)$ appearing in Eq. (5). These deviations can be parametrized in a systematic manner, for example following the framework analogous to parameterized post-Newtonian.

Second, the propagation of the direct wave from the near-horizon region to infinity is governed by the black hole's potential, which determines the transmissivity (gray-body factor). This aspect can also be parametrized by introducing controlled deformations of the potential. In particular, the screening of horizon modes in GR may not hold in certain beyond-GR scenarios, leading to measurable modifications.

These two ingredients can be combined into a parametrized framework that directly confronts the observed direct-wave signal with GR predictions.

Beyond direct-wave-only tests, consistency checks across **direct waves, QNMs, and inspiral signals** provide an additional and complementary avenue, analogous in spirit to IMR consistency tests. With improved waveform modeling, one can infer remnant and binary parameters using the direct wave alone and compare them against those obtained from QNMs and inspiral. Alternatively, a joint post-inspiral model incorporating both direct-wave and QNM contributions can be used to perform self-consistent tests of GR.

Finally, we emphasize that the results presented in this manuscript are not themselves intended as a test of GR. Rather, our goal is to establish the **observational relevance** of direct waves as a distinct and measurable component of the post-inspiral gravitational-wave signal, and to demonstrate their consistency across analytic modeling, numerical relativity, and observational data. This constitutes a necessary first step toward developing a new framework for testing GR using direct waves in future high-signal-to-noise detections.

We have added "Such tests could be enabled by joint merger-ringdown analyses in which parametrized direct-wave and quasinormal-mode models are fitted simultaneously, allowing consistency tests of remnant and near-horizon properties once theoretical and observational systematics are carefully accounted for." to the Conclusion section.

My concern is that the direct wave may not open up tests of general relativity as the authors claim. Figure 4 in the manuscript helps to explain my reasoning: there is a straightforward prediction of the horizon mode frequency and damping rate, but this is not what is observed due to a screening from the potential barrier. So, any discrepancy between what is measured and what is predicted can be attributed to this screening. Unfortunately, it seems this screening (at least for binary mergers with comparable masses) is not sufficiently well understood for any predictions to be tested with data. This can be compared, for example, with tests of general relativity involving quasinormal modes: in the same figure the quasinormal mode frequencies and damping rates are also indicated. In contrast to the direct wave and horizon mode, the quasinormal mode frequencies/damping rates are exactly what is measured by an observer, thus allowing quantitative statements about the consistency of observations with theory. The author's thoughts on what quantitative statements the direct wave and horizon mode could offer would help to see its potential impact.

We thank the referee for this important concern. We first wish to clarify that the interpretation of Fig. 4 in the referee's comment is not correct. The offset between the inferred frequency/damping rate and the horizon mode shown in Fig. 4 is not primarily a consequence of late-time screening by the potential barrier. Rather, it arises because Fig. 4 probes an *early* portion of the direct-wave regime (approximately from -7M to -3M), where the instantaneous frequency and damping rate are not expected to coincide with the horizon mode. This is demonstrated explicitly in Supplementary Fig. 9, which shows that the instantaneous direct-wave frequency and damping rate approach the horizon mode only at much later times (t ~ 20M). We have better clarified this in the third last paragraph before conclusion, when discussing the results in Fig. 4.

We agree with the referee that Fig. 4 by itself does not provide a sharp quantitative test of general relativity (GR). This is a consequence of the **simplified analysis strategy** adopted there: within a given time window, we approximate the direct wave using a QNM-like ansatz with constant frequency and damping rate. **The purpose of this approach is intentionally qualitative and model-agnostic**, in order to demonstrate that the extracted averaged frequencies and decay rates cluster around, though do not coincide with, the horizon mode. By sliding the fitting window to later times, the inferred decay rate increases, consistent with continued radial acceleration of the source (see Supplementary Fig. 9). Meanwhile, the same qualitative behavior is observed when applying the identical analysis to numerical-relativity waveforms (shown as plus signs in Fig. 4).

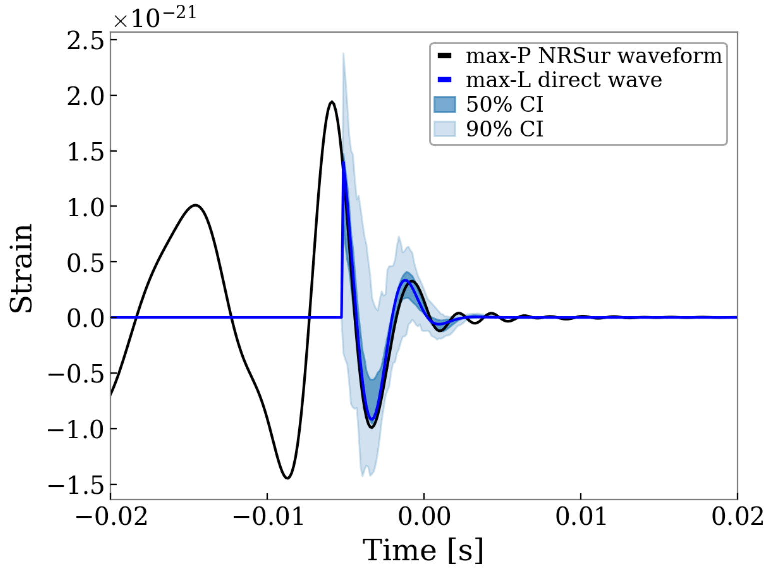
The referee is correct that a more quantitative framework is required to enable tests of GR. This motivates our second, complementary analysis based on **matched filtering** (page 5 and Fig. 2). In this approach, we employ the physically motivated analytic model of Eq. (5), which explicitly captures the time-dependent dynamics of the direct wave through ω_G(t). We can then visually examine how the best-fit waveforms compare to the NRSur waveform, and consider the posterior distributions of the inferred parameters (added in the Supplementary Material as Fig. 6).

We then examine the residual after subtracting the best-fit waveform (as shown by the reconstructed "max-L" direct-wave waveform in blue below). This strategy is better suited for systematic extensions, including parametrized deviations from GR, as discussed earlier in our response.

We emphasize that the two approaches serve distinct but complementary purposes. The QNM-like analysis in Fig. 4 is deliberately model-agnostic and data-driven, allowing qualitative features, such as the time evolution of the decay rate, to be inferred without committing to a specific dynamical model. In contrast, the matched-filtering approach provides a natural pathway toward a quantitative testing framework through controlled parametrization of near-horizon dynamics and wave propagation.

Finally, we reiterate that the primary goal of this manuscript is **not** to perform a test of GR, but rather to establish the **observational relevance** of direct waves in GW250114. By addressing the referee's concern here and above, we believe we have outlined a concrete and doable avenue to test GR through direct waves.

As an additional remark, although the horizon mode itself is screened, its associated **nonlinear effects**, such as couplings with quasinormal modes, may not be screened and could manifest in higher harmonics (e.g., the l=m=4 mode). This is another promising direction for future investigations.



Data & methodology: validity of approach, quality of data, quality of presentation

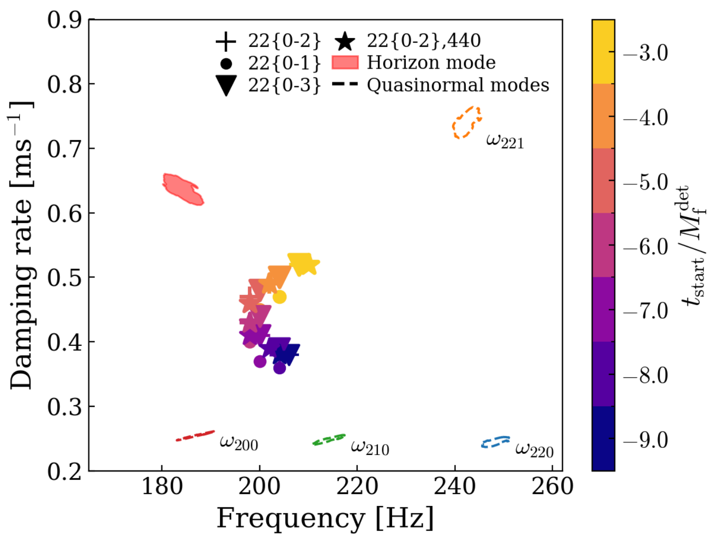
I found the manuscript well written and well presented. I believe the quality of the gravitational-wave data is of no concern; it corresponds to a high-SNR event which has been well studied (see Refs. [40,41] in the manuscript). The core of the method relies on an analysis method known as a quasinormal mode rational filter. This approach is well tested in terms of performing ringdown analysis and used in multiple works (including works by the LIGO Scientific, Virgo, and KAGRA Collaborations). I do, however, have some questions/concerns regarding the use of this method for the direct wave analysis:

2. Firstly, I have a concern that not all of the quasinormal modes in the data are being removed. Ref. [82] in the manuscript demonstrates that as many as nine overtones are excited early in the ringdown, but throughout this work only the first two overtones have been removed. It is true that the SNR of GW250114 only allows the first overtone amplitude to be measured confidently (with some tentative evidence for the second overtone), but I would argue that the higher overtones are still present in the data (even if not individually resolveable). Would it not make sense to remove contributions from all overtones before searching for the direct wave?

According to Eqs. (11) and (12) in arXiv: 2301.06639, applying filters to a signal inevitably suppresses the amplitudes of other QNMs, since QNMs are not orthogonal to each other. In this case, removing the first two overtones simultaneously reduces the amplitude of the third overtone by ~98%. As a result, even though numerical-relativity studies demonstrate that many overtones can be excited at very early times, filtering only the dominant ones is sufficient to suppress the collective QNM contribution to a level where the direct wave can be identified. This is the strategy we adopt throughout, including in Fig. 3, where only two overtones are filtered when analyzing numerical-relativity waveforms.

For observational data such as GW250114, only a limited subset of these modes can meaningfully contribute given the finite signal-to-noise ratio. Higher overtones are not individually resolvable. To assess the impact of this modelling choice, we explicitly tested the sensitivity of our results to the treatment of additional modes by repeating the analysis with several alternative subtractions. In addition to our default removal of the 220+221+222 modes, we also considered removing (i) 220+221 only, (ii) 220+221+222+223, and (iii) 220+221+222+440, using the NRSur waveform (see plot attached below). In all cases, the inferred frequencies and damping rates remain qualitatively unchanged across the range of start times shown in Fig. 4.

This robustness arises because the direct-wave is less impacted (suppressed) by the filters.

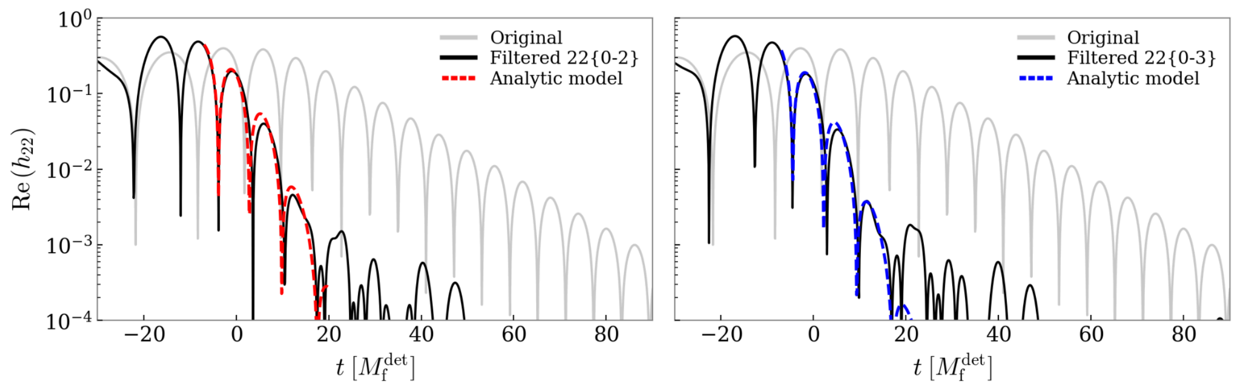


This is also true for NRSur7dq4 - in Fig. 3 only the first two overtones were removed, but presumably SNR is not an issue here and so higher overtones can certainly also be removed? In Fig. 3 we are presumably seeing some contribution from the higher overtones in the filtered data, so I am concerned the direct wave model is instead fitting to them. The author's comments on this issue would be appreciated.

As discussed earlier, removing the first two overtones already reduces the amplitude of the third overtone by ~98%, making it negligible even when no explicit filter for that mode is applied.

We attached a pair of plots below with and without subtracting the 223 mode and find that the QNM-subtracted waveforms are fitted equally well by the analytic direct-wave model in both cases. The small differences do not affect any of the results or conclusions, as shown in the fitting comparison plot above. For GW250114, there is no evidence for the presence of the 223 mode in the data, so the choice to subtract only the 220+221+222 modes is well motivated and consistent with the analytic model used.

Finally, the model-agnostic frequency and decay-rate inference shown in Fig. 4 provides an independent check. The extracted parameters lie close to the theoretical expectation for the direct wave and are well separated from the expected locations of higher overtones, which are far outside the inferred contours.



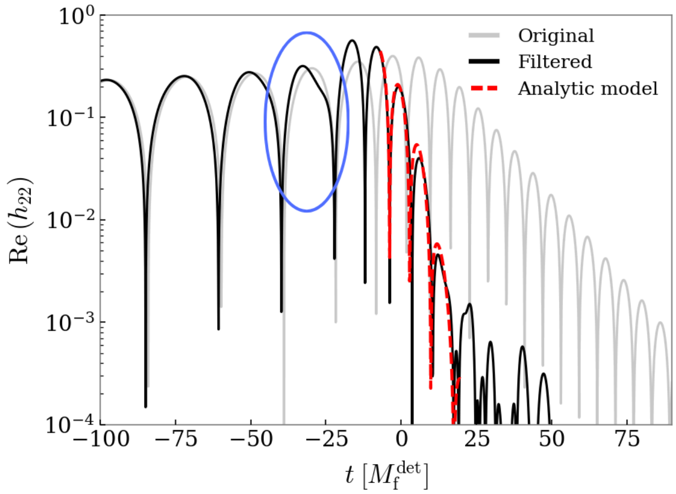
3. Secondly, I was curious about the impact of the filter on the pre-ringdown data. As explained in Ref. [60] in the manuscript, applying the filter introduces a "flipped ringdown" to the data immediately preceding the ringdown. I was surprised to see this was not mentioned in the manuscript, especially since the work involves fitting data slightly at or slightly before the time of peak strain. Was it checked if an artificial flipped ringdown was being introduced to the data, and if this interfered with the fitting of the direct wave?

Indeed, filters introduce a time-reversed ("flipped") QNM component prior to the nominal start time of the ringdown. We emphasize, however, that **there is no first-principles requirement that the QNM start time be tied to the peak of the strain**. To date, determinations of QNM start times have been entirely phenomenological, relying on the success or failure of QNM-only fits.

From the perspective of the present study, which explicitly incorporates direct waves, this provides a useful retrospective view on earlier QNM analyses. Those analyses did not account for direct-wave contributions, and it is therefore unsurprising that QNM-only models tend to fail near the strain peak. Such a failure does not imply that QNMs have not yet been excited; an equally plausible interpretation is that the signal near the peak consists of a superposition of QNMs and direct waves.

In Fig. 3 (also attached below), several cycles of the direct wave become visible after filtering. We argue that these cycles are not contaminated by the "flipped ringdown" (or equivalently a "ringup") artifact for the following reasons:
(a) the direct-wave amplitude decreases with time, whereas a ringup contribution would increase;
(b) the agnostic fits to the decay rate shown in Fig. 4 are positive, while a ringup component would correspond to a negative decay rate (reflecting growth); and
(c) the extracted direct-wave cycles agree well with our theoretical model (red dashed curve), which is built without the ringup artifact.

From a result-driven perspective, the fact that the "flipped ringdown" components do not contaminate the identified direct waves suggests that QNMs are already excited at relatively early times. For example, we do observe that one late-inspiral cycle (highlighted by the circled region below) is distorted by the filter. This distortion can be attributed to the ringup component, indicating the existence of a narrow time window around the strain peak in which QNMs coexist with other dynamical contributions, including direct waves, dynamical excitation of QNMs, and potentially nonlinear effects.



Regarding reproducibility:

4. I don't believe the priors are provided on the frequency and damping rate used in the main direct-wave analysis (configuration files for the specific analyses carried out would help with reproducibility).
We have added the prior ranges in Methods Sec D: "..., assuming uniform priors on the frequency (160–240 Hz) and damping rate (0.05–0.9 ms^-1)."

5. The authors should provide more details on the procedure for obtaining the detection threshold. For example, what priors are used when applying the analysis to noise? How many trials are performed?

We have added the following text to clarify how the detection statistic is computed in Methods D:

"We determine the detection threshold by performing injection studies in which a QNM-only signal (l=m=2,n=0,1,2) is added to the detector noise surrounding the event. These QNMs are injected with amplitudes drawn uniformly from 4×10^{-22} to 2×10^{-21} and phases from 0 to 2\pi at the reference time, while their complex frequencies are fixed to the maximum-a-posteriori remnant parameters inferred from the inspiral-merger-ringdown analysis. For each simulated realization, we compute the detection statistic comparing a model that includes an additional damped sinusoid (representing a direct-wave component) against a baseline model in which only the three injected QNMs are removed, with their parameters fixed to the injected values. The prior ranges and parameter resolution for the additional damped-sinusoid search are identical to those used in the direct-wave search of the event data. This setup therefore mirrors our search for an additional damped sinusoid on top of the three QNMs with parameters determined from the inspiral-merger-ringdown analysis. Repeating this procedure for 300 independent noise realizations (drawn at 64-s intervals within the 3 hours before and after the event, excluding the 10 s surrounding the event itself) yields the distribution of the detection statistic under QNM-only null hypothesis. We define the threshold D_1% as the value corresponding to a 1% false-alarm probability. An observed D - D_1%>0 therefore indicates that the preference for a direct-wave signal is unlikely to arise from noise fluctuations alone."

6. It is also not clear how the analytic model for the direct wave is fitted to the NRSur7dq4 model (I believe this is referred to as the "benchmarked waveform" in the Methods). Is there an amplitude and phase being fitted for? The same applies to the waveform without screening shown in Fig. 9.

Indeed, after constructing the direct-wave analytic models (both with and without screening), we align them to the NRSur7dq4 waveform by applying a time shift and fitting a single complex-valued amplitude, which accounts for both overall amplitude and phase differences. The fit is performed using a least-square method (np.linalg.lstsq in Python). We have clarified this fitting procedure in the Methods section E of the revised manuscript.

7. It is not explained how the sky-location marginalization is performed for the result in Fig. 5.
We have clarified how the sky-location marginalization is performed in the caption of Fig. 5, along with the newly added results from marginalizing over remnant black hole parameters. The sentence in the caption reads: "The red (purple) contour marginalizes over sky location (remnant black hole parameters). The sky-location marginalization effectively propagates uncertainty in the reference time at which the strain amplitude peaks, while the remnant-parameter marginalization accounts for possible variations in the QNM (l = m = 2, n = 0,1,2) frequencies."

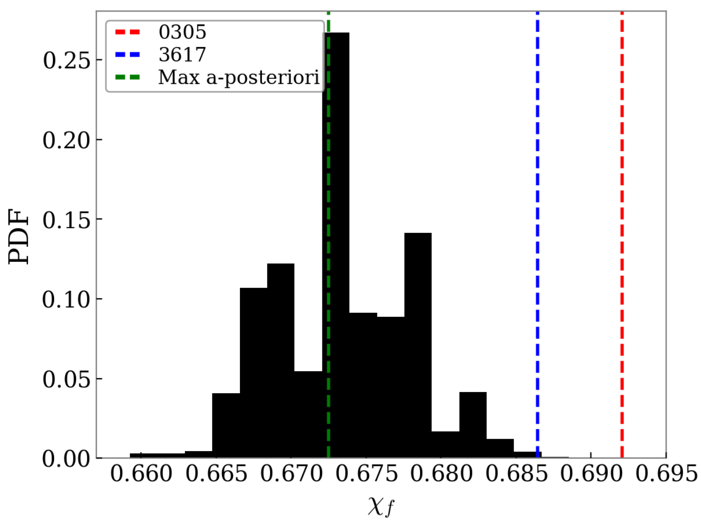
Finally:

8. The manuscript also makes use of the waveform model NRSur7dq4 to illustrate the presence of the direct wave. However, using a numerical relativity simulation would be more accurate, and since just one set of parameters are used for NRSur7dq4 I see no reason why a numerical relativity simulation couldn't be used for the results in Figs. 3 and 9 (it may not make a significant difference, but since NRSur7dq4 is known to suffer from some loss of accuracy in the ringdown I believe it would at least be worth checking). The GW250114 companion paper, Ref. [41] in the manuscript, used the simulation SXS:BBH:3617.

We thank the referee for this comment. While a full NR waveform would in principle provide the most accurate description of the signal, there is currently no NR simulation that closely matches the inferred remnant parameters of GW250114.

In particular, although Ref. [41] uses SXS:BBH:3617 (with intrinsic parameters lying within the 90% credible region of GW250114), its remnant properties, especially the final spin which the direct wave signal is sensitive to, lie toward the tail of the posterior distribution for GW250114 (see attached plot below), and we find that it not accurate enough to be used as a waveform for this event.

We therefore adopt the maximum-a-posteriori NRSur7dq4 waveform. Although NRSur7dq4 is known to be less accurate deep in the ringdown, the direct-wave signal analysed here is concentrated at earlier times where the signal-to-noise ratio is highest. The subdominant systematics in a surrogate waveform are not expected to impact our results, as reflected by the clear residual features shown in Fig. 3.



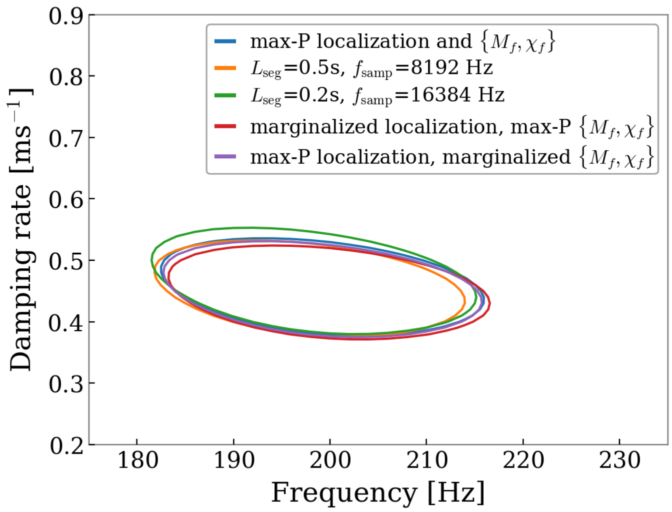
Appropriate use of statistics and treatment of uncertainties

9. This analysis relies upon removing the quasinormal-mode content from the data before searching for the direct wave. This removal of quasinormal modes has some inherent uncertainty associated with it, which is not propagated throughout this analysis. Specifically, the filters used to remove the quasinormal modes are constructed using the remnant black hole mass and spin, which has some uncertainty. One could perhaps consider jointly searching for the direct wave and filtering out the quasinormal modes (drawing values for the remnant mass and spin from the posterior) - did the authors consider this? It may also be true that the uncertainty on the remnant mass and spin is sufficiently small to not impact the direct wave measurement shown in Fig. 4, but this is not stated in the manuscript.

This above also applies to the matched-filtering analysis used to obtain the direct wave SNRs.

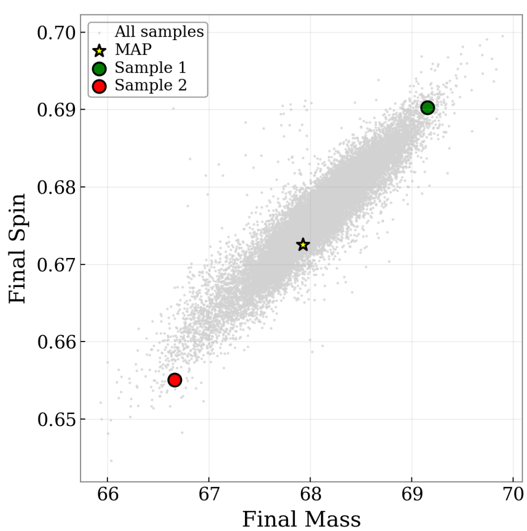
We agree that the subtracted QNMs are associated with uncertainties. To assess the impact of this uncertainty, we repeated the agnostic analysis at -7M shown in Fig. 5 (also attached below) to include results obtained by sampling the remnant mass and spin from the IMR posterior distribution.

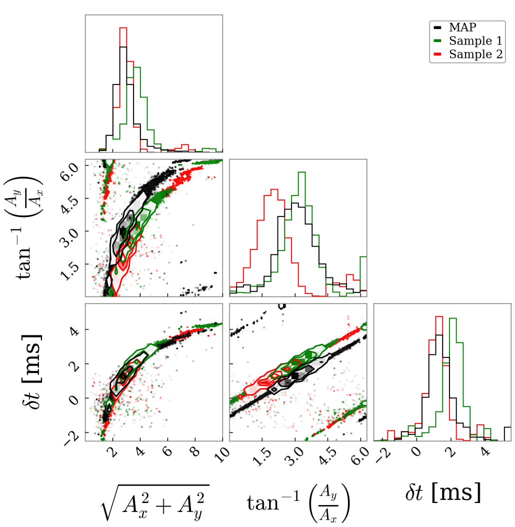
As shown in the plot, the inferred frequency and damping time from the agnostic search after marginalizing over the {M, chi} from different posterior samples (shown in purple) are consistent with the analysis using only the maximum-a-posteriori choice of {M, chi} and so the uncertainty of the remnant parameters does not affect the direct wave measurement.



For the matched-filtering analysis, fully propagating the complete posterior distribution of the remnant mass and spin is computationally demanding. Instead, we provide a comparison plot showing the posterior distribution of the direct-wave parameters obtained using the maximum-a-posteriori (MAP) remnant properties, together with results derived from two representative samples drawn from the tails of the mass-spin posterior. The inferred parameters are consistent across these cases (see plots attached below).

Jointly searching for direct waves and QNMs using the remnant posterior distribution would require generating analytic direct-wave templates across parameter space, which in turn requires a flexible, parameterized direct-wave waveform model. Developing such a model is beyond the scope of this first study and is considered for future work. We have expanded on this future direction in the conclusion of the revised manuscript.





10. I am also curious about what uncertainties are present in the construction of the analytic model used in Fig. 3. Certain parameters have to be chosen (such as the energy and angular momentum of the orbit), but it is not clear to me how much flexibility there is with picking them. The agreement between the analytic

model and the filtered waveform is one of the key pieces of evidence for the direct wave, so it is important to know how robust that agreement is to choices of parameters.

As explained in Sec. C of Supplementary Material, the spin of the remnant black hole is the only parameter that enters the analytic model. The energy and angular momentum are uniquely given by the values associated with the innermost stable circular orbit (ISCO), which read

$$E = \frac{r_I^{3/2} - 2r_I^{1/2} + \chi}{r_I^{3/4}\sqrt{r_I^{3/2} - 3r_I^{1/2} + 2\chi}},$$

$$L_z = \frac{r_I^2 - 2\chi r_I^{1/2} + \chi^2}{r_I^{3/4}\sqrt{r_I^{3/2} - 3r_I^{1/2} + 2\chi}}.$$

where r_I is the prograde ISCO radius. In practice, we start the particle at r_I - 10^(-5) to allow it to plunge into the black hole. We have checked that the small offset does not impact the result since the direct wave regime corresponds to relatively late times and is insensitive to the transition from quasi-circular motion to plunge.

We have added the expressions of E and L_z to Sec. C of Supplementary Material.

I believe the detection statistic is a suitable statistical test (provided some more details on the exact procedure are provided). A couple of other comments:

11. It would be more accurate to label the red line in Fig. 2 “Maximum-likelihood waveform reconstruction”, since the reconstruction would have some uncertainty associated with it.

We have updated the label in Fig. 2, additionally, we have also updated the analysis to use the maximum a posteriori values of the remnant mass, spin, and start time, which are more naturally motivated within a Bayesian framework. This update does not qualitatively affect the results.

12. It is not stated what level the contours correspond to in Fig. 5.

We have added “90% credible region” to the caption of Fig. 5

Conclusions: robustness, validity, reliability

Overall, provided the concerns above are addressed, I believe the results in Fig. 3 and Fig. 4 provide compelling evidence for the detection of a direct wave component (and I consider this to be the main argument for publication in Nature). The authors have provided some checks of the robustness of the result, but as explained above I believe additional checks should be performed. Finally, as explained in point 1, the impact of this result on tests of general relativity is not clear to me (especially when compared to the usual quasinormal-mode analysis) and could use more explanation in the manuscript.

We have individually discussed the additional checks in response to earlier comments. We have also clarified the impact of these results on the test of general relativity in conclusion (as summarized above).

Other comments

13. There are a couple of places in the manuscript where the time of peak strain is referred to as the merger time, but I would avoid saying this since the merger time is not very well defined — simply saying t=0 is set to the time of peak strain is sufficient.

We have now fixed this throughout the manuscript.

14. I think Fig. 2 could be improved by providing the time axis in units of the (detector frame) maximum-likelihood remnant black hole mass (perhaps with a second axis at the top of the figure). This would aid with comparing Fig. 2 to Fig. 3.

We have added a second axis at the top of Fig. 2.

15. It is stated that the inferred direct wave frequency does not fully asymptote to omega_H in part due to "the limited signal-to-noise ratio at late times". However, we also see that the fits to NRSur7dq4 (plus symbols in Fig. 4) also don't asymptote to omega_H, despite the SNR presumably not being an issue here. Clarification on this would be appreciated.

The offset between the NR results and the horizon mode in Fig. 4 is as expected. The instantaneous frequency and damping rate for $t < 0$ do not coincide with the horizon mode, but instead evolve toward it asymptotically at later times. This behaviour is illustrated explicitly in Fig. 9, where the instantaneous frequency and damping rate approach the horizon mode only at much later times ($t \sim 20M$). However, in this late-time regime, the direct-wave amplitude is strongly suppressed by near-horizon redshift effects and additional screening effects, making the signal effectively unobservable.

The behavior of instantaneous frequency and damping rate from -7M to -3M shown in Fig. 4 closely matches the evolution shown in Fig. 9 for the same regime. In particular, the decay rate is smaller than kappa but increases monotonically with time, while the real frequency is roughly 10% above Omega_H. We have added clarifications in the third last paragraph before conclusion to better clarify the results when discussing Fig. 4.

We thank the referees for their careful review of our revised manuscript and for their helpful additional suggestions. We have addressed these points in the manuscript and summarized them below. We also note that Fig. 4 has been slightly updated relative to the previously submitted version, as the earlier submission inadvertently used a slightly outdated version of the figure. This update is minor and does not affect any of the results, interpretations, or conclusions of the manuscript.

Referee #1 (Remarks to the Author):

Regarding the sky position marginalisation: the authors' reply is correct only under the assumption that the uncertainty in the absolute time of arrival (e.g. at the geocenter) is negligible, as its effect is distinct wrt time delays, now marginalised over via the sky position. This statement needs to be supported by a quantitative argument, or this uncertainty included as well.

We thank the referee for this additional comment. In our analysis, we marginalize directly over the reference time at which the strain amplitude peaks in each detector (Hanford and Livingston), rather than assuming a fixed geocentric arrival time. This captures the relevant timing degrees of freedom, which can equivalently be expressed in terms of geocentric peak time and sky location, and therefore includes the associated uncertainties. We agree that referring to this as sky-position marginalization was imprecise, and we have revised the legend, caption, and relevant text to clarify this point. We attach the updated figure and caption below for reference.

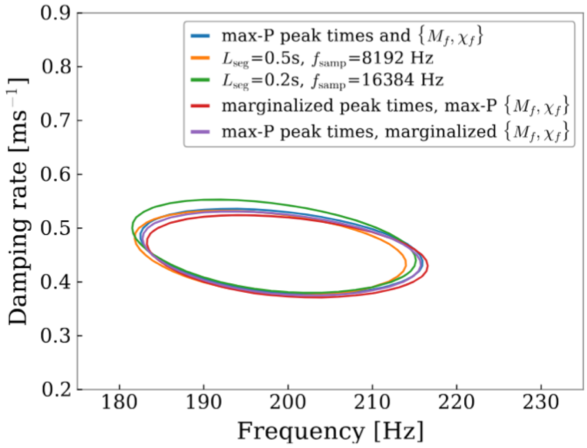


Figure 5: Inferred frequency and damping rate (90% credible region) of the direct-wave signal at $t_{start} = -7M_f^{det}$ under different analysis settings for L_{seg}, f_{samp}, strain peak times, and remnant black hole properties $\{M_f, \chi_f\}$ (with "max-P" denoting maximum-a-posteriori values). The red (purple) contour marginalizes over the strain peak time in each detector (remnant black hole parameters). The timing uncertainties reflect variations in geocentric peak time and sky location, while the remnant-parameter marginalization accounts for possible variations in the QNM ($\ell = m = 2, n = 0, 1, 2$) frequencies. The main results presented in the paper correspond to the settings in blue. The analysis is robust against variations in these settings.

Regarding 5.: This is an acceptable statement, but the fact that the correct inclusion of this effect is unknown, and the consequent assumption of ignoring it, should be explicitly made in the manuscript. The reader should be cautioned more explicitly about this ignorance, and that these statements are based on empirical numerical observations rather than a well-understood model.

We thank the referee for the suggestion. We have added a statement in the discussion after Eq. (5) to caution the reader, which reads "In this analysis, we fix the remnant mass and spin in the analytic model to their asymptotic values. The impact of their time dependence due to merger nonlinearities is not yet quantitatively understood; however, the empirical agreement with numerical-relativity waveforms in Fig. 3 indicates that this approximation is sufficient for describing the late-time direct-wave signal considered here."

6. I also request to add the waveform reconstruction to the manuscript, as it is very informative.

We thank the referee for this suggestion. We have added the waveform reconstruction to the manuscript in the "Template-based search for direct waves" section (Methods). In the process, we identified and corrected a bug in the code used to generate the plot provided in our previous response, where a subset of posterior samples was inadvertently drawn from the tail of the distribution. We now draw 10,000 samples uniformly from the full posterior and reconstruct the waveforms. With this correction, the inferred uncertainty is significantly reduced, as reflected in the updated figure in the manuscript (also attached below).

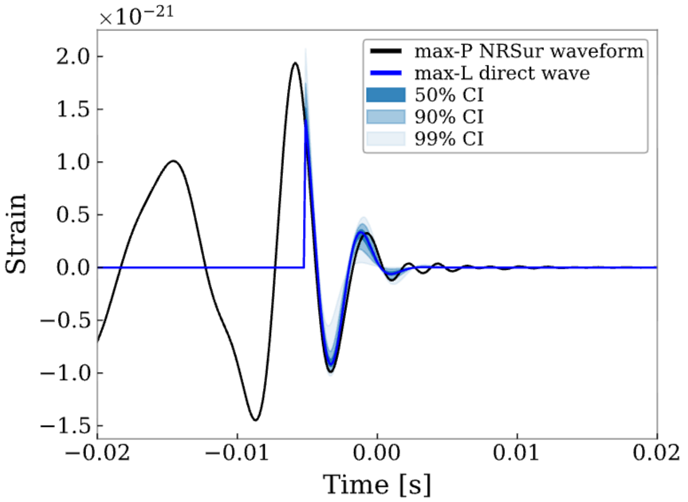


Figure 7: Waveform reconstruction. Black: maximum-a-posteriori ("max-P") NRSur7dq4 waveform from the inspiral-merger-ringdown analysis, with the ($\ell = m = 2, n = 0, 1, 2$) QNMs removed. Blue: maximum-likelihood ("max-L") direct-wave reconstruction. Shaded bands: 50%, 90%, and 99% credible intervals.

Once these points are incorporated, in my opinion the manuscript can proceed to publication.

Referee #1 (Remarks on code availability):

If the authors will release the scripts and data as promised in the report, I believe that this would allow for full reproducibility.

We have prepared the data and code for publication on Zenodo and GitHub (information included in the manuscript).

Referee #3 (Remarks to the Author):

I thank the authors for the detailed responses to my comments and the changes made to the paper, which address my main concerns. I am happy to recommend the paper for publication in Nature. Below I give additional comments on the paper.

The paper provides compelling evidence for the detection of a direct-wave in the GW250114 signal, and establishes its observational relevance. I believe the identification of the direct wave will have implications for waveform modeling, fitting ringdown quasinormal modes, and will open new avenues for testing general relativity. However, the paper does not perform any such tests. It is my understanding that to perform robust tests of general relativity more sophisticated calculations for the evolving direct-wave frequency omega_G are needed (I believe this is at least partly what the sentence "Such tests could be enabled … once theoretical and observational systematics are carefully accounted for" on page 6 is referring to). To reiterate, my concern is that the omega_G currently employed in the paper is calculated in the test-mass limit, and any discrepancy with general relativity could be attributed to applying a test-mass limit calculation to an equal-mass system (I apologise for incorrectly claiming the offset between the final horizon frequency and the measured frequency was due to the screening - my main point was that a robust prediction for the evolving frequency is required to test general relativity).

I believe the analysis method is robust and the conclusions are sound. I would just note (as also mentioned in the paper) that ideally one would fit for the ringdown quasinormal modes and direct wave simultaneously in a self-consistent inference. The current analysis, by effectively pre-processing the data to remove quasinormal modes, misses any correlations that exist between the quasinormal modes and direct wave. Related to my previous point, any high-precision analysis and test of general relativity should perform the fitting simultaneously. But, for the sake of establishing observational relevance, the current analysis is sufficient.

We have revised the conclusion to explicitly state this:

"It marks a first step towards future tests of general relativity with direct waves. Such a framework could include joint merger-ringdown analyses, in which parametrized direct-wave and quasinormal-mode models are fitted simultaneously to assess the consistency of remnant and near-horizon properties, together with comparisons of extracted direct-wave features against general-relativity predictions."

This paragraph is followed by descriptions of required theoretical development: "Realizing this framework will require more accurate waveform templates for direct waves, including extensions to precessing binaries and the incorporation of nonlinear effects. Our current analysis relies on an analytic description derived in the test-mass limit, with the remnant mass and spin fixed to their asymptotic values. To date, studies of nonlinear ringdown have largely focused on quadratic quasinormal modes. Extending these efforts to capture nonlinear contributions to direct waves, including time dependence of the remnant mass and spin, will be an important direction for future work."

Finally, I thank the authors for the additions made to Fig. 5, which helps to make the conclusions robust. I just request clarification on one point: when the analysis is performed with different sky locations, is the geocenter peak time being fixed? In the main text below Fig. 5 it is stated the geocentric signal start time is varied, but this isn't stated in the Fig. 5 caption. Instead in the caption it is stated that varying the sky location effectively propagates uncertainty in the peak time, but since the geocenter peak time will have its own uncertainty associated with it I am not sure if this is true.

We thank the referee for this comment. In the analysis, we marginalize directly over the reference time at which the strain amplitude peaks in each detector (Hanford and Livingston). This captures the relevant timing degrees of freedom, which can equivalently be expressed in terms of geocentric peak time and sky location, without fixing either quantity. The contours shown in Fig. 5 therefore incorporate the combined timing uncertainty across both detectors. We have revised the legend, caption, and relevant text to make this point explicit. We attach the updated figure and caption below for reference.

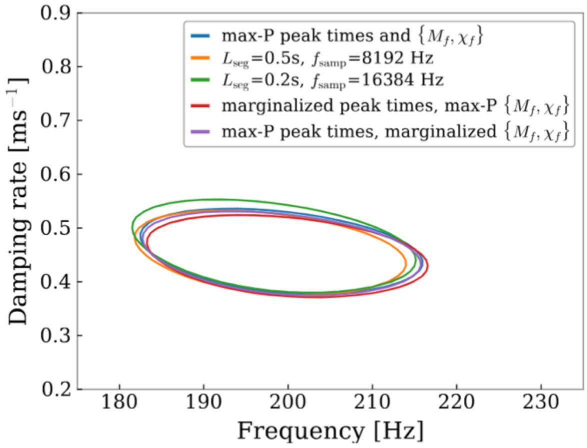


Figure 5: Inferred frequency and damping rate (90% credible region) of the direct-wave signal at $t_{\rm start} = -7M_{\rm f}^{\rm det}$ under different analysis settings for $L_{\rm seg}$, $f_{\rm samp}$, strain peak times, and remnant black hole properties $\{M_f, \chi_f\}$ (with "max-P" denoting maximum-a-posteriori values). The red (purple) contour marginalizes over the strain peak time in each detector (remnant black hole parameters). The timing uncertainties reflect variations in geocentric peak time and sky location, while the remnant-parameter marginalization accounts for possible variations in the QNM ($\ell = m = 2, n = 0, 1, 2$) frequencies. The main results presented in the paper correspond to the settings in blue. The analysis is robust against variations in these settings.