Peer Review File

Manuscript Title: Optical measurement of superluminal motion in the neutron-star merger GW170817

Reviewer Comments & Author Rebuttals

Reviewer Reports on the Initial Version:

Referee expertise:

Referee #1: observations of transients

Referee #2: observations and theory of transients

Referee #3: theory of transients

Referees' comments:

Referee #1 (Remarks to the Author):

This is an interesting paper - the author's use a careful analysis of HST imaging of the kilonova counterpart of GW170817, to establish the exact location of the merger, and hence provide better constraints on the motion of the radio afterglow (driven by a relativistic jet) seen in VLBI observations. A crucial development is the availability of GAIA DR2 and EDR3 releases, which ensures an excellent tie between the radio and optical frames. The authors use this measurement to draw improved conclusions about the physical properties of the jet, including initial Lorentz factor, angle to the line of sight etc.

The paper is mostly well written and well referenced, and the discussion on interpretation convincing. The authors deserve particular credit for the very careful astrometric analysis. It is certainly impressive to get substantially better than 1mas positional uncertainties with standard HST imaging. Given the considerable interest in the jet produced by the GW170817 merger (thanks to the uniquely powerful observational data-sets), I would recommend the paper for publication in Nature, even though I don't really buy the narrative that it is a breakthrough that will necessarily open up future "new avenues" - in part because observing jet motion is itself not new, and in part because such opportunities are rare and the available options may differ from one case to another.

Some other comments:

1) opening sentence of para 2 reads somewhat oddly, especially given that the observations were archival. Better wording e.g.
"with the Hubble Space Telescope (HST) images taken ~8 d" (although this could also be a good place to indicate that the observations were obtained over ranges of dates)

2) A minor criticism is that while the use of GAIA results is discussed in detail, the importance of GAIA in enabling this kind of study is not credited/explained as prominently as it should be.

3) I find the whole paragraph "The constraint on the initial Lorentz factor of the jet cleanly separates GW170817 from the majority of Galactic systems..." to add little to the paper. These conclusions have already been reached previously from other constraints on ultra-relativistic jets in GRBs, so while it is still of interest that the jet in this case does indeed appear to have high initial Lorentz factor, I don't see that much is gained from the comparison to other jetted sources here.

4) Similarly, the discussion of the future role of JWST in making similar measurements, may be over-emphasised. Yes, JWST should do better than HST (assuming systematics can be controlled as well), but it is apparent that events even within ~100 Mpc and otherwise similar to GW170817 are going to be quite rare, and may happen only occasionally or not at all during the JWST lifetime. So some of this discussion might be trimmed e.g. it is first argued that the jet motion could be seen with JWST, but then conceded that VLBI is likely to be better for the afterglow, so perhaps no need to go round that circle.

Referee #2 (Remarks to the Author):

A. Summary of the key results

Mooley et al. present in this paper an estimate of the superluminal speed of the GW170817 jet based on HST observations taken 8 days after explosion. By combining this detection of superluminal motion, the first ever made based on optical data, with their already published radio data, they find a variable apparent speed (decreasing beta parameter), revise and refine estimates of the variable Lorentz factor of the wing and core of the jet, and set new and more cogent constraints on the orientation of the binary neutron star merger and observer viewing angle toward the structured jet axis.

B. Originality and significance: if not novel, please include reference

Not only is this augmenting significantly our understanding and knowledge of this single electromagnetic and gravitational study case of binary compact star merger, but it can also be used in the future for further improvement of the system as a standard siren and determination of the Hubble constant (as done for instance in Hototkezaka et al. 2019, Nature Astronomy, https://doi.org/10.1038/s41550-019-0820-1).
I believe the authors have illustrated in a clear way a clever novel method for superluminal motion determination and for derivation of physical parameters, supported by accurate modelling of the jet behaviour. The novelty of the tecniques, the importance of the findings and the relevance of the matter for several branches - both observational and theoretical - of astrophysics and cosmology warrant publication of this manuscript in Nature.

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

The authors have a number of previous papers on this subject, based on their accurate radio monitoring with various facilities. The data are accurate, well presented and all uncertainties spellt out realistically.

D. Appropriate use of statistics and treatment of uncertainties

The uncertainties, both statistical and systematic, were well discussed and are conservative.

E. Conclusions: robustness, validity, reliability

This work builds on previous work and brings about a new, unprecedented element of measuring in optical a quantity that is traditionally subject of radio observations. The treatment is sufficiently well described in Main Text and Methods that it appears solid and reliable.
With the advent of JWST, future more sensitive, high angular resolution infrared measurements of this kind of conveniently closeby systems will afford a detailed description of the jet geometry and dynamics and, when used together with gravitational data, will make it possible to set accurate constraints on the physics of these sources and on the cosmological parameters.

F. Suggested improvements: experiments, data for possible revision

I have several minor suggestions to improve readability:

- perhaps the authors should mention in the main text the distance they assume for the host galaxy, for immediate reference

- Caption to Figure 2: I would delete the expression "lab frame" as it is confusing. The authors explain clearly what is intended, without the need to use this expression.

G. References: appropriate credit to previous work?

All necessary prior work and papers were cited.

H. Clarity and context: lucidity of abstract/summary, appropriateness of abstract, introduction and conclusions

The information is conveyed in a clear and logic way and in good English and style.

Referee #3 (Remarks to the Author):

Dear editor,
The paper “Hubble Space Telescope Measurement of Superluminal Motion in GW170817” presents a new astrometric measurement using HST observation of the kilonova of GW170817. When combined with the VLBI observations this measurement provides additional constraints on the motion of the afterglow image on the plane of the sky, which in turn can be used to improve the constraints on the observing angle and on the jet geometry. I cannot comment on the quality of the HST and VLBI analysis, but if it is sound then this is a very impressive measurement. Unfortunately, I cannot give the same compliment to the theoretical analysis presented in the paper based on this measurement. I agree with the authors that their observations can, in principle, be used to derive improved constraints on the viewing angle. However, as I explain below, contrary to the authors claims, the constraints that they obtain on the viewing angle are not more stringent nor more accurate than those obtained in previous studies. Therefore, I cannot recommend publication of this paper to Nature. You may decide that the observation itself, without the constraints on the geometry, is enough for publication in Nature. But, if you decide that high impact implications of the observations are necessary, then even with a more accurate analysis I am not sure that the paper merits a publication in Nature. It depends on the quality of the constraints and their implications.

The authors present an analysis of the new data, with which they constrain the observing angle to be 17-24deg in 1sigma confidence level. However, this is just the statistic uncertainty which assumes that their model is accurate. However, as the authors properly acknowledge, they are making quite a few approximations. Using approximations is perfectly fine, but as I discuss next, in this case the uncertainty that the approximations introduce are at least comparable, and probably larger, than the statistical uncertainties. Thus, when the systematic uncertainties are added to the error budget, the constraints obtained in this paper are not better than those obtained in previous studies.

The uncertainties introduced by the authors approximations are as follows:
1) Point source approximations: the core of the jet is approximated to be a point source. The half opening angle of the jet is 3-4 deg and its full opening angle is 6-8deg. The location of the image centroid relative to the jet axis changes with time. Before the peak the image is closer to the observer (at a smaller angle) than the jet axis. As time propagates the image is moving from the side of the core that is closer to the observer to the side that is farther than the observer and after the peak it starts moving back towards the jet axis. Thus, the location of the image at 175day is not necessarily the jet axis. Given that the size of the core is 6-8deg, I estimate the uncertainty introduced by this approximation to be between one and a few degrees.
2) Neglecting sideways spreading: There is a common misconception that relativistic jets are not spreading until their Lorentz factor drops to ~3. The correct statement is that they do spread, but not as fast as predicted by the naive expectation that at all time \Gamma\theta_j=1. For example, Duffel & Laskar 2018 find that a jet with an initial opening angle of 0.05rad doubles its opening angle by the time that the Lorentz factor drops to 3. Thus, it is most likely that the spreading of the jet in GW170817 was substantial by day 230. Also, while the effect of the spreading on the light curve is

not very large, its effect on the image location can be substantial, as was shown recently by Fernández et al. (arXiv:2101.05138). Therefore, this approximation can introduce a significant error. Based on Fernández et al. I estimate that this approximation introduces an error of a few degrees.
3) As far as I understand Eq. 9 in the methods section ignores the fact that the emission does not depend only on x. In particular, there is a very strong dependence on Gamma itself. By ignoring that, the authors basically assume some unjustified assumption on the jet structure. Moreover, the authors' analysis ignores the best constrain that we have on the jet structure, which is the light curve. Thus, there is a bias, which I cannot estimate, between the authors analysis and the more accurate analysis of other papers that did fit the light curve and the VLBI data simultaneously.

Taking together all these sources of uncertainties the systematic error of the author's analysis is probably at least a few degrees. This should be added to the result reported by the authors, 17-24deg in 1sigma. This can be compared, for example, to the previous work of the leading author, Mooley et al. 2018, which carried out a simultaneous fit to the VLBI data and the light curve using a set of full hydrodynamic simulations that are not subject to any of the systematic uncertainties mentioned above. They find that simulations with viewing angle that ranges between 14 and 28 degrees agree with the data to within 2 sigma. Thus, although the new paper led by Mooley includes new data, the constraints on the viewing angle of the previous paper are at least as good, and most likely better.

Author Rebuttals to Initial Comments:

We thank the referees and editor for their valuable comments. We have color coded the revisions made to the paper (based on referee / editor comments).

Referees' comments:

Referee #1 (Remarks to the Author):

This is an interesting paper - the author's use a careful analysis of HST imaging of the kilonova counterpart of GW170817, to establish the exact location of the merger, and hence provide better constraints on the motion of the radio afterglow (driven by a relativistic jet) seen in VLBI observations. A crucial development is the availability of GAIA DR2 and EDR3 releases, which ensures an excellent tie between the radio and optical frames. The authors use this measurement to draw improved conclusions about the physical properties of the jet, including initial Lorentz factor, angle to the line of sight etc.

The paper is mostly well written and well referenced, and the discussion on interpretation convincing. The authors deserve particular credit for the very careful astrometric analysis. It is certainly impressive to get substantially better than 1mas positional uncertainties with standard HST imaging. Given the considerable interest in the jet produced by the GW170817 merger (thanks to the uniquely powerful observational data-sets), I would recommend the paper for publication in Nature, even though I don't really buy the narrative that it is a breakthrough that will necessarily open up future "new avenues" - in part because observing jet motion is itself not new, and in part because such opportunities are rare and the available options may differ from one case to another.

We agree that observing jet motion is not new, and we clarify that the “new avenues” is referring to the *optical* positional measurement that provides the location of the center of explosion (that is generally not accessible for GRBs). The center of explosion is important in establishing the baseline and provides complementary information compared to radio VLBI, which tracks the motion of the synchrotron source only.

We foresee this to be useful for not just for binary NS mergers (gravitational-wave sources and including kilonovae from LSST, for example), but also for orphan long-GRBs from LSST (where the

supernova marks the position of jet launch) and jetted tidal disruption events (where the accretion disk marks the position of jet launch).

We estimate that proper motion measurements with the JWST and radio VLBI will be able to provide reasonable constraints on the viewing angle and jet parameters for events occurring within 150-200 Mpc, implying detection rates of off-axis NS mergers and long-GRBs to be ~1 yr^-1 in each category. Hence, the opportunities may not be extremely rare. The rate of off-axis jetted TDEs is currently highly uncertain, but their optical and radio emissions are much brighter which enables proper motion measurements to larger distances.

Some other comments:

1) opening sentence of para 2 reads somewhat oddly, especially given that the observations were archival. Better wording e.g.

"with the Hubble Space Telescope (HST) images taken ~8 d" (although this could also be a good place to indicate that the observations were obtained over ranges of dates)

Agreed, thanks, fixed.

2) A minor criticism is that while the use of GAIA results is discussed in detail, the importance of GAIA in enabling this kind of study is not credited/explained as prominently as it should be.

We have now explicitly referenced GAIA in the second paragraph of the main text.

3) I find the whole paragraph "The constraint on the initial Lorentz factor of the jet cleanly separates GW170817 from the majority of Galactic systems..." to add little to the paper. These conclusions have already been reached previously from other constraints on ultra-relativistic jets in GRBs, so while it is still of interest that the jet in this case does indeed appear to have high initial Lorentz factor, I don't see that much is gained from the comparison to other jetted sources here.

Proper motion measurements of X-ray binaries and AGN jets exist, but this is the first such measurement for a GRB jet that implies Gamma>>10. Hence we feel that this paragraph nicely

summarizes the direct implication of this observation. If this is still problematic or unnecessary then we can remove or reword this paragraph.

4) Similarly, the discussion of the future role of JWST in making similar measurements, may be over-emphasised. Yes, JWST should do better than HST (assuming systematics can be controlled as well), but it is apparent that events even within ~100 Mpc and otherwise similar to GW170817 are going to be quite rare, and may happen only occasionally or not at all during the JWST lifetime. So some of this discussion might be trimmed e.g. it is first argued that the jet motion could be seen with JWST, but then conceded that VLBI is likely to be better for the afterglow, so perhaps no need to go round that circle.

We agree that GW170817-like events will be rare. In fact, many of the systems are expected to be in higher density environments (up to n~1 cm^-3 as for Galactic DNS systems per Pol 2020arXiv200803842P or at least n>~1e-2 cm^-3 like short-GRBs per Fong et al. 2015 and Duque et al. 2020A%26A...639A..15D). The high ISM density implies higher peak flux density, so proper motion can be detected to larger distances and/or larger observing angles, compared to GW170817.

If the afterglow is bright then JWST-only measurements of proper motion may be very useful. The main advantage in using JWST-only measurements (without relying on radio VLBI) is that uncertainties related to GAIA/ICRS-frame conversion are eliminated.

We estimate that, using JWST+VLBI, GW170817-like events up to ~150 Mpc can deliver strong constraints on the geometry and jet morphology. Given the LIGO-Virgo rate of ~500 Gpc^-3 yr^-1, this implies at least 1 event per year. We also note that the number of EM counterparts within 150 Mpc will increase substantially starting ~2025 when LIGO-India becomes operational. Hence, several NS merger events can be expected over the lifetime of JWST. Additionally, the case of long-GRBs and TDEs with off-axis jets was also presented above.

Referee #2 (Remarks to the Author):

A. Summary of the key results

Mooley et al. present in this paper an estimate of the superluminal speed of the GW170817 jet based on HST observations taken 8 days after explosion. By combining this detection of superluminal motion, the first ever made based on optical data, with their already published radio data, they find a variable apparent speed (decreasing beta parameter), revise and refine estimates of the variable Lorentz factor of the wing and core of the jet, and set new and more cogent constraints on the orientation of the binary neutron star merger and observer viewing angle toward the structured jet axis.

B. Originality and significance: if not novel, please include reference

Not only is this augmenting significantly our understanding and knowledge of this single electromagnetic and gravitational study case of binary compact star merger, but it can also be used in the future for further improvement of the system as a standard siren and determination of the Hubble constant (as done for instance in Hototkezaka et al. 2019, Nature Astronomy, https://doi.org/10.1038/s41550-019-0820-1).

I believe the authors have illustrated in a clear way a clever novel method for superluminal motion determination and for derivation of physical parameters, supported by accurate modelling of the jet behaviour. The novelty of the tecniques, the importance of the findings and the relevance of the matter for several branches - both observational and theoretical - of astrophysics and cosmology warrant publication of this manuscript in Nature.

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

The authors have a number of previous papers on this subject, based on their accurate radio monitoring with various facilities. The data are accurate, well presented and all uncertainties spellt out realistically.

D. Appropriate use of statistics and treatment of uncertainties

The uncertainties, both statistical and systematic, were well discussed and are conservative.

E. Conclusions: robustness, validity, reliability

This work builds on previous work and brings about a new, unprecedented element of measuring in optical a quantity that is traditionally subject of radio observations. The treatment is sufficiently well described in Main Text and Methods that it appears solid and reliable.

With the advent of JWST, future more sensitive, high angular resolution infrared measurements of this kind of conveniently closeby systems will afford a detailed description of the jet geometry and dynamics and, when used together with gravitational data, will make it possible to set accurate constraints on the physics of these sources and on the cosmological parameters.

F. Suggested improvements: experiments, data for possible revision

I have several minor suggestions to improve readability:

- perhaps the authors should mention in the main text the distance they assume for the host galaxy, for immediate reference

Done.

- Caption to Figure 2: I would delete the expression "lab frame" as it is confusing. The authors explain clearly what is intended, without the need to use this expression.

Done.

G. References: appropriate credit to previous work?

All necessary prior work and papers were cited.

H. Clarity and context: lucidity of abstract/summary, appropriateness of abstract, introduction and conclusions

The information is conveyed in a clear and logic way and in good English and style.

Thanks for these comments, we have addressed all of them.

Referee #3 (Remarks to the Author):

The main concern of the referee is on the theoretical analysis presented in the previous version of the paper, which we address using an extensive set of hydrodynamic simulations.

Dear editor,

The paper “Hubble Space Telescope Measurement of Superluminal Motion in GW170817” presents a new astrometric measurement using HST observation of the kilonova of GW170817. When combined with the VLBI observations this measurement provides additional constraints on the motion of the afterglow image on the plane of the sky, which in turn can be used to improve the constraints on the observing angle and on the jet geometry. I cannot comment on the quality of the HST and VLBI analysis, but if it is sound then this is a very impressive measurement. Unfortunately, I cannot give the same compliment to the theoretical analysis presented in the paper based on this measurement. I agree with the authors that their observations can, in principle, be used to derive improved constraints on the viewing angle. However, as I explain below, contrary to the authors claims, the constraints that they obtain on the viewing angle are not more stringent nor more accurate than those obtained in

previous studies. Therefore, I cannot recommend publication of this paper to Nature. You may decide that the observation itself, without the constraints on the geometry, is enough for publication in Nature. But, if you decide that high impact implications of the observations are necessary, then even with a more accurate analysis I am not sure that the paper merits a publication in Nature. It depends on the quality of the constraints and their implications.

The authors present an analysis of the new data, with which they constrain the observing angle to be 17-24deg in 1sigma confidence level. However, this is just the statistic uncertainty which assumes that their model is accurate. However, as the authors properly acknowledge, they are making quite a few approximations. Using approximations is perfectly fine, but as I discuss next, in this case the uncertainty that the approximations introduce are at least comparable, and probably larger, than the statistical uncertainties. Thus, when the systematic uncertainties are added to the error budget, the constraints obtained in this paper are not better than those obtained in previous studies.

The uncertainties introduced by the authors approximations are as follows:

1) Point source approximations: the core of the jet is approximated to be a point source. The half opening angle of the jet is 3-4 deg and its full opening angle is 6-8deg. The location of the image centroid relative to the jet axis changes with time. Before the peak the image is closer to the observer (at a smaller angle) than the jet axis. As time propagates the image is moving from the side of the core that is closer to the observer to the side that is farther than the observer and after the peak it starts moving back towards the jet axis. Thus, the location of the image at 175day is not

necessarily the jet axis. Given that the size of the core is 6-8deg, I estimate the uncertainty introduced by this approximation to be between one and a few degrees.

2) Neglecting sideways spreading: There is a common misconception that relativistic jets are not spreading until their Lorentz factor drops to ~3. The correct statement is that they do spread, but not as fast as predicted by the naive expectation that at all time \Gamma\theta_j=1. For example, Duffel & Laskar 2018 find that a jet with an initial opening angle of 0.05rad doubles its opening angle by the time that the Lorentz factor drops to 3. Thus, it is most likely that the spreading of the jet in GW170817 was substantial by day 230. Also, while the effect of the spreading on the light curve is not very large, its effect on the image location can be substantial, as was shown recently by Fernández et al. (arXiv:2101.05138). Therefore, this approximation can introduce a significant error. Based on Fernández et al. I estimate that this approximation introduces an error of a few degrees.

3) As far as I understand Eq. 9 in the methods section ignores the fact that the emission does not depend only on x. In particular, there is a very strong dependence on Gamma itself. By ignoring that, the authors basically assume some unjustified assumption on the jet structure. Moreover, the authors' analysis ignores the best constrain that we have on the jet structure, which is the light curve. Thus, there is a bias, which I cannot estimate, between the authors analysis and the more accurate analysis of other papers that did fit the light curve and the VLBI data simultaneously.

Taking together all these sources of uncertainties the systematic error of the author's analysis is probably at least a few degrees. This should be added to the result reported by the authors, 17-24deg in 1sigma. This can be compared, for example, to the previous work of the leading author, Mooley et al. 2018, which carried out a simultaneous fit to the VLBI data and the light curve using a set of full hydrodynamic simulations that are not subject to any of the systematic uncertainties mentioned above. They find that simulations with viewing angle that ranges between 14 and 28 degrees agree with the data to within 2 sigma. Thus, although the new paper led by Mooley includes new data, the constraints on the viewing angle of the previous paper are at least as good, and most likely better.

We agree with the referee on the uncertainties introduced by the approximations made in our semi-analytic model: (1) point-source emitter, (2) sideways spreading. In order to obtain robust constraints on the parameters we carried out hydrodynamic simulations.

Strengths of our hydro simulations are:

(1) We have a flexible jet structure model: it can accommodate a narrow Gaussian-like structure or a broad structure. Our posteriors do not suffer from the degeneracies in jet/environment parameters found in some of the previous works that have used hydrodynamical simulations.

(2) We self-consistently take into account lateral expansion of the jet.

(3) We have carried out ~1e6 direct simulations under an efficient Monte Carlo sampler (no extrapolations used).

From our simulations we find that the viewing angle (in the range 19-25 degrees, 90% confidence), which is more precise than found in previous works. We believe that our new method reduces the systematic error substantially. The viewing angle constraint from hydrodynamical analysis indicates that the point source model gives consistent results.

While the hydrodynamical analysis gives the most robust measurements, the point source model has been retained for illustrative purposes: it provides an intuitive link between the jet proper motion, the viewing angle and Lorentz factors associated with the structured jet.

Finally, we have removed the sentence saying that lateral expansion is only important after the jet has decelerated to Lorentz factor of 3. We have added the references to the papers (Duffel & Laskar 2018, Fernandez et al. 2022) mentioned by the referee.

Reviewer Reports on the First Revision:

Referees' comments:

Referee #1 (Remarks to the Author):

I have read the revised version of the paper, and the authors' response to the concerns raised by the reviewers. The paper is indeed improved, particularly thanks to the extra modelling work done in response to the pointed criticisms of referee 3. The increased precision and confidence with which important parameters of the GW170817 jet have been determined do seem to me to merit publication in Nature.

I still think that the discussion of rates in the paper and the response is rather optimistic. 500 BNS mergers/Gpc/yr is plausible, although would suggest that we were quite unlucky in the O3 run not to detect events at d<150 Mpc (190425 being possibly in that range). However, even then, to perform this kind of proper motion study, the GW network needs to be operating (obviously!) and one needs to locate an EM counterpart, for it to be within the JWST field of regard, for the jet axis angle to be in an interesting range (too great, and "afterglow" emission will only be detected when the emission becomes very isotropic, and the motion will not appear "superluminal"). So I doubt a rate as high as 1 "interesting" BNS/BHNS per year in the JWST lifetime will be achieved, but I agree that there may well be occasional opportunities at d<150 Mpc. I do take the points that cases with higher medium density, and cases where JWST observes both the KN (or SN) and the afterglow in the near-IR, would be particularly promising for the science return. [A final note, it is true that the addition of LIGO-India should improve the fraction for which EM counterparts are discovered, but otherwise should not increase the GW detection rate of nearby events significantly]

Small point: the right hand axis in Fig. 4 would be better if the minor tick-marks were removed, since it tends to give the impression to the reader that they continue to mean something, even after the astrometric precision reaches the assumed floor (even better would be to indicate on the axis where the floor was reached, rather than continuing to label tick marks).

Referee #2 (Remarks to the Author):

I have reviewed this resubmitted version of the ms "Hubble Space Telescope Measurement of Superluminal
2 Motion in GW170817" by Kunal Mooley and collaborators.

The main observational result of the paper consists in combining the HST milliarcsec-precision detection of the optical counterpart of GW170817 with radio VLBI observations to characterize in an accurate way the superluminal motion of the source. The theoretical implication is that some stringent constraints can be set on the jet geometry and physical behavior. Most importantly, an initial Lorentz factor larger than 40 is determined. While indirect observational evidence, based on fundamental physics considerations and on very early observations of afterglows, exists that the Lorentz factors of both long and short GRBs must be in the range 100-1000 - the highest for any

macroscopic astrophysical source - this is the first observational proof of such a high Gamma in any astrophysical source (XRB, AGN, TDE Gammas being generally no larger than 10-20).

Concerning my comments to the earlier version of this ms (explicit reference to distance of host galaxy), I confirm that they have been taken into account.
With respect to that ms version, the authors have accomplished extra work and simulations adopting the JEDI code for relativistic jets, including jet lateral expansion. My understanding is that the JEDI code confirms the results of the semi-analytic approach, although the finding of a Lorentz factor Gamma > 40 hinges essentially on the use of this code. Therefore, the robustness of the result depends on the reliability of the code, which should be guaranteed by the fact that one of the developers, Wenbin Lu, is a coauthor of this paper.

I believe that this result, if firm and solid, warrants publication in Nature. Perhaps the authors could be a bit more explicit (a sentence in the main text may suffice) about the specific superiority of the code in returning this parameter with respect to the semi-analytic approach.

All uncertainties throughout the paper have been spellt out clearly and they provide sufficient quantitative information.

Trivia: the sentence on page 4:

"If we simply assume that we observe the region (dominating the emission) satisfying Gamma = 1/thetaat any given time prior to the peak of the afterglow light curve."

does not seem to be connected in an obvious way with the preceding or following one, something is missing.

Referee #3 (Remarks to the Author):

I read the revised manuscript and the theoretical modeling is significantly improved. There are still some significant points that should be corrected, but after this is done the theoretical part of the paper will be sound. I find it hard to decide whether the manuscript (after corrected) will merit publication in Nature. Therefore, after discussing my comments on the theoretical modeling, I list below the pros and the cons for publication in Nature (as I see them) without making a final recommendation.

Comments on the theoretical modeling:
The main point that should be amended in the manuscript is the extent of conclusions that the authors derive from their analysis. The authors claim that they find three new constraints on GW170817 at a confidence level of 90%: the observing angle is 19-26deg, the jet core angle is 4-6deg and the initial Lorentz factor (LF) of the core is >40. This is however only correct if the jet structure follows the parametrization of equations 11 and 12 in the methods (which it almost certainly doesn't). Since we do not have a reliable a priori information on the jet structure the largest uncertainty is the effect of the jet structure on each of these constraints, which is unknown.

Out of these three constraints the limit on the Lorentz factor is clearly not robust. By the time that we see the jet core it already decelerated significantly, and I cannot see a way that the available observations constrain its initial LF to be larger than 40. Putting it differently, I expect that there are jets that are fully consistent with the observations with initial core LF that is much lower than 40 (e.g., if the core size of the LF is different than the core size of the energy distribution). They are improbeable (or not explored at all) when the parametrization and the priors taken by the authors are assumed. But this is merely a result of the authors assumptions and not a conclusion that is based on the observations. Therefore, the statements on the constraints of the core LF should be removed, unless the authors can show that this is indeed a robust constraint that is independent of their assumed jet structure, which as I wrote above, I doubt very much.
The constraints on the viewing angle and the jet opening angle are much less sensitive to the jet structure and are therefore more robust. The authors still should emphasize in the main text that the confidence level that they report are under the assumption that eqs. 11 and 12 describe the jet structure reasonably well.
Additional point that the authors should be accurate about is the description of the method that they are using. The code used by the authors is not really a relativistic hydro axisymmetric code as implicitly advertised in the current manuscript. A hydro axisymmetric code solves the hydrodynamic equations under the assumption of axisymmetric and is therefore 2D in nature. The code used by the authors solves the 1D equations derived by Kumar and Granot in 2003, which are an approximation (of whom accuracy was never properly tested to the best of my knowledge) of the full 2D equations.
Final minor issue is the statement in line 80 that the material that gave rise to the gamma-ray emission had an initial Lorentz factor of 5-7. This statement is unfounded. We have only a lower limit on the Lorentz factor of the gamma-ray emitting region. There is no robust upper limit and the LF in the gamma-ray emitting region can be significantly larger than 7.

Pros and cons for publication in Nature:
Pros: The determination of the viewing angle and the jet opening angle of binary neutron star mergers is of great importance, both for an independent measurement of the Hubble constant and for deciphering the physical processes that take place during and following the merger. This paper presents a significant contribution in that front. Most important is the unique HST observation and its astrometric calibration with the VLBI measurements. This is a demonstration of the great potential of similar methods for future merger observations with JWST. The new observations also enable obtaining tighter constraints on the viewing angle of the jet from GW170817.

Cons: The benefit from the improved constraints on the viewing angle are incremental when compared, for example, to the original Nature paper on the VLBI measurements. The VLBI data provided conclusive evidence of a successful jet, constraints on its properties and a solution for the entire system geometry, thereby landing a significant support to the relation of mergers to short GRBs. The new manuscript provides an incremental improvement of those constrainers. Similarly, the VLBI data led to a significant improvement of the constraints on H0, reducing the error from a measurement that is based on GW170817 from ~20% to 7% (Hotokezaka et al 2019). The improved constraints on the viewing angle obtained in this manuscript have a negligible effect on the measurement of H0, since the error is this measurement is not dominated anymore by the viewing angle (it is now dominated by the peculiar velocity of the host and by the GW signal to noise ratio).

I find it hard to decide if these results are enough to merit publication in Nature, but I believe that I gave enough information for the editor to decide.

Author Rebuttals to First Revision:

Referee #1 (Remarks to the Author):

I have read the revised version of the paper, and the authors' response to the concerns raised by the reviewers. The paper is indeed improved, particularly thanks to the extra modelling work done in response to the pointed criticisms of referee 3. The increased precision and confidence with which important parameters of the GW170817 jet have been determined do seem to me to merit publication in Nature.

I still think that the discussion of rates in the paper and the response is rather optimistic. 500 BNS mergers/Gpc/yr is plausible, although would suggest that we were quite unlucky in the O3 run not to detect events at d<150 Mpc (190425 being possibly in that range). However, even then, to perform this kind of proper motion study, the GW network needs to be operating (obviously!) and one needs to locate an EM counterpart, for it to be within the JWST field of regard, for the jet axis angle to be in an interesting range (too great, and "afterglow" emission will only be detected when the emission becomes very isotropic, and the motion will not appear "superluminal"). So I doubt a rate as high as 1 "interesting" BNS/BHNS per year in the JWST lifetime will be achieved, but I agree that there may well be occasional opportunities at d<150 Mpc. I do take the points that cases with higher medium density, and cases where JWST observes both the KN (or SN) and the afterglow in the

near-IR, would be particularly promising for the science return. [A final note, it is true that the addition of LIGO-India should improve the fraction for which EM counterparts are discovered, but otherwise should not increase the GW detection rate of nearby events significantly]

Small point: the right hand axis in Fig. 4 would be better if the minor tick-marks were removed, since it tends to give the impression to the reader that they continue to mean something, even after the astrometric precision reaches the assumed floor (even better would be to indicate on the axis where the floor was reached, rather than continuing to label tick marks).

Author Response: We thank the referee for the cautionary notes on the actual detection rate of nearby BNS and BHNS mergers by JWST (e.g., the field of regard at a given time is only 40% of the sky). The physical rates are currently highly uncertain and only the upcoming LIGO-Virgo observing runs will inform us about the actual rates. At this point, we believe that our rate estimate is not overly optimistic. There is a good chance that JWST astrometry of a nearby kilonova in the future would greatly improve the precision of VLBI proper motion measurements of the afterglow emission. In any case, the rate estimate is not the main part of the paper, which focuses on GW170817 and the new technique based on the implication of precision astrometry of the kilonova emission.

Following the referee's suggestion, we have removed the minor ticks in the right axis of Fig 4.

nature portfolio

Referee #2 (Remarks to the Author):

I have reviewed this resubmitted version of the ms "Hubble Space Telescope Measurement of Superluminal Motion in GW170817" by Kunal Mooley and collaborators.

The main observational result of the paper consists in combining the HST milliarcsec-precision detection of the optical counterpart of GW170817 with radio VLBI observations to characterize in an accurate way the superluminal motion of the source. The theoretical implication is that some stringent constraints can be set on the jet geometry and physical behavior. Most importantly, an initial Lorentz factor larger than 40 is determined. While indirect observational evidence, based on fundamental physics considerations and on very early observations of afterglows, exists that the Lorentz factors of both long and short GRBs must be in the range 100-1000 - the highest for any macroscopic astrophysical source - this is the first observational proof of such a high Gamma in any astrophysical source (XRB, AGN, TDE Gammas being generally no larger than 10-20).

Concerning my comments to the earlier version of this ms (explicit reference to distance of host galaxy), I confirm that they have been taken into account.

With respect to that ms version, the authors have accomplished extra work and simulations adopting the JEDI code for relativistic jets, including jet lateral expansion. My understanding is that the JEDI code confirms the results of the semi-analytic approach, although the finding of a Lorentz factor Gamma > 40 hinges essentially on the use of this code. Therefore, the robustness of the result depends on the reliability of the code, which should be guaranteed by the fact that one of the developers, Wenbin Lu, is a coauthor of this paper.

I believe that this result, if firm and solid, warrants publication in Nature. Perhaps the authors could be a bit more explicit (a sentence in the main text may suffice) about the specific superiority of the code in returning this parameter with respect to the semi-analytic approach.

All uncertainties throughout the paper have been spellt out clearly and they provide sufficient quantitative information.

Trivia: the sentence on page 4:

"If we simply assume that we observe the region (dominating the emission) satisfying Gamma = 1/theta at any given time prior to the peak of the afterglow light curve."

does not seem to be connected in an obvious way with the preceding or following one, something is missing.

Author Response: The advantage of the 1D JEDI code over the semi-analytic point-source (or one-zone) model is that it has the full jet angular structure. This allows us to directly constrain the jet angular structure (although within our power-law jet parameterization in eq. 11 and 12) by fitting to the full dataset, which is not possible for the semi-analytic point-source model. Another advantage of our 1D JEDI code is that it is much faster than other 2D codes in that each simulation only takes a few seconds on a CPU core, which makes it possible to do millions of simulations to fit the data in a Monte Carlo manner.

Due to the 1500 word limit for the Main Text, and since the "specific superiority of the code" statements are relevant only for the very small subset of readers having the appropriate technical expertise, we have opted to place these statements in the Methods section.

Thank you for pointing out the seemingly misplaced sentence "If we simply assume…". There was a typo in this sentence that we have now fixed.

Referee #3 (Remarks to the Author):

I read the revised manuscript and the theoretical modeling is significantly improved. There are still some significant points that should be corrected, but after this is done the theoretical part of the paper will be sound. I find it hard to decide whether the manuscript (after corrected) will merit publication in Nature. Therefore, after discussing my comments on the theoretical modeling, I list below the pros and the cons for publication in Nature (as I see them) without making a final recommendation.

Comments on the theoretical modeling:

The main point that should be amended in the manuscript is the extent of conclusions that the authors derive from their analysis. The authors claim that they find three new constraints on GW170817 at a confidence level of 90%: the observing angle is 19-26deg, the jet core angle is 4-6deg and the initial Lorentz factor (LF) of the core is >40. This is however only correct if the jet structure follows the parametrization of equations 11 and 12 in the methods (which it almost certainly doesn't). Since we do not have a reliable a priori information on the jet structure the largest uncertainty is the effect of the jet structure on each of these constraints, which is unknown.

We agree with the referee on potential systematic errors, i.e., the true jet structure might not be well described by our power-law model. We have added in the Main Text that our conclusion of "core LF>40" is obtained based on the power-law jet structure model. In the revised Methods section, we also emphasized that the flexibility in the power-law parametrization of the jet structure (with 5 parameters) is superior to the Gaussian jet model (with 3 parameters) or the "boosted fireball" model (with 4 parameters) commonly adopted in the literature. We believe our power-law parameterization should encompass a broader range of physical jet structures than other alternatives in the literature. Additionally, recent numerical simulations (Gottlieb et al. 2021, MNRAS, 500, 3511) of the jet propagation inside the merger ejecta strongly support the power-law model.

It is in principle possible to add more parameters to the model to make it even more flexible. However, since the current model already provides a good fit to the full dataset (with reduced chi-squared sufficiently close to 1), adding more parameters would cause the issue of overfitting the data. If one thinks about the dataset more carefully, there are only ~7 independent measurements (despite the fact that there are ~100 data points): (1) peak time of the lightcurve, (2) peak flux of the lightcurve, (3) flux rising slope before the peak time, (4) flux decaying slope after the peak time, (5) spectral slope, (6) and (7) two independent proper motions. Thus, models with much more than 7 parameters should not be considered. Our jet model has 8 parameters (apart from the source distance, which is independently constrained by additional information of the host galaxy), but one of them --- the initial core Lorentz factor --- only has a lower limit. If we were to

add more parameters, e.g., the core size in Lorentz factor distribution being different than the core size in energy distribution (as the referee mentioned), then this will cause some of the parameters not to be strongly constrained by the data and hence the conclusion will be even more dependent on the priors. We don't foresee a better way of solving the problem in the near future (without fully understanding the physics that determines the jet structure --- but those pieces of physics are highly uncertain due to the lack of data).

To summarize, our model parametrization is optimally chosen based on physical motivation as well as the available data, and the effects of the priors on the inferred posterior distribution are minimized by our choices.

Out of these three constraints the limit on the Lorentz factor is clearly not robust. By the time that we see the jet core it already decelerated significantly, and I cannot see a way that the available observations constrain its initial LF to be larger than 40. Putting it differently, I expect that there are jets that are fully consistent with the observations with initial core LF that is much lower than 40 (e.g., if the core size of the LF is different than the core size of the energy distribution). They are improbeable (or not explored at all) when the parametrization and the priors taken by the authors are assumed. But this is merely a result of the authors assumptions and not a conclusion that is based on the observations. Therefore, the statements on the constraints of the core LF should be removed, unless the authors can show that this is indeed a robust constraint that is independent of their assumed jet structure, which as I wrote above, I doubt very much.

We agree with the referee that the conclusion of "core LF > 40" depends on our parametrization and prior. This point has now been explicitly pointed out in the Methods Section 7 of the revised paper. Since the 0d-75d mean apparent speed is ~7c and the 0d-230d mean apparent speed is ~5c, the inferred "core LF > 40" comes from extrapolation in polar angle and in time. Below, we explain how the core LF is constrained by our sophisticated modeling and why the lower limit should be robust.

The key of the argument is that the afterglow flux rises roughly as a power-law of F ~ t^{~0.8} from t~10d (initial detection) to about t~155d (the peak time). The physical picture is that from 10d to 155d, the observer sees the part of the jet that is progressively closer to the axis, i.e. at smaller and smaller polar angles. In order to produce such a large dynamical range of a factor of 15 in time, the part of the jet dominating the afterglow flux at t=10d must be sufficiently far from the jet core and initially sufficiently relativistic such that its beaming angle only reaches our line of sight at t=10d (but not much before that).

However, suppose the initial core LF is only 10 and the core angular size is 5 degrees (note that the core angular size is constrained by the lightcurve decay after the peak time), then the material at >10 degrees away from the axis would only be mildly relativistic. This means that the region beyond 10 degrees away from the jet axis would have been seen by the observer (located at a viewing angle of ~20 degrees away from the jet axis) all the time. This scenario cannot produce the large dynamical range of a factor of 15 in the rise time of the lightcurve. Moreover, it will also require fine-tuning to reproduce the mean apparent speed of ~7c between 0d and 75d.

The constraints on the viewing angle and the jet opening angle are much less sensitive to the jet structure and are therefore more robust. The authors still should emphasize in the main text that the confidence level that they report are under the assumption that eqs. 11 and 12 describe the jet structure reasonably well.

Additional point that the authors should be accurate about is the description of the method that they are using. The code used by the authors is not really a relativistic hydro axisymmetric code as implicitly advertised in the current manuscript. A hydro axisymmetric code solves the hydrodynamic equations under the assumption of axisymmetric and is therefore 2D in nature. The code used by the authors solves the 1D equations derived by Kumar and Granot in 2003, which are an approximation (of whom accuracy was never properly tested to the best of my knowledge) of the full 2D equations.

We have now added a sentence in the Main Text clarifying/emphasizing the power-law parametrization of the jet structure. We have also added a sentence in the Methods section 7 clarifying that, although the jet lateral expansion is intrinsically a 2D problem, the fact that the forward shock-compressed region is very thin in the radial direction motivates an effective 1D solution. We further note that the Jedi code has been tested against another code JetFit that is based on full 2D simulations (see Fig. 12 in the code method paper arXiv:2005.10313). Due to the 1500 word limit we prefer not to place all technical details in the Main Text.

Final minor issue is the statement in line 80 that the material that gave rise to the gamma-ray emission had an initial Lorentz factor of 5-7. This statement is unfounded. We have only a lower limit on the Lorentz factor of the gamma-ray emitting region. There is no robust upper limit and the LF in the gamma-ray emitting region can be significantly larger than 7.

We agree with the referee that the initial LF of the gamma-ray emitting region only has a lower limit (based on the compactness argument). Previous authors' claim that this LF is 5 to 7 was based on additional assumptions on the jet angular structure. We have removed this ambiguous sentence for consistency.

Pros and cons for publication in Nature:

Pros: The determination of the viewing angle and the jet opening angle of binary neutron star mergers is of great importance, both for an independent measurement of the Hubble constant and for deciphering the physical processes that take place during and following the merger. This paper presents a significant contribution in that front. Most important is the unique HST observation and its astrometric calibration with the VLBI measurements. This is a demonstration of the great potential of similar methods for future merger observations with JWST. The new observations also enable obtaining tighter constraints on the viewing angle of the jet from GW170817.

Cons: The benefit from the improved constraints on the viewing angle are incremental when compared, for example, to the original Nature paper on the VLBI measurements. The VLBI data provided conclusive evidence of a successful jet, constraints on its properties and a solution for the entire system geometry, thereby landing a significant support to the relation of mergers to short GRBs. The new manuscript provides an incremental improvement of those constrainers. Similarly, the VLBI data led to a significant improvement of the constraints on H0, reducing the error from a measurement that is based on GW170817 from ~20% to 7% (Hotokezaka et al 2019). The improved constraints on the viewing angle obtained in this manuscript have a negligible effect on the measurement of H0, since the error is this measurement is not dominated anymore by the viewing angle (it is now dominated by the peculiar velocity of the host and by the GW signal to noise ratio).

This paper does not focus on improving the constraints on H0. Instead, the main theoretical results are: we have improved the viewing angle constraints from "14-28 deg" to "19-25 deg" (shrinking the error by a factor of 2), and we have updated the core Lorentz factor constraints from ">4" to ">40" (improving by a factor of 10). We have also presented a new method (based on hydrodynamic + Monte Carlo simulations) that can robustly constrain the jet structure as well as the other geometrical parameters of the system --- paving the road for future precise measurements of H0.

On the observational side, we have demonstrated the crucial role of the astrometric position of the kilonova. In the near future, JWST observations of kilonovae from other neutron star mergers will achieve better astrometric precision than in this work, and this combined with VLBI measurements will extract the maximum amount of information from such rare but precious events (and hence reduce the otherwise very large errors in H0 for distant events).