Extracting the Gravitational Recoil from Black Hole Merger Signals

Vijay Varma, Maximiliano Isi, and Sylvia Biscoveanu

Phys. Rev. Lett. 124, 101104 — Published 13 March 2020

DOI: 10.1103/PhysRevLett.124.101104
Extracting the gravitational recoil from black hole merger signals

Vijay Varma, 1,* Maximiliano Isi, 2,+ and Sylvia Biscoveanu 2,†

1 TAPIR 350-17, California Institute of Technology, 1200 E California Boulevard, Pasadena, CA 91125, USA
2 LIGO Laboratory, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

(Dated: February 13, 2020)

Gravitational waves carry energy, angular momentum, and linear momentum. In generic binary black hole mergers, the loss of linear momentum imparts a recoil velocity, or a “kick”, to the remnant black hole. We exploit recent advances in gravitational waveform and remnant black hole modeling to extract information about the kick from the gravitational wave signal. Kick measurements such as these are astrophysically valuable, enabling independent constraints on the rate of second-generation merger. Further, we show that kicks must be factored into future ringdown tests of general relativity with third-generation gravitational wave detectors to avoid systematic biases. We find that, although little information can be gained about the kick for existing gravitational wave events, interesting measurements will soon become possible as detectors improve. We show that, once LIGO and Virgo reach their design sensitivities, we will reliably extract the kick velocity for generically precessing binaries—including the so-called superkicks, reaching up to 5000 km/s.

Introduction As existing gravitational wave (GW) detectors, Advanced LIGO [1] and Virgo [2], approach their design sensitivities, they continue to open up unprecedented avenues for studying the astrophysics of black holes (BHs). One such opportunity is to experimentally study the gravitational recoil in binary BH mergers. It is well known that GWs carry away energy and angular momentum, causing the binary to shrink during the inspiral; however, in addition to this, GWs also carry away linear momentum, shifting the binary’s center of mass in the opposite direction [3–6]. Learning about this effect from GW data would be of high astrophysical significance.

During a binary BH coalescence, most of the linear momentum is radiated near the time of the merger [7–13], resulting in a recoil or “kick” imparted to the remnant BH. The end state of the remnant is entirely characterized by its mass \(m_1\) and \(m_2\), spin \(\mathbf{\chi_1}\) and \(\mathbf{\chi_2}\); all additional complexities (“hair”) [14, 15] are dissipated away in GWs during the ringdown stage that follows the merger. The remnant mass and spin have already been measured from GW signals and used to test general relativity [16–23]. However, a measurement of the kick has remained elusive.

Measuring the kick velocity from binary BHs would have important astrophysical applications—particularly for precessing binaries, where the component BH spins have generic orientations with respect to the orbit. For these systems, the spins interact with the orbital angular momentum as well as with each other, causing the orbital plane to precess [24]. The kick velocity of these systems can reach up to 5000 km/s for certain fine-tuned configurations [25–30], earning them the moniker of “superkicks”. Such velocities are larger than the escape velocity of even the most massive galaxies. This can have dramatic consequences for mergers of supermassive BHs residing at galactic centers. The remnant BH can be significantly displaced or ejected [31, 32], impacting the galaxy’s evolution [33–35], and event rates [36] for the future LISA mission [37].

The kick velocity is also important for second-generation stellar-mass mergers, where one of the component BHs originated in a previous merger. This scenario has attracted much attention recently [38–45] because the GW event GW170729 [46, 47] may have a component BH that is too massive to originate in a supernova explosion [48, 49], the typical formation scenario for stellar-mass BHs. A second-generation merger could explain this, as the first merger would have led to a remnant BH more massive than the original stellar-mass progenitors. If we could measure the kick velocity from GW signals, we could place independent constraints on rates of second-generation mergers.

In this Letter, we present the first method to extract the kick magnitude and direction from generically precessing GW signals. We demonstrate that kicks will be measured reliably once LIGO and Virgo reach their design sensitivities, and possibly even earlier. The key is being able to accurately measure the spins of the individual BHs in the binary, from which the kick velocity can be inferred. This is made possible by two advances in GW modeling achieved in the past few years: numerical relativity (NR) surrogate models for both gravitational waveforms [50, 51] and remnant properties [50, 52], suitable for generically precessing binaries. These models capture the effects of spin precession at an accuracy level comparable to the NR simulations, and are the most accurate models currently available in their regime of validity [50].

Methods We use the surrogate waveform model NRSur7dq4 [50] to analyze public GW data [46, 53], as well as simulated signals in synthetic Gaussian noise corresponding to the three-detector advanced LIGO-Virgo network at its design sensitivity [54–56].

NRSur7dq4 is trained on NR simulations with mass ratios \(q = m_1/m_2 \leq 4\) and component spin magnitudes

\[ v_{\text{varma@caltech.edu}} \]
\[ \text{maxisi@mit.edu; NHFP Einstein fellow} \]
\[ \text{sbsico@mit.edu} \]
$|\chi_1|, |\chi_2| \leq 0.8$ with generic spin directions. The index 1 (2) corresponds to the heavier (lighter) BH, with $m_1 \geq m_2$. The spin components are specified at a reference GW frequency $f_{\text{ref}} = 20$ Hz, in a source frame defined as follows: the $z$-axis lies along the instantaneous orbital angular momentum, the $x$-axis points from the lighter to the heavier BH, and the $y$-axis completes the right-handed triad. We use all available spin-weighted spherical harmonic modes for this model ($\ell \leq 4$). The inclination angle $i$ and azimuthal angle $\phi_{\text{ref}}$ indicate the location of the observer in the sky of the source, and take different values for each injection.

We obtain Bayesian posteriors on the signal parameters using the LALINFERENCE package [57], part of the LIGO Algorithm Library (LAL) Suite [58]. Because of restrictions on the duration of the injection, we choose to analyze data with a minimum Fourier frequency $f_{\text{low}} = 20$ Hz. Waveform length also restricts the higher-order-mode content of our NRSur7dq4 injections and templates in such way that modes with azimuthal harmonic number $m$ contribute with a starting frequency $f_{\text{min}}^{(m)} = m f_{\text{low}} / 2$. This means that our sensitivity projections are conservative, as detectors are expected to access information starting at lower frequencies than our simulations. NR injections are handled via the dedicated infrastructure in LAL [59].

Given the posteriors distributions for the component parameters $\Lambda = \{m_1, m_2, \chi_1, \chi_2\}$, we use the remnant-properties surrogate model NRSur7dq4Remnant [50] to predict the mass $m_f$, spin $\chi_f$, and kick velocity $\mathbf{v}_f$ of the remnant. Trained on the same simulations as NRSur7dq4, NRSur7dq4Remnant uses Gaussian Process Regression [52, 60] to model the remnant properties. NRSur7dq4Remnant improves upon previous remnant properties models by at least an order of magnitude in accuracy [50]. NRSur7dq4Remnant models the full kick velocity vector and can, therefore, predict both the kick magnitude and direction. To assess whether a meaningful kick measurement has been made, we compare this posterior distribution with the corresponding effective prior distribution, estimated by drawing component parameters $\Lambda$ from the prior. The priors used for the component parameters are described in the Supplement [61].

Comparison to previous methods—The challenge of measuring the kick velocity from GW signals has been tackled before. The recoil may Doppler shift the final portion of the GW signal. Ref. [62] showed that it will not be possible to measure the kick velocity from this effect alone until third-generation GW detectors become active in the 2030s [63–66]. Ref. [67] proposed a method to extract the kick based on direct comparison against NR simulations, showing that current detectors are sufficient for a kick measurement; however, that study was restricted to non-precessing systems, where we do not expect very large kicks ($\gtrsim 300$ km/s). Ref. [68] compared GW150914 data against NR simulations, including precessing systems, to place bounds on the kick of GW150914. However, both Refs. [67] and [68] relied on a discrete bank of NR simulations, which does not allow for a full exploration of the multidimensional posterior for the system parameters.

Our procedure for measuring kicks is more widely applicable than those of Refs. [62, 67, 68] in a few ways. Since the surrogate models accurately reproduce the NR simulations, we are potentially sensitive to effects of the recoil other than simple Doppler shifts (e.g. acceleration of the center of mass near merger [7–13], or phase aberration [69]). Therefore, rather than rely on Doppler shifts in the

![Figure 1](image1.png)

Figure 1. Kick magnitude measurement using different remnant BH models in conjunction with the NRSur7dq4 waveform model, for an injected NR signal at the design sensitivity of LIGO and Virgo. The signal parameters are given in the inset text and the corresponding kick magnitude is indicated by the dashed gray line. The effective prior is shown as a dashed histogram.

![Figure 2](image2.png)

Figure 2. Kick measurement for the GW events GW150914 and GW170729. We find only marginal differences between the posterior and the effective prior, suggesting that very little information about the kick can be gained from these events. We quantify this via the KL divergence, shown in the upper-right insets.
We inject \textit{NRSur7dq4} signals corresponding to the three markers and measure the kick velocity using our method. The posteriors for the measured kick magnitudes are shown in the right panel; the colors correspond to the markers in the left panel. The true kick magnitudes are shown as dashed vertical lines, and the effective prior is shown as a dashed histogram. In all three cases, the kick velocity is well recovered.

\textit{NR simulation}—We first demonstrate our method by injecting an NR waveform into noise from a simulated LIGO-Virgo network at design sensitivity. The signal parameters are given in the inset text of Fig. 1. We choose a luminosity distance consistent with that of GW150914 [70], \(d_L = 440\) Mpc. Using the \textit{NRSur7dq4} waveform model, we recover the signal with a signal-to-noise ratio (SNR) of 102. Here and throughout this paper, reported SNRs correspond to the network matched-filter, maximum \(a\)-posteriori values. Further, all masses are reported in the detector frame.

Our method successfully recovers the injected kick magnitude, as seen from the posterior in Fig. 1. We find that the use of the remnant surrogate model \textit{NRSur7dq4Remnant} is critical. To show this, we consider an alternate kick formula developed in Refs. [7–10, 71], as summarized in [72]. Using this formula (which we label “RIT”) on the same \textit{NRSur7dq4} samples yields a totally uninformative posterior on the kick. We note that the NR waveform used here (with identifier SXS:BBH:0137 [73–75]) was not used to train the surrogate models.

\textit{Kick measurement from existing GW events}—Next, we apply our method to GWTC-1 [46] by reanalyzing the publicly available data released by the LIGO-Virgo Collaborations [53, 76]. Figure 2 shows the posteriors we recover for the kick magnitude for the GW150914 [70] and GW170729 [46] events. These are compared with the prior for the kick magnitude. Not much information about the kick can be gained for the GWTC-1 events, as measured by the Kullback–Leibler (KL) divergence from the prior. A KL divergence of 0.16 and 0.22 bits. This can be compared with Ref. [46] where \(\sim 0.13\) bits of information is shown in the left panel of Fig. 3 (see Ref. [82] for visualizations of the sinusoidal dependence

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure3}
\caption{A demonstration of the measurability of superkicks at the design sensitivity of LIGO and Virgo. We consider the fine-tuned binary configuration shown in the inset of the left panel. The kick velocity has a sinusoidal dependence on the angle \(\alpha\) as shown in the left panel. We inject the \textit{NRSur7dq4} signals corresponding to the three markers and measure the kick velocity using our method. The posteriors for the measured kick magnitudes are shown in the right panel; the colors correspond to the markers in the left panel. The true kick magnitudes are shown as dashed vertical lines, and the effective prior is shown as a dashed histogram. In all three cases, the kick velocity is well recovered.}
\end{figure}
and superkicks. We use \texttt{NRSur7dq4Remnant} to find the value of \( \alpha \) that yields the maximum kick for the chosen spin magnitude. We consider the \( \alpha \) values that lead to the superkick (\( \| \mathbf{v}_f \| = 1814 \) km/s), half of the superkick (\( \| \mathbf{v}_f \| = 907 \) km/s), and a minimum kick magnitude (\( \| \mathbf{v}_f \| = 35 \) km/s) [83]. The right-panel of Fig. 3 shows the kick magnitude posteriors obtained by applying our method to \texttt{NRSur7dq4} injections corresponding to those three configurations. We are able to clearly distinguish the kick velocity between these injections, which have otherwise nearly identical parameters. This is in agreement with Ref. [29], where a mismatch comparison was used to assess distinguishability between similar configurations. The kick magnitude can be reliably recovered in all three cases, demonstrating our ability to accurately measure superkicks at the design sensitivity of LIGO and Virgo.

### Measuring kicks from generic systems

The large kicks explored in the previous section required some fine-tuning of the component parameters. For generic systems that are more likely to occur in nature, typical kicks are much smaller [9, 84]. We now explore the measurability of the kick velocity of arbitrary systems by injecting randomly chosen signals and studying the recovered kicks. We perform 60 \texttt{NRSur7dq4} injections uniformly sampled from mass ratios \( q \in [1, 3] \), spin magnitudes \( |\chi_1|, |\chi_2| \in [0, 0.8] \), arbitrary spin directions, total masses \( M \in [70, 150] \), luminosity distances \( d_L \in [400, 2000] \) Mpc, inclination angles \( \iota \in [0, \pi] \), and reference phases \( \phi_{\text{ref}} \in [0, 2\pi] \). These ranges are chosen to fall within the training region of current surrogate models [50].

The recovered posteriors for the kick magnitude are shown in the top-panel of Fig. 4 for a subset of 10 representative cases. Our method reliably recovers the kick magnitude for these generic systems; biases away from the true value are consistent with statistical error, as shown in the Supplement [61].

Figure 5 shows the measurement uncertainty in the recovered kick magnitude for all 60 random cases. In general, a larger SNR leads to a better measurement of the kick magnitude, but the specific choice of injected parameters also plays a role, causing the spread in Fig. 5. In some cases a good measurement can be made at SNRs as low as 20. This suggests that kick velocities can be measured using our method even before LIGO and Virgo achieve their design sensitivities.

Our method measures the full kick vector. To gauge how well we can recover the kick direction, we consider the angle between the measured kick direction \( \hat{\mathbf{v}}_f \) and the injected kick direction \( \hat{\mathbf{v}}_{f\text{inj}} \), namely \( \cos^{-1}(\hat{\mathbf{v}}_f \cdot \hat{\mathbf{v}}_{f\text{inj}}) \). We refer to this angle as the kick-direction “bias”; for the true injection value, this angle is zero. The bottom-panel of Fig. 4 shows the distribution of this quantity as derived from the full kick-vector posteriors corresponding to the same 10 cases as the top-panel. For all cases where the injected kick magnitude is \( \gtrsim 300 \) km/s we recover the kick direction, i.e. \( \cos^{-1}(\hat{\mathbf{v}}_f \cdot \hat{\mathbf{v}}_{f\text{inj}}) \approx 0 \). For smaller kick magnitudes, \texttt{NRSur7dq4Remnant} is known to have larger intrinsic errors in the kick direction [50], which results in correspondingly higher kick-direction posterior biases. This comes from similar errors in the underlying NR simulations on which the surrogate model is trained [50], and should thus be fixed by more accurate simulations.
Applications— Based on Fig. 5, we generally expect an uncertainty of $\lesssim 500 \text{ km/s}$ at SNR $\sim 50$ in measuring the kick magnitude at the 68.27\% credible level ($\sim 1\sigma$). This can be used to place meaningful constraints on the retention rate of the remnant for different types of galaxies. For example, a kick measurement of the type shown in Fig. 1 would lead us to conclude that the remnant of such a binary would be ejected from most globular clusters, which typically have escape velocities $\lesssim 50 \text{ km/s}$ [31, 85].

In Fig. 6, we use the projection of the full kick vector along the line of sight to compute the kick’s effect on the remnant BH mass [62] that would be inferred by an analysis of the Doppler-shifted ringdown signal. As detectors become more sensitive, this effect will need to be accounted for to avoid systematic biases in tests of general relativity, especially for third-generation detectors and remnants with large kick velocities along the line of sight. Our method will prevent these issues, as we discuss in the Supplement [61].

Conclusion— We present the first method to accurately extract both the kick magnitude and direction of generically precessing binary BHs. This is made possible by recent NR surrogate models for the gravitational waveform and properties of the merger remnant (Fig. 1).

We find that the SNR for existing GWTC-1 events is not sufficient to make a confident measurement of the kick velocity (Fig. 2). However, our results indicate that the kick velocity will be reliably measured once LIGO and Virgo reach their design sensitivities. This includes systems with arbitrary parameters (Fig. 4), as well as configurations fine-tuned to produce superkicks with $|v_f| \sim 1000 \text{ km/s}$ (Fig. 3). Measuring such kicks was previously estimated to be only possible with third-generation GW detectors [62]. On the contrary, we find that accurate waveform and remnant surrogate models will soon enable this with existing detectors (Fig. 5). This is in agreement with Ref. [67], which made compatible predictions for nonprecessing systems, for which $|v_f| \lesssim 300 \text{ km/s}$.

Kick measurements obtained with our method can be used to place independent constraints on the retention rate of the remnant BH in binary BH mergers, which is directly related to the rate of second-generation mergers. In addition, we show (Fig. 6 and Supplement [61]) that kicks must be factored into ringdown tests of general relativity with third-generation GW detectors to avoid systematic biases.

In this study, we focused on projected measurements by LIGO and Virgo at design sensitivity. Since the kick velocity is very well recovered in some moderate-SNR cases, we expect that our method may yield a successful kick measurement before design sensitivity is achieved. This would mark the first time a gravitational recoil is experimentally studied with GWs, providing a brand new observable for astrophysics.

Acknowledgments— We thank Juan Calderon Bustillo for a review and comments on the manuscript. We thank Nathan Johnson-McDaniel, Ajith Parameswaran, Davide Gerosa, Matt Giesler, Leo Stein, Saul Teukolsky, Gregorio Carullo, Aaron Zimmerman, and Bala Iyer for useful discussions. V.V. is supported by the Sherman Fairchild Foundation, and NSF grants PHY – 170212 and PHY – 1708213 at Caltech. M.I. is supported by NASA through the NASA Hubble Fellowship grant No. HST-HF2-51410.001-A awarded by the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., for NASA, under contract NAS5-26555. This research made use of data, software and/or web tools obtained from the Gravitational Wave Open Science Center [76], a service of the LIGO Laboratory, the LIGO Scientific Collaboration and the Virgo Collaboration. Computations were performed on the Alice cluster at ICTS and the Wheeler cluster at Caltech. This paper carries LIGO document number LIGO-P2000030.


See Supplemental Material here, for details about the Doppler shifted remnant mass, a study of biases in the kick measurement, and priors used for the component parameters. This further includes Refs. [86–99]. ([).


[83] The overall minimum should be zero, but a 35 km/s limit arises from numerical noise in the simulations on which NRSur7dq4Remnant is trained. In spite of this, NRSur7dq4Remnant is more noise than alternate kick models by an order-of-magnitude [50].


