New approach to semileptonic tags in B-meson semi-invisible decays

Gaetano de Marino,[1,*] Diego Guadagnoli,[2,†] Chan Beom Park,[3,4,‡] and Karim Trabelsi[1,§]

[1]*Université Paris-Saclay, CNRS/IN2P3, IJCLab, 91405 Orsay, France*
[2]*LAPTh, Université Savoie Mont-Blanc et CNRS, 74941 Annecy, France*
[3]*Center for Theoretical Physics of the Universe, Institute for Basic Science (IBS), 34126 Daejeon, Korea*
[4]*Department of Physics, Chonnam National University, Gwangju 61186, Korea*

(Received 18 September 2022; accepted 13 February 2023; published 9 March 2023)

Semi-invisible decays of mesons are a prime strategy for searches of new particles or forbidden decays. At current and upcoming facilities beauty mesons are produced in pairs, and the second decay often provides a "tag" to enhance signal reconstruction. Typical tags are fully reconstructed, hadronic decays, whereas one tends to discard "semileptonic" modes, clean and abundant, but with an elusive neutrino. We introduce a new strategy that makes semileptonic tags competitive with hadronic tags, potentially implying a substantial increase in the usable event statistics for *any* semi-invisible meson or τ-lepton decay. Our strategy rests on the use of appropriate kinematic quantities, collectively denoted as M_2, that had never been applied in the context of meson decays. We introduce M_2 definitions that leverage two key features of high-intensity meson-decay events: the known decaying-meson mass and its flight direction—accessible thanks to the exquisite vertexing capabilities available. We benchmark these definitions in the $B \to K\tau\mu$, a null test of the Standard Model of great current interest. In an unrefined application we already observe a substantial improvement in sensitivity—halfway to the ideal case of using fully reconstructed semileptonic tags with true kinematics.

DOI: 10.1103/PhysRevD.107.055010

Pair-produced particles decaying to partly invisible final states are a ubiquitous type of events at high-intensity experiments such as LHCb and Belle II. A prototype example are $B - \bar{B}$ pairs, each B decaying to final states including at least one neutrino. Of these B-meson pairs, one ($B_{\rm sig}$) produces the "signal" decay, and the other ($B_{\rm tag}$) may be used as a "tag." Common tag decays are modes that can be fully reconstructed, as is the case for many hadronic tags. This way, even if the signal includes elusive particles, the number of constraints is large enough to close the kinematics. Such strategy, somewhat by definition, excludes from consideration tag decays that contain undetected particles, for instance most semileptonic tags. This is unfortunate because semileptonic tags are often clean, thus affording high efficiencies, and most importantly they have large branching fractions—e.g., almost 20% for only four $D^{(*)}\ell\nu$ modes. In this study we introduce a strategy that achieves, with appropriately designed kinematic quantities, a full event reconstruction in semileptonic tags, thus making them competitive in resolution with hadronic tags, and thereby substantially increase the usable statistics in, potentially, *any* semi-invisible meson or τ-lepton decay.

Precisely events where both $B_{\rm sig}$ and $B_{\rm tag}$ decay semi-invisibly lend themselves to the use of kinematic variables fit for purpose. Consider B mesons pair-produced in electron-positron collisions at the $\Upsilon(4S)$ resonance, and decaying as

$$B_1 B_2 \to V_1(p_1)\chi_1(k_1) + V_2(p_2)\chi_2(k_2), \tag{1}$$

where V_i are visible and χ_i invisible (sets of) particles. The center-of-mass energy $\sqrt{s}$ is fixed, and the *full* momentum of the e^+e^- system is known. Therefore, besides the visible $p_{1,2}$, in such events one measures the *full* missing three-momentum $\boldsymbol{P}^{\rm miss}$ and can thereby construct the variable known as M_2 [1–3]

$$M_2 = \min_{\boldsymbol{k}_1+\boldsymbol{k}_2=\boldsymbol{P}^{\rm miss}} [\max\{M_{(1)}, M_{(2)}\}], \tag{2}$$

where $M_{(i)}$ is the total invariant mass on decay branch $i = 1, 2$, i.e., $M^2_{(i)} = (p_i + k_i)^2$. Note that these momenta correspond to those in Eq. (1) and that the $\boldsymbol{P}^{\rm miss}$ constraint reduces to *three* the variables over which the minimization

*demarino@ijclab.in2p3.fr
†diego.guadagnoli@lapth.cnrs.fr
‡cbpark@jnu.ac.kr
§karim.trabelsi@in2p3.fr

is performed, e.g., $(\boldsymbol{k}_1)_{x,y,z}$. The construction in Eq. (2) follows a well-defined logic [1]. The choice of the maximum between the invariant masses of the two decay branches is designed for M_2 to access, through its endpoint as a distribution, the heaviest physics scale of the decay, in our case the $B_{1,2}$ mass m_B. The subsequent minimization guarantees that $M_{(i)}$—calculated for $\boldsymbol{k}_i$ that are separately unknown, hence may take on unphysical values—do *not* result in unphysical parent-mass values, in particular ones that are larger than m_B.

This construction comes with built-in features that are key in our context. The first is the possibility to estimate the *separate* invisible momenta for the two decay chains. In fact, it was shown that the minimum of Eq. (2), obtained for *unique* values of $\boldsymbol{k}_{1,2}$—so-called M_2-assisted on-shell invisible momenta, or MAOS momenta [3–6]—is distributed around the *physical* $\boldsymbol{k}_{1,2}$. Second, the minimization in Eq. (2) can be performed in the presence of *additional* constraints, so the definition of M_2 is extensible. Crucially the larger their number, the more MAOS momenta approach the physical momenta [6].

We can use such 'extensibility' to our advantage. Specifically, we introduce two new kinematic requirements that can be enforced on the M_2 definition. These constraints are new in the following respects: (i) they "overload" M_2, i.e., they reduce to zero the number of kinematic degrees of freedom in the M_2 distribution. Then M_2 is no more a distribution—its minimization becomes equivalent to finding the unique solution of the event's kinematic equations; (ii) we consider kinematic requirements other than on-shell mass constraints—in particular, constraints on the parent-$B_{\rm sig}$ (or equivalently -$B_{\rm tag}$) flying *directions*, inferred from *vertex*-reconstruction information. A third novelty is the actual application, to namely B-meson decays with final-state τ leptons, greatly topical at present. We next discuss these three novelties in turn.

A first natural constraint at lepton colliders is the known total collision energy

$$(p_1 + k_1 + p_2 + k_2)^2 = s. \tag{3}$$

In our application to meson decays, a further constraint are the known decaying-parent masses

$$(p_1 + k_1)^2 = (p_2 + k_2)^2 = m_B^2. \tag{4}$$

Note that the constraints (3) and (4), taken together, reduce to *zero* the number of d.o.f. in the M_2 minimization—i.e., make the M_2 solution equivalent to finding the unique root of a complete set of kinematic equations for the event. We then introduce a first M_2 definition, denoted as M_{2sB}, and constructed from Eq. (2), with the minimization further constrained via Eqs. (3) and (4). In the literature, the constraint in Eq. (4) has been discussed *without* the last equality [7], i.e., without assuming a known mass for the parent particles. In this case M_2 is known to have the same minimum as M_{2s} [7]. To our knowledge (and surprise), a discussion of M_2 with the full constraint in Eq. (4) is missing in the literature. This may be due to various reasons. M_2 was born as a variable for direct searches of new—hence with unknown mass—pair-produced resonances decaying semi-invisibly; in fact, one of the defining features of M_2 is that its endpoint as a distribution allows us to measure the decaying-parent's mass. On the contrary, our M_{2sB} is no more a distribution—its minimum is an exact solver of the kinematic equations, event by event—as the constraints (4) plus (3) close the event kinematics.

We next consider a further, qualitatively different, class of constraints than those hitherto discussed in the literature —namely constraints on the decaying-parent flight *direction*. Such constraints may well be valuable because of the accurate vertex-reconstruction information available at present and upcoming high-intensity colliders. Note that imposing, e.g., the $B_{\rm sig}$ flight direction renders the $B_{\rm tag}$ counterpart redundant [8], so we will discuss the former only.

The $B_{\rm sig}$ flight direction is determined, event by event, as $\hat{\boldsymbol{v}}_{\rm sig} = (\boldsymbol{r}_{\rm sig} - \boldsymbol{r}_0)/|\boldsymbol{r}_{\rm sig} - \boldsymbol{r}_0|$, where $\boldsymbol{r}_0$ and $\boldsymbol{r}_{\rm sig}$ are the locations of the primary and respectively the $B_{\rm sig}$-decay vertices. Concretely, each component of $\boldsymbol{r}_0$ as well as $\boldsymbol{r}_{\rm sig}$ comes with an error, which implies that the measured $B_{\rm sig}$ flight vector spans a set of directions. If we denote $\hat{\boldsymbol{v}}_{\rm sig}$ as the vector obtained with the central values for $\boldsymbol{r}_0$ and for $\boldsymbol{r}_{\rm sig}$, we may impose the following constraint

$$\arccos(\hat{\boldsymbol{p}}_{B_{\rm sig}} \cdot \hat{\boldsymbol{v}}_{\rm sig}) \leq \delta_{\rm sig}, \tag{5}$$

where $\hat{\boldsymbol{p}}_{B_{\rm sig}}$ denotes the unit vector corresponding to $\boldsymbol{p}_{B_{\rm sig}} = \boldsymbol{p}_{\rm sig} + \boldsymbol{k}_{\rm sig}$, with $\boldsymbol{k}_{\rm sig}$ estimated through MAOS [4,5]. The quantity on the right-hand side (rhs), $\delta_{\rm sig}$, can be determined by scanning over $\hat{\boldsymbol{v}}_{\rm sig}$ as the measured $\boldsymbol{r}_0$ and $\boldsymbol{r}_{\rm sig}$ are varied within their standard deviations σ^{r_0} and $\sigma^{r_{\rm sig}}$. One can thus expect that $\delta_{\rm sig}$ will be of the same order as the largest between the relative uncertainties for $\boldsymbol{r}_0$ and $\boldsymbol{r}_{\rm sig}$. Reference values for these uncertainties will be discussed in the text around Fig. 1. Intuitively, the constraint in Eq. (5) would dictate that the $B_{\rm sig}$ direction form a cone of maximal aperture $2\delta_{\rm sig}$ with $\hat{\boldsymbol{v}}_{\rm sig}$. As the error on $\boldsymbol{r}_{0,\rm sig}$ tends to zero, so does $\delta_{\rm sig}$ [9]. Importantly, Eq. (5) amounts to an *inequality* constraint, whose application does *not* reduce the number of d.o.f. in the M_2 minimization. Inequality constraints may be very powerful for the purpose of overloading M_2, namely of bounding its minimization with a number of requirements that equal or exceed the number of kinematic d.o.f. available. We will return to this point later. In our circumstances, we can follow a simpler procedure. Event by event, we replace the true $\hat{\boldsymbol{v}}_{\rm sig}$ with a vector estimated by smearing with motivated distributions both $\boldsymbol{r}_0$ and $\boldsymbol{r}_{\rm sig}$ around their true values. We then impose the

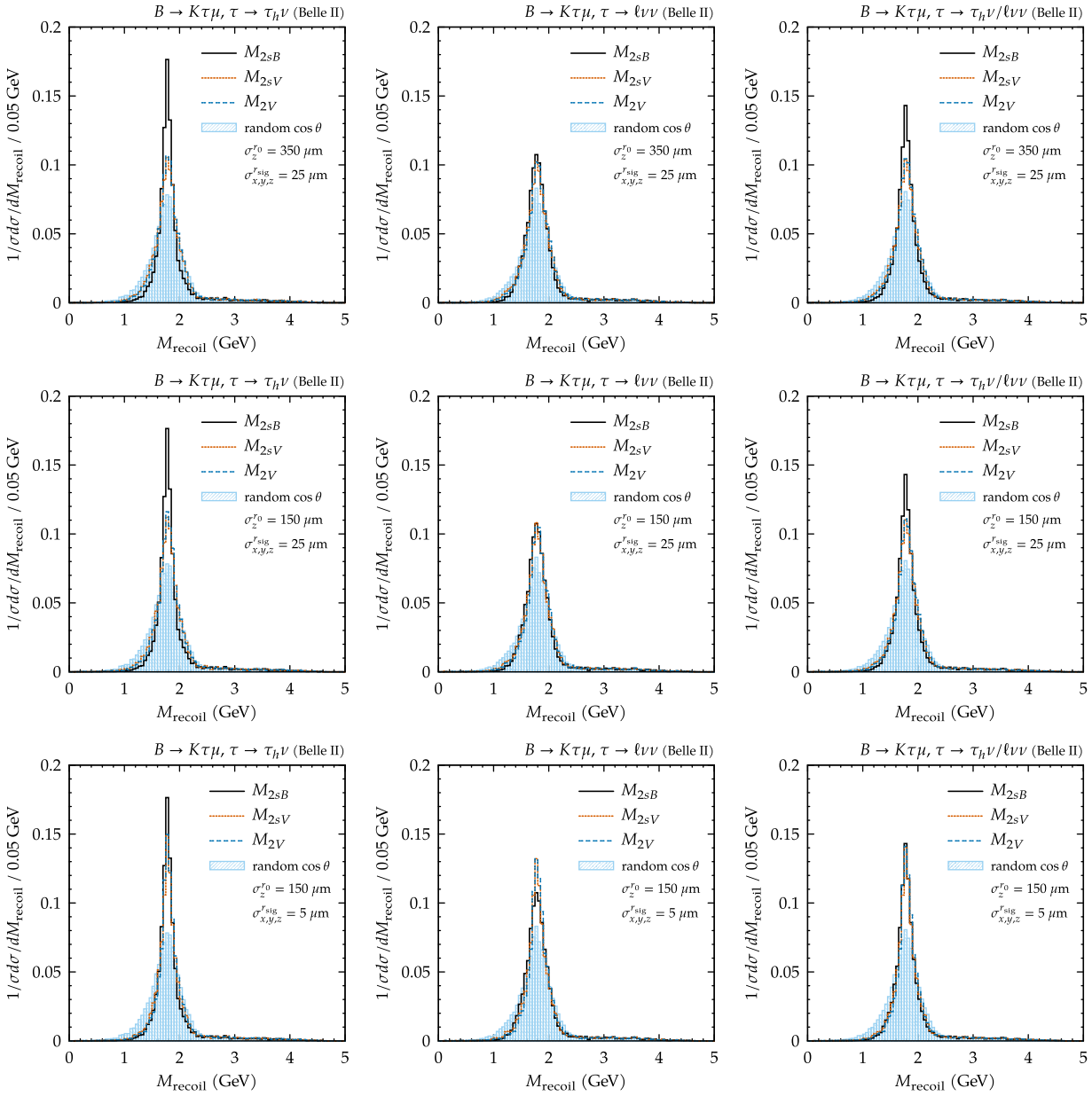

FIG. 1. $M_{\rm recoil}$ distributions for different experimental setups, detailed in the text, and for different $\tau_{\rm sig}$ decay modes (leftmost to rightmost panels: hadronic, leptonic τ decays, or both).

thus-estimated $\hat{\boldsymbol{v}}_{\rm sig}$ as an *equality* constraint (henceforth V) on the $B_{\rm sig}$ flight direction.

"Directional" constraints such as Eq. (5) have, to our knowledge, never been considered in connection with M_2 and siblings. The M_2 definition of Eq. (2), plus the equality V constraint just discussed will be referred to as M_{2V} [16]. M_{2V} reduces the number of d.o.f. available in the M_2 minimization by two units—the constraint fixes one direction in 3-dimensional space. Note that, if we further use the knowledge of s, we can fix not only the orientation, but also the magnitude of $\boldsymbol{p}_{B_{\rm sig}}$, and thereby close the kinematics. In this case, that we will denote as M_{2sV}, the minimization must necessarily land at $\boldsymbol{k}_{\rm sig} = \boldsymbol{k}_{\rm sig}^{\rm true}$. On the other hand, if one does *not* use the s information, as in M_{2V}, s is an outcome of the M_2 algorithm.

We next turn to the actual application, the $B^\pm \to K^\pm\tau\mu$ search at Belle and Belle II. This decay is a null test of the Standard Model, hence any signal would represent a clear signature of new physics coupled to the third generation of down-type quarks and leptons [17].

The search strategy in place at B factories (see, e.g., Refs. [18,19]) for $B^\pm \to K^\pm\tau\mu$ is based on $B^+_{\rm sig} \to K^+_{\rm sig}\tau\ell_{\rm sig}$ events, with 1-prong τ decays to $\ell\nu\nu$, $\pi\nu$ and $\rho\nu$ (making up over 70% of all τ decays), and associated with a *fully reconstructed, hadronic* tag decay, e.g., $B^-_{\rm tag} \to D^0(\to K^-\pi^+)\pi^-$ [20]. This $B_{\rm tag}$ decay is referred to as

hadronic B-tagging, and allows us to completely reconstruct the $B_{\rm sig}$ decay as well. There exists an entire set of additional, independent *semileptonic* (SL) tag decays such as $B^-_{\rm tag} \to D^0(K\pi)\ell^-\bar{\nu}$, which at present cannot be exploited within the above strategy. In fact, these decays involve escaping neutrinos, which hinder full event reconstruction. Our strategy is key precisely in these SL-tag decays, and makes them competitive to hadronic-tag ones, as we will show in terms of increase in statistics and of gain in branching-ratio sensitivity. Our chosen decay is meant to benchmark the strategy, and our results suggest it is exploitable in numerous other applications, some of which we mention at the end of the paper.

The overall resolution of a given set of $B_{\rm sig}B_{\rm tag}$ decay channels—whether the tag is hadronic or SL—may be quantified by $M_{\rm recoil}$. In our chosen decay, this has the crucial advantage of reducing the search to a "bump hunt" in the total invariant mass of the signal-side τ decay products. In fact, if correctly estimated, $M_{\rm recoil}$ must peak at m_τ. $M_{\rm recoil}$ can be constructed as

$$\begin{aligned} M^2_{\rm recoil} &\equiv (p^*_{e^+e^-} - p^*_{B_{\rm tag}} - p^*_{K_{\rm sig}\ell_{\rm sig}})^2 \\ &= m^2_{B_{\rm tag}} + m^2_{K_{\rm sig}\ell_{\rm sig}} \\ &\quad - 2(E^*_{B_{\rm tag}}E^*_{K_{\rm sig}\ell_{\rm sig}} + |\boldsymbol{p}^*_{B_{\rm tag}}||\boldsymbol{p}^*_{K_{\rm sig}\ell_{\rm sig}}|\cos\theta). \end{aligned} \quad (6)$$

Here asterisks denote the center-of-mass frame, where $M_{\rm recoil}$ takes a simple form, although it is Lorentz-invariant by definition; besides, $E^*_{B_{\rm tag}} = \sqrt{s}/2$, and θ is the angle between $\boldsymbol{p}^*_{B_{\rm tag}}$ and $\boldsymbol{p}^*_{K_{\rm sig}\ell_{\rm sig}}$. Equation (6) allows us to immediately identify the main current limitation of SL tags: since the tag side is not fully reconstructed, θ is unknown. For SL-tag analyses, the cosine of this angle is currently taken as uniformly distributed. We find that, for this reason, the current SL-tag resolution is about 5 times worse than the hadronic tag's [21].

Thus, our stated objective can be recast as the following task in our considered decay: improving the $\cos\theta$ estimate in SL tags, and quantifying the improvement in $M_{\rm recoil}$ as well as in the ensuing branching-ratio sensitivity. To that end, we adopt a numerical procedure whose technical details are relegated to a supplemental section.

A complete comparison of the performance of M_{2sB}, M_{2V}, and M_{2sV} in the $M_{\rm recoil}$ distribution is presented in the histogram array of Fig. 1. The random-$\cos\theta$ case is shown as baseline. To ease readability, we note that the array's columns refer to the different 1-prong decay modes considered for the signal τ: from left to right, hadronic, leptonic τ decays, or both. The choice of the decay modes affects all of M_{2sB}, M_{2V}, and M_{2sV}. The array's rows, in turn, represent different scenarios for the V constraint, defined by the uncertainties σ^{r_0} and $\sigma^{r_{\rm sig}}$ in the primary vertex and in the $B_{\rm sig}$ decay vertex, respectively. Upper to lower rows represent the current setup at Belle II; the setup expected at the Belle II design luminosity; a hypothetical future scenario [22], respectively. By definition, these different scenarios affect $M_{2(s)V}$ only, not M_{2sB}. The M_{2sB} distributions are shown in every figure row only for comparison with $M_{2(s)V}$.

This comparison shows that M_{2sB} performs better than $M_{2(s)V}$ in the hadronic $\tau_{\rm sig}$-decay case, whereas M_{2sB} and $M_{2(s)V}$ are comparable in the leptonic-decay instance. This implies a somewhat better M_{2sB} performance when the channels are combined. These conclusions hold in the "current Belle-II" scenario (first row) and to a lesser degree in the "Belle-II design-luminosity" scenario (second row). We see from the lower two rows that, if the $\sigma^{r_{\rm sig}}$ value were to halve with respect to the 25 μm figure, the $M_{2(s)V}$ performance would be very close to M_{2sB} in the hadronic $\tau_{\rm sig}$-decay case, and even superior to it in the leptonic-decay case, implying a comparable performance between M_{2sB} and $M_{2(s)V}$ in the combined-channel case.

These findings suggest that $M_{2s(B)}$ and $M_{2(s)V}$ have distinct advantages and disadvantages: the former has a strong sensitivity to the $m_{\nu\bar{\nu}}$ constraint; the latter has little sensitivity in that respect, and allows us to profitably use vertex-reconstruction information—to the extent that it is accurate enough. In fact, the choice of $M_{2(s)V}$ over M_{2sB} hinges on the considered detector's vertexing capabilities, and in case of comparable performances the best strategy would be a combined analysis, where M_2 is overloaded with all of the s, B, and V constraints.

The $M_{\rm recoil}$ distributions can finally be translated into an upper limit on $\mathcal{B}(B \to K\tau\mu)$. This is calculated at 90% confidence level with an established frequentist method (see, e.g., Ref. [24]). For each given $M_{\rm recoil}$ distribution, we apply our selection to a Monte Carlo sample consisting of all possible backgrounds, of overall size equal to the Belle dataset. This allows us to constrain the background shape in the signal region beyond a simple sideband extrapolation. Given the superiority of M_{2sB} over $M_{2(s)V}$ within the Belle-II setup in the foreseeable future, for this study we deploy M_{2sB} alone, with the $\hat{m}_{\nu\bar{\nu}}$ ansatz for $\tau_{\rm sig} \to \ell + \nu\bar{\nu}$. The analysis uses somewhat simplifying assumptions: 1-prong $\tau_{\rm sig}$ decays are reconstructed as $\pi/\mu/e$, i.e., the ρ is currently not being reconstructed; besides cross-feed across the different categories is neglected. These effects, however, are not expected to sizeably change the overall picture. With these simplifications, we get a 90% C.L. upper bound on $\mathcal{B}(B^\pm \to K^\pm\tau^\pm\mu^\mp) = 1.2 \times 10^{-5}$ with 710 fb^{-1}. Within our approximations, this limit equals the one that we obtain with the hadronic tag. The corresponding limits obtained with $M_{\rm recoil}$ (random $\cos\theta$) and with $M^{\rm true}_{\rm recoil}$ (using the true $\cos\theta$) are respectively 2.0 and 0.6 both in units of 10^{-5} (see also [21]). Hence a no-frills application of M_{2sB} leads *per se* to an improvement already halfway between the current strategy and the fully-reconstructed SL case.

Our approach opens several lines of development. First, the constraints discussed can be applied to M_2, Eq. (2), in a number of combinations, according not only to the constraints that are actually included, but also to whether they are imposed exactly (i.e., as equalities) or as inequality relations. In particular, one may exceed the total number of kinematic d.o.f. available—thus overloading M_2—by implementing constraints as inequalities. In this respect, V constraints are especially promising, also in that they can be deployed at hadronic facilities, where s is not available, but accurate vertex reconstruction typically is. Second, the $B \to K\tau\mu$ analysis we discussed may be carried over to many additional channels of topical interest. One example is $B \to \tau\mu$, that features at least two clear advantages: the τ is monochromatic in the parent-B rest frame, and the signal has lower combinatorial backgrounds. Another example is $B \to K\nu\bar{\nu}$, a key constraint for model-building [25] and even for searches of light supersymmetric states (e.g., [26]). We expect our approach to make SL tags as practicable as hadronic tags in these and other modes. Our work is in the direction of what may be denoted as the "efficiency frontier." The next step toward improving tag-based analyses even further is then to enlarge the set of hadronic tags used—in what is instead the "resolution frontier." This step seems however very dependent on the specific decay considered, i.e., it does not seem to admit a universal approach such as the one laid out here.

ACKNOWLEDGMENTS

This work is supported by ANR under Contract No. 202650 and by IBS under the project code IBS-R018-D1.

APPENDIX: TECHNICAL DETAILS

Here we collect information that is not necessary for the general understanding of our approach, but is needed for fully reproducing our results.

The first point concerns our numerical setup and assumptions. Our results use phase-space events populated through the EvtGen Monte Carlo [27]. For consistency with the on-going Belle analysis [18], and as already mentioned, we restrict to the 1-prong τ decays to $\ell\nu\nu$, $\pi\nu$ and $\rho\nu$. We note that our thus-generated event sample includes realistic detector smearing, which in our case affects $\boldsymbol{p}_i$ [see Eq. (1)], $\boldsymbol{P}^{\text{miss}}$, and the locations of the interaction point and of the decay vertices. We also assume that the combinatorial ambiguity due to the assignment of the visible particles to one of the two decay chains has completely been resolved. The purpose is to single out the signal quality degradation due *solely* to the presence of undetected particles on the tag side.

The second point concerns the actual construction of one of the inputs to M_2, Eq. (2) The constraints (3) and (4) require knowledge of $k^2_{1,2}$, the invariant masses of the invisible systems on the signal and tag sides. While $k^2_2 = 0$, k^2_1 is unknown for a leptonically decaying signal-side τ—it is the invariant mass squared of the two final-state neutrinos. A simple ansatz often adopted in the literature is $k^2_1 = 0$, see in particular [4,5]. This is however not realistic, as the physical $m^2_{\nu\bar{\nu}}$ distribution peaks around 1 GeV^2, as shown in Fig. 2 (left). We then consider an improved ansatz for k^2_1, expecting that it will make more effectual the M_2 constraints where k^2_1 enters. The M_2 definition proper is, on the other hand, expected to be less affected by the k^2_1 ansatz [4,5,28]. We construct our k^2_1 ansatz as follows. We start from the B in its rest frame, and boost it with the total beams momentum. We then obtain k_1 by subtracting the sum of the visible final-state momenta. This approximation neglects the back-to-back momentum of the $B - \bar{B}$ pair in the center-of-mass frame, which is however small with respect to the boost induced by the beams asymmetry. As a result of our ansatz, the k^2_1 distribution is neatly close to the physical one for about 86% of the events, and yields unphysical negative values for the remaining 14%. For these events we switch the overall k^2_1 sign. The resulting ansatz for the $\nu\bar{\nu}$ invariant mass, to be denoted as $\hat{m}_{\nu\bar{\nu}}$, gives rise to the distribution shown in Fig. 2.

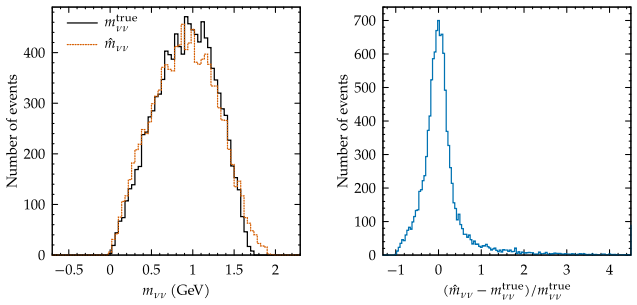

FIG. 2. Distribution for $\hat{m}_{\nu\nu}$ as defined in the text, and comparisons with the physical distribution.

We implemented $\hat{m}_{\nu\bar{\nu}}$ in all of $M_{2s(B)}$ and $M_{2(s)V}$ and tested the improvement directly in the M_{recoil} distribution. In Fig. 3 we show these four M_2 definitions, plus the random-cos θ distribution as reference. The three panels correspond to the three choices $m_{\nu\bar{\nu}} = \{m^{\text{true}}_{\nu\bar{\nu}}, \hat{m}_{\nu\bar{\nu}}, 0\}$, respectively. As the figure shows in terms of the distributions' height/width, the improvement due to $\hat{m}_{\nu\bar{\nu}}$ with respect to $m_{\nu\bar{\nu}} = 0$ is significant for the $M_{2(s)V}$, whereas it is only slight for M_{2sB}. However, for M_{2sB} the $\hat{m}_{\nu\bar{\nu}}$ ansatz has the other effect of removing a bias present in the peak position with the $m_{\nu\bar{\nu}} = 0$ ansatz. All things considered, the accuracy of the $m_{\nu\bar{\nu}}$ ansatz appears to impact especially M_{2sB}. This sensitivity seems to be due to the B constraint. In fact, the quality of the peak in M_{recoil} calculated with M_{2sB} improves dramatically from the $\hat{m}_{\nu\bar{\nu}}$ to the $m^{\text{true}}_{\nu\bar{\nu}}$ cases, whereas the corresponding improvement in M_{2s} or $M_{2(s)V}$ is marginal or absent.

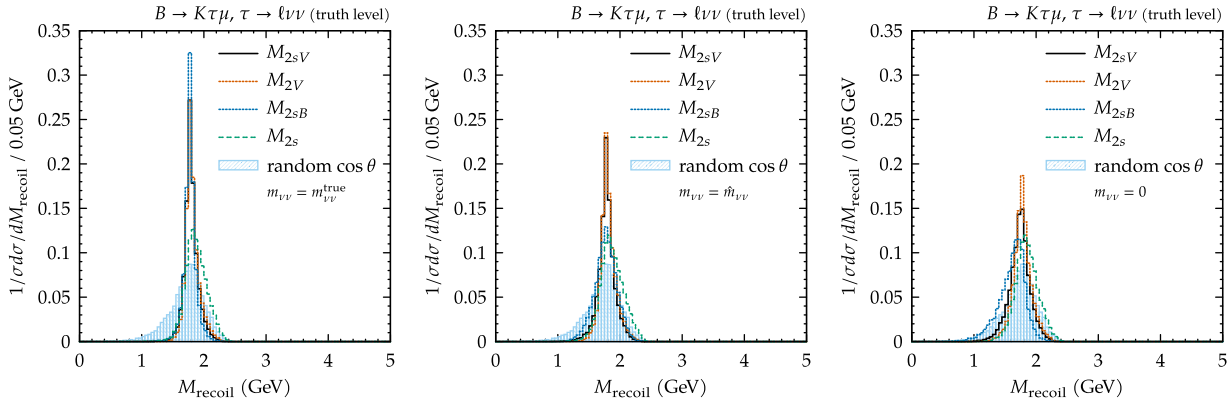

FIG. 3. M_{recoil} distributions for $M_{2s(B)}$ and $M_{2(s)V}$ (see lines in each plot's legend), as a function of $m_{\nu\bar{\nu}} = \{m^{\text{true}}_{\nu\bar{\nu}}, \hat{m}_{\nu\bar{\nu}}, 0\}$ (panels left to right). The random-cos θ distribution is also shown as reference.

[1] A. J. Barr, T. J. Khoo, P. Konar, K. Kong, C. G. Lester, K. T. Matchev, and M. Park, Guide to transverse projections and mass-constraining variables, Phys. Rev. D **84**, 095031 (2011).

[2] G. G. Ross and M. Serna, Mass determination of new states at hadron colliders, Phys. Lett. B **665**, 212 (2008).

[3] W. S. Cho, J. S. Gainer, D. Kim, K. T. Matchev, F. Moortgat, L. Pape, and M. Park, On-shell constrained M_2 variables with applications to mass measurements and topology disambiguation, J. High Energy Phys. 08 (2014) 070.

[4] W. S. Cho, K. Choi, Y. G. Kim, and C. B. Park, M(T2)-assisted on-shell reconstruction of missing momenta and its application to spin measurement at the LHC, Phys. Rev. D **79**, 031701 (2009).

[5] C. B. Park, Reconstructing the heavy resonance at hadron colliders, Phys. Rev. D **84**, 096001 (2011).

[6] D. Kim, K. T. Matchev, F. Moortgat, and L. Pape, Testing invisible momentum ansatze in missing energy events at the LHC, J. High Energy Phys. 08 (2017) 102.

[7] P. Konar and A. K. Swain, Mass reconstruction with M_2 under constraint in semi-invisible production at a hadron collider, Phys. Rev. D **93**, 015021 (2016).

[8] In the center-of-mass (CM) frame, $\boldsymbol{p}_{B_{\text{tag}}} = -\boldsymbol{p}_{B_{\text{sig}}}$ and that, at Belle (II), the boost from the lab to the CM frame is known from $\boldsymbol{p}_{\Upsilon(4S)}$. As a consequence, imposing the constraint (5) also on the tag side is redundant.

[9] The inequality constraint in Eq. (5) may be naturally implemented using the sequential quadratic programming (SQP) method [10–14], which is the main algorithm in the YAM2 software library [15] utilized throughout this work.

[10] R. B. Wilson, A simplicial algorithm for concave programming, Ph.D. thesis, Graduate School of Business Administration, Harvard University, 1963.

[11] U. M. G. Palomares and O. L. Mangasarian, Superlinearly convergent quasi-Newton algorithms for nonlinearly constrained optimization problems, Math. Program. **11**, 1 (1976).

[12] S.-P. Han, Superlinearly convergent variable metric algorithms for general nonlinear programming problems, Math. Program. **11**, 263 (1976).

[13] S. P. Han, A globally convergent method for nonlinear programming, J. Optimiz. Theory App. **22**, 297 (1977).

[14] M. J. D. Powell, Algorithms for nonlinear constraints that use lagrangian functions, Math. Program. **14**, 224 (1978).

[15] C. B. Park, YAM2: Yet another library for the M_2 variables using sequential quadratic programming, Comput. Phys. Commun. **264**, 107967 (2021).

[16] We verified that M_2 plus the inequality constraint of Eq. (5) yields, in the small-δ_{sig} limit, the same solution as the respective equality constraint. Imposing an equality constraint is however faster and less subject to numerical instabilities.

[17] S. L. Glashow, D. Guadagnoli, and K. Lane, Lepton Flavor Violation in B Decays?, Phys. Rev. Lett. **114**, 091801 (2015).

[18] S. Watanuki *et al.* (Belle Collaboration), Search for the lepton flavour violating decays $B^+ \to K^+\tau^\pm\ell^\mp$ ($\ell = e, \mu$) at Belle, arXiv:2212.04128.

[19] B. Aubert *et al.* (*BABAR* Collaboration), Search for the Decay $B^+ \to K^+\tau^\mp\mu^\pm$, Phys. Rev. Lett. **99**, 201801 (2007); J. P. Lees *et al.* (*BABAR* Collaboration), A search for the decay modes $B^{+-} \to h^{+-}\tau^{+-}l$, Phys. Rev. D **86**, 012004 (2012).

[20] Here we specified only correlated charge assignments. The search includes the case of flipped signs.

[21] G. de Marino, Search for the Lepton Flavor Violating B+ → K+ τ (e, μ) decays with the Belle and Belle II detectors, thesis, Université Paris-Saclay, 2022, https://theses.hal.science/tel-03945632.

[22] We note that the beam has a non-negligible size mostly in the e^+-beam axis z, hence the primary-vertex' components in the plane orthogonal to that axis are determined very accurately. At Belle $\sigma_z^{r_0} \sim 4$ mm, about one o.o.m. larger than the B decay length—which makes the $\hat{\boldsymbol{v}}_{\text{sig}}$ constraint,

Eq. (5), ineffectual. Conversely, $\sigma_z^{r_0} \simeq 350$ μm at Belle II, and expected to further improve to 150 μm at the design luminosity [23]. The secondary vertex $\boldsymbol{r}_{\rm sig}$ is determined with a spread in each coordinate of about 45 μm at Belle and 25 μm at Belle II. We estimate $\boldsymbol{r}_{0,\rm sig}$ as normally distributed random numbers, with standard deviations $\sigma^{r_0,r_{\rm sig}}$.

[23] T. Abe *et al.* (Belle-II Collaboration), Belle II technical design report, arXiv:1011.0352.

[24] S. Choudhury *et al.* (Belle Collaboration), Test of lepton flavor universality and search for lepton flavor violation in $B \to K\ell\ell$ decays, J. High Energy Phys. 03 (2021) 105.

[25] A. J. Buras, J. Girrbach-Noe, C. Niehoff, and D. M. Straub, $B \to K^{(*)}\nu\bar{\nu}$ decays in the standard model and beyond, J. High Energy Phys. 02 (**2015**) 184.

[26] C. O. Dib, J. C. Helo, V. E. Lyubovitskij, N. A. Neill, A. Soffer, and Z. S. Wang, Probing R-parity violation in B-meson decays to a baryon and a light neutralino, arXiv:2208.06421.

[27] D. J. Lange, The EvtGen particle decay simulation package, Nucl. Instrum. Methods Phys. Res., Sect. A **462**, 152 (2001).

[28] D. Guadagnoli, C. B. Park, and F. Tenchini, $\tau \to \ell +$ invisible through invisible-savvy collider variables, Phys. Lett. B **822**, 136701 (2021).