

Vertex Fitting in the Belle II Analysis Framework

Francesco Tenchini^{*†}

University of Melbourne

E-mail: francesco.tenchini@unimelb.edu.au

The Belle II experiment at SuperKEKB is rapidly approaching its data taking phase, where it will be used to perform precision measurements of the Standard Model. Accurate vertex reconstruction is an essential tool for precise measurements of particle observables and for background suppression in such an experiment. We discuss the available vertex fitting methods in the Belle II Analysis Software Framework and their performance, as well as introduce a Kalman-based global fitting approach aimed at improving resolutions for underconstrained decay trees.

*The 25th International workshop on vertex detectors
September 26-30, 2016
La Biodola, Isola d'Elba, ITALY*

^{*}Speaker.

[†]On behalf of the Belle II collaboration.

1. SuperKEKB and Belle II

The Belle II experiment at KEK [1] has recently begun commissioning and is scheduled to start its first physics run in 2018, with the aim of performing precision measurements of Standard Model processes in the search for new physics over an intended sample of 50 ab^{-1} , fifty times the original Belle dataset. To this end the original KEKB accelerator, an asymmetrical electron-positron collider operating at energies of 8.0 and 3.5 GeV respectively, has been redesigned for a $\times 40$ increase in luminosity thanks to the increased current and to the improved beam focusing. Named SuperKEKB, the new accelerator circulated its first beams in 2016, at energies of 7.0 (e^-) and 4.0 GeV (e^+). Such an upgrade not only provides an increased collision rate, but also an increase in beam background rates. This must be dealt with accordingly through an upgrade of the detector, readout, trigger and reconstruction techniques. The inner tracker has been significantly modified to accommodate for the new beam characteristics. The original Belle vertex detector, composed of 4 layers of double-sided silicon detectors (DSSDs) has been replaced by a new one, composed of a 2-layered Pixel Detector (PXD) and a 4-layered Silicon Vertex Detector (SVD) with an inner radius reduced from 2.0 to 1.4 cm. This new design is aimed at reducing occupancy while simultaneously tracking charged particles closer to the interaction point.

2. Vertex Fitting in Belle II

Vertex reconstruction in B-factories aims at extracting the parameters of composite particles — momentum, invariant mass, decay vertex position — from the final state particles that constitute its daughters. It is an essential tool for background separation in physics analysis, in particular in the case of high combinatorics. A typical vertex fit in Belle II begins with using reconstructed tracks to determine the vertex coordinates, while the decaying particle kinematics are constrained through the requirement of four-momentum conservation. Neutral particles, such as photons reconstructed in the calorimeter, can be added to the fit but impose less stringent conditions on the vertex; no additional position information is provided although they fully contribute to the kinematics. Additional constraints can be subsequently applied to improve the measurement, such as setting the mass of one or more particles to the nominal value, or requiring the track to originate from the interaction region in the case of fits with a single track. One of the fundamental applications of vertex fitting is the extraction of the decay time difference of a B meson pair originating from the decay of a $\Upsilon(4S)$. This measurement is performed by boosting the rest frame of the decay and measuring the spatial separation in the \hat{z} direction of the lab frame; this is possible through the use of an asymmetric beam. Because of the new beam energies at which SuperKEKB operates, the Lorentz boost factor of Belle II ($\beta\gamma = 0.28$) is lower than that of Belle ($\beta\gamma = 0.425$). Therefore it is even more important to provide improved vertex resolutions, in particular along the beam \hat{z} axis.

Two implementations of vertex fitting are available at Belle II: *KFit* [2], originally developed for Belle and based on a least square minimisation approach; and *RAVE* [3], a new package introduced as a standalone implementation of CMS libraries and relying on a Kalman filter approach [6]. A third algorithm, *TreeFitter* is being developed in order to globally refit decay chains, and will be discussed in Sec. 4.1.

3. Vertex Fits Performance

In order to gauge the achievable vertex reconstruction performance, we consider a sample of $5 \times 10^5 B^0 \bar{B}^0$ pairs generated with GEANT4 [4] as part of the Belle II Monte Carlo 2015 campaign, one of which (the *signal*) decays in the “golden” mode $B^0 \rightarrow J/\psi(\rightarrow \mu^+ \mu^-) K_s^0(\rightarrow \pi^+ \pi^-)$ while the other (the *tag*) decays generically. Final state charged particles are selected from good tracks (track fit p-value > 0.001) with loose particle identification requirements, then fitted with RAVE to extract parameters for $J/\psi \rightarrow \mu^+ \mu^-$, $K_s^0 \rightarrow \pi^+ \pi^-$ and finally $B^0 \rightarrow J/\psi K_s^0$. The resulting z resolution for fitted J/ψ is shown in Fig. 1 (left). In this case the J/ψ is a short-lived resonance and decays inside the interaction region, close to the B^0 vertex. Therefore $\approx 95\%$ of its daughter muons cross the inner detector, producing hits in the PXD (Fig. 2) Rejecting those events can further improve the resolution, down to $\sigma_z = 19.8 \mu\text{m}$ with a $1.9 \mu\text{m}$ bias (Fig. 1, right). Compared to similar results in Belle ($\sigma_z^{Belle} = 43 \mu\text{m}$, $0.2 \mu\text{m}$), this is a significant improvement; the increased bias is to be attributed to as yet unresolved track fitting issues.

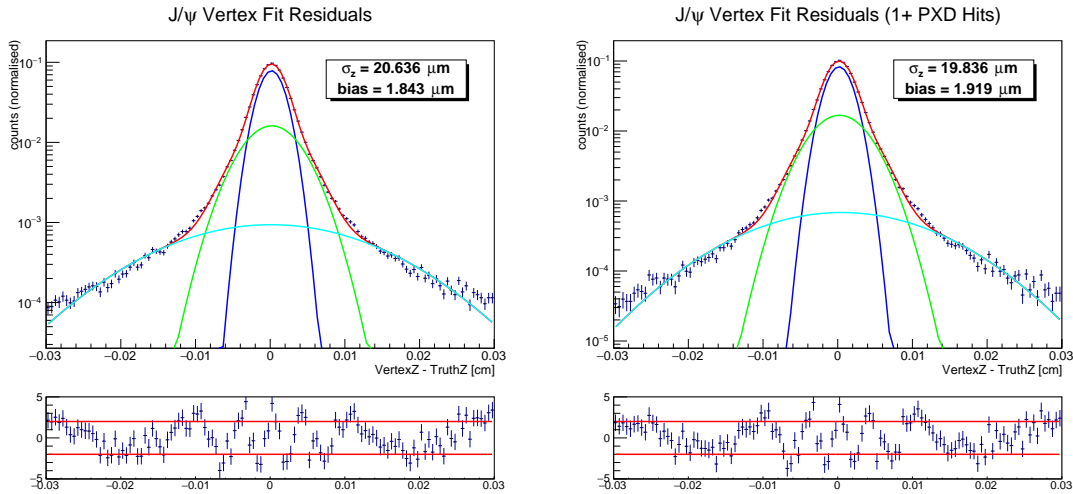


Figure 1: Longitudinal position resolution for reconstructed $J/\psi \rightarrow \mu^+ \mu^-$ vertices (*left*) and with the additional requirement of 1+ PXD hits per track (*right*).

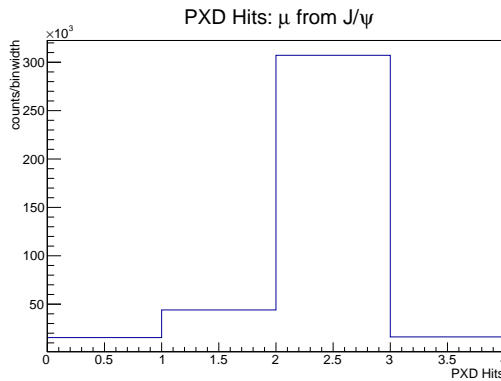


Figure 2: PXD hit count distribution of muons from reconstructed $J/\psi \rightarrow \mu^+ \mu^-$ decays.

The case of long-lived particles must be considered separately, as they usually decay away the interaction region. As an example, K_S^0 from the same process decay on average ≈ 6 cm away from the interaction point, well outside the PXD radius, and 77% of its daughter pions do not produce PXD hits (Fig. 3). Therefore when comparing vertices with and without the requirement of at least one PXD hit, the difference in resolution is much larger (Fig. 4).

These tracks from displaced vertices traverse less detector material than those originating within the beam pipe. In order to avoid performing material extrapolation during the analysis phase, long lived neutral particle are fitted with RAVE as part of the reconstruction process, after the fitting of charged tracks.

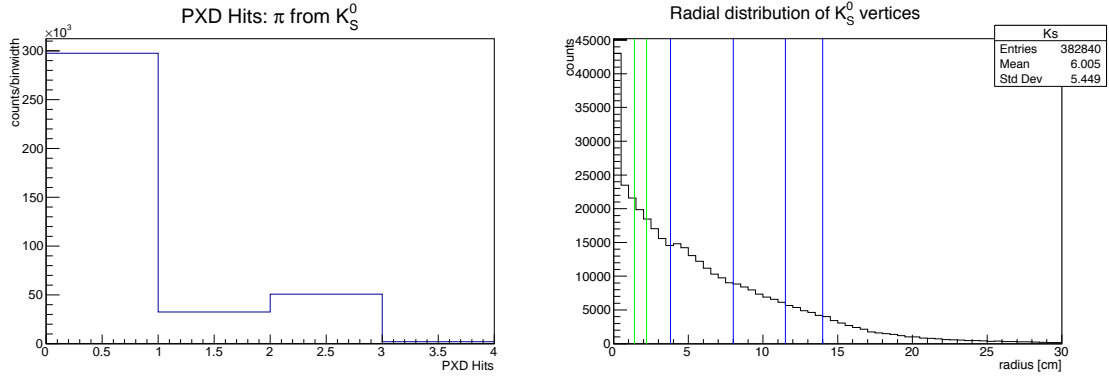


Figure 3: PXD hit count of daughter pions (*left*) and radial distribution of decay vertices (*right*) for reconstructed $K_S^0 \rightarrow \pi^+ \pi^-$. The colored lines mark the radius of the PXD (*green*) and SVD (*blue*).

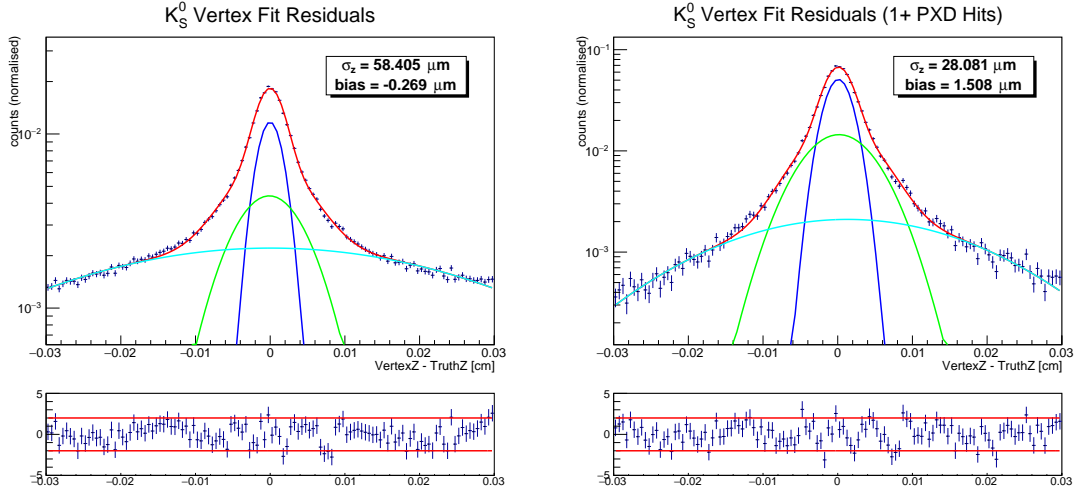


Figure 4: Longitudinal position resolution for reconstructed $K_S^0 \rightarrow \pi^+ \pi^-$ displaced vertices (*left*) and with the additional requirement of 1+ PXD hits per track (*right*).

When determining the B decay time difference for the purpose of time-dependent CP violation studies, the dominant contribution originates from the position of the generically-decaying B vertex. This is reconstructed by fitting (with RAVE) all the tracks with at least 1 PXD hit that do not belong to a displaced vertex nor to the fully reconstructed signal-side B. The resolutions are

similarly improved, going from the $\sigma_{tag,z} = 89 \mu\text{m}$ ($\sigma_{\Delta t} = 0.70$ ps) of Belle to $\sigma_{tag,z} = 52.7 \mu\text{m}$ ($\sigma_{\Delta t} = 0.63$ ps) in Belle II. Good time resolution is preserved, even with a reduced Lorentz boost.

4. Global Fitting

The example in the previous section describes the most commonly used approach to fully reconstructing a decay chain. Vertices are individually fitted, starting from final particle states up to the mother particle, usually a B or D meson or a τ . This is called *leaf-by-leaf* fitting and has the benefit of being simple and computationally light; however it is usually not optimal. Since each vertex in the chain is fit independently, the fit is blind to information which could contribute to the fit but is located upstream in the decay chain. As an example, in the decay $B^0 \rightarrow J/\psi K_s^0$, $K_s^0 \rightarrow \pi^0 \pi^0$, we have no knowledge of the (well-resolved) $J/\psi \rightarrow \mu^+ \mu^-$ vertex while fitting the K_s^0 . This could be used to better constrain the K_s^0 mass.

An alternate approach, called a *global fit* consists in parametrising and fitting the full decay tree. This method is computationally expensive; the execution of the fit requires the inversion of a large matrix due to the many parameters involved. It can however circumvent the limitations of a leaf-by-leaf fit thanks to having access to the measurements from the whole tree. This includes the full decay covariance matrix, which can be extracted as a natural output of the fit.

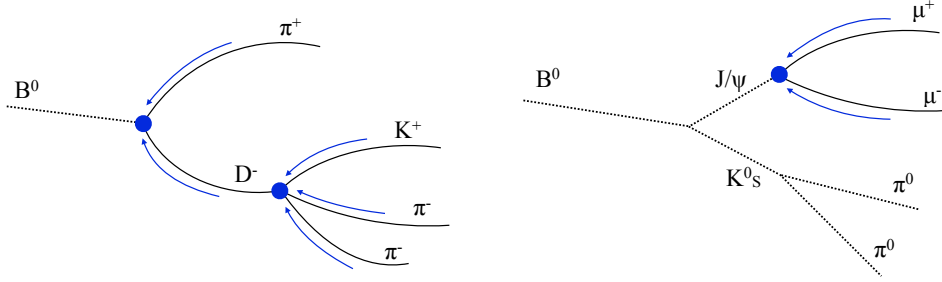


Figure 5: *Left:* Leaf-by-leaf vertex fitting approach. Vertices with neutral particles (dashed lines) and charged tracks (continuous curves) are fit individually. *Right:* Decay requiring global treatment.

4.1 TreeFitter

A solution to this problem was discussed in [5] and first used in BaBar. It is being implemented in the Belle II Analysis Software Framework (basf2) as the *TreeFitter* package. Let us consider a parametrisation x of the tree and a set of constraints $\{m_0, m_1, \dots, m_k, \dots\}$ which can be written as a function of the parameters, $m_k = h_k(x)$. Constraints are applied one by one using a Kalman filter approach with the gain formalism [6]. After $k - 1$ steps, applying the constraint m_k gives a new estimate for the parameters

$$x_k = x_{k-1} + K_k r_k^{k-1} \quad (4.1)$$

where $r_k^{k-1} = m_k - h(x_{k-1})$ is the *residual* of the k -th prediction, while the gain matrix is written as

$$K_k = C_{k-1} H_k^T (V_k + H_k C_{k-1} H_k^T)^{-1} \quad (4.2)$$

where C_{k-1} is the covariance matrix of the tree estimated at the $(k-1)$ -th step, $H_k = \partial h / \partial x|_{x_{k-1}}$ and V_k is the covariance of the constraint. The full derivation can be found in the cited paper. In this form, the matrix to be inverted is of the dimension of the constraint, rather than the dimension of the parametrisation, which significantly speeds up execution times. This is valid under the assumption of the constraint being linear. In the case of significantly nonlinear constraints, the procedure must be iterated to achieve convergence. $\mathcal{O}(1)$ iterations are generally sufficient. As the procedure is blind to higher order contribution, the final result may depend on constraint ordering. Carefully arranging constraints in order of increasing non-linearity will alleviate the problem. For further improvement, we must alter the traditional Kalman approach by introducing referencing. In this variant of the algorithm, we compute the projection matrix H_k using the result of the last full iteration, rather than a prediction x_{k-1} from the previous constraints. This limits the impact of ordering, improves fit stability and results in a faster convergence.

4.2 Physics Applications

The global approach described above can be used in a wide variety of physics analyses, but is uniquely suited to treating modes where parts of the tree would be otherwise underconstrained, such as time-dependent CP violation measurements in which a B^0 decays into neutral particles [7, 8, 9]. Missing particles —such as neutrinos— can be incorporated into the fit provided the decay chain is sufficiently constrained to handle the extra degrees of freedom. As an example we mention the semileptonic modes $B \rightarrow D^* l \nu$, where the ability to treat missing neutrino(s) could be used to improve the rejection of combinatorial background [10, 11, 12, 13]. Here the $B \rightarrow D^* \tau \nu$ sub-mode displays a 4σ excess that could hint at the existence of new physics. Additionally, as previously mentioned, the full covariance matrix of the decay can be directly extracted as part of the global approach. This is of particular use in the treatment of uncertainties in studies such as Dalitz plot analyses.

It is worth mentioning that analysis of Belle data within the Belle II framework has recently become possible thanks to the introduction of the the Belle to Belle II conversion module. Therefore development of global algorithms can also impact studies in Belle. A version of the *TreeFitter* has been implemented in basf2 and is currently being tested for performance.

5. Algorithm Timing

Physics analyses over large data sets imply a large number of repetitions of the vertex fit. As a consequence, the computational speed of the algorithms on a single event becomes an important point of optimisation. The timing performance of basf2 vertex fit algorithms was evaluated in the latest release, using once again the $B^0 \rightarrow J/\psi(\rightarrow \mu^+ \mu^-) K_s^0(\rightarrow \pi^+ \pi^-)$ decay. The average execution time per fit is shown in Tab. 1. The timing performance of RAVE is affected by its implementation, which requires interfacing to external libraries. For basic kinematic fits the use of kFit is therefore recommended. TreeFitter speed is comparable to kFit when fitting single vertices, but scales up to RAVE levels when performing a global fit. Performance on more complex modes is non-trivial to extrapolate a priori and is analysis dependent.

It is interesting to estimate, in the case above, the impact of the Kalman approach to the global fit as opposed to the inversion of a single large matrix. This difference is heavily dependent on the

Decay vertex	kFit	RAVE	TreeFitter
$J/\psi \rightarrow \mu^+ \mu^-$	0.9	22.8	1.6
$K_s^0 \rightarrow \pi^+ \pi^-$	2.2	40.0	-
$B^0 \rightarrow J/\psi K_s^0$	0.6	13.4	-
Total	3.7	76.2	71.8

Table 1: Execution time of vertex fit algorithms in HEPspec2006-ms/event[14].

way matrix inversion is performed; we find that it can vary between one order of magnitude (for an optimised algorithm) and two (for a naive $\mathcal{O}(n^3)$ inversion).

6. Conclusions

Vertex fitting serves a fundamental role in improving decay parameter measurements and improving background rejection. We discussed its use and application within the scope of the Belle II experiment and described the algorithms used for this purpose, as well as their performance in resolving decay vertex positions, where the role of the upgraded Silicon Pixel Detector is salient.

We presented a global algorithm, currently under implementation and aimed at refitting full decay chains. Its physics implications were discussed and computational performance was compared to existing approaches.

References

- [1] T. Abe *et al.* [Belle-II Collaboration], arXiv:1011.0352 [physics.ins-det].
- [2] J. Tanaka, Belle Note 194.
- [3] W. Waltenberger, IEEE Trans. Nucl. Sci. **58** (2011) 434.
- [4] S. Agostinelli *et al.* [GEANT4 Collaboration], Nucl. Instrum. Meth. A **506** (2003) 250.
- [5] W. D. Hulsbergen, Nucl. Instrum. Meth. A **552** (2005) 566.
- [6] R. Fruhwirth, Nucl. Instrum. Meth. A **262** (1987) 444.
- [7] B. Aubert *et al.* [BaBar Collaboration], Phys. Rev. Lett. **93** (2004) 131805.
- [8] B. Aubert *et al.* [BaBar Collaboration], Phys. Rev. Lett. **95** (2005) 011801.
- [9] K. Abe *et al.* [Belle Collaboration], arXiv:0708.1845 [hep-ex].
- [10] J. P. Lees *et al.* [BaBar Collaboration], Phys. Rev. D **88** (2013) no.7, 072012.
- [11] M. Huschle *et al.* [Belle Collaboration], Phys. Rev. D **92** (2015) no.7, 072014.
- [12] R. Aaij *et al.* [LHCb Collaboration], Phys. Rev. Lett. **115** (2015) no.11, 111803
Addendum: [Phys. Rev. Lett. **115** (2015) no.15, 159901].
- [13] The Heavy Flavor Averaging Group, <http://www.slac.stanford.edu/xorg/hfag/>, 2016.
- [14] HEP-SPEC06 Benchmark Suite, w3.hepik.org/benchmarks/doku.php