Search for a $\tau^+\tau^-$ resonance in $e^+e^- \to \mu^+\mu^-\tau^+\tau^-$ events with the Belle II experiment

Belle II Collaboration

Numerical results

We provide a text file with numerical results of the observed cross section of $e^+e^- \to X\,(\to \tau^+\tau^-)\;\mu^+\mu^-$, where $X = Z', S, \mathrm{ALP}$, as well as of the observed 90% CL upper limit on the cross section, g', ξ, and $|C_{\ell\ell}|/\Lambda$ with $\ell = e, \mu, \tau$ as functions of the mass.

Discriminant variables

Discriminating variables used as inputs of the MLP neural networks can be grouped in three classes: variables sensitive to the presence of a resonance in the final state; variables sensitive to the production mechanism, since the resonance is emitted as FSR off one of the two tagging muons; and variables sensitive to the presence of a $\tau^+\tau^-$ pair in the final state.

The first two classes contain variables expressed in the c.m. frame and mostly related to the kinematic properties of the two tagging muons.

Variables belonging to the third class are expressed in the reference frame where the recoil system against the two tagging muons is at rest (the Z' rest frame, in case of signal).

Variables sensitive to the presence of a resonance in the final state.

- The momenta of the two tagging muons are combined in two different variables, A and L. The two-dimensional distribution of the magnitudes of the two tagging muons is shown in Fig. S1, for signal and background: the distribution is confined within a straight line and a hyperbola, both analytically deducible from kinematic properties. Background processes populate the edges of the distribution, while signal is more uniformly distributed. To exploit this feature, an asymmetry-like discriminant variable A, shown in Fig. S2, is defined as $A = (d_1 - d_2)/(d_1 + d_2)$ where d_1 and d_2 are shown in Fig. S1. The second discriminant variable L is related to the position of the generic point of the distribution along the straight line (see the segment ℓ in Fig. S1), and is defined as ℓ scaled by the maximum recoil momentum kinematically reachable: background events cluster on the extremes of the distribution, while signal preferentially populates the central part.
- The invariant mass of the two additional charged particles (other than the two tagging muons), and the sum of the magnitudes of their momenta.

Variables sensitive to the production mechanism

The components of the recoil momentum transverse to the momentum direction of the tagging muon with maximum and minimum momentum, respectively, called $p_{T,\mathrm{recoil}}^{\mu-\mathrm{max}}$ and $p_{T,\mathrm{recoil}}^{\mu-\mathrm{min}}$ [1]. In the case of signal, these are the transverse momenta of the Z' with respect to the momentum direction of each of the two tagging muons. Since the Z' is radiated off one of the two muons, these variables are sensitive to the signal FSR production. The $p_{T,\mathrm{recoil}}^{\mu-\mathrm{max}}$ and $p_{T,\mathrm{recoil}}^{\mu-\mathrm{min}}$ variables are combined to form two different variables that use polar coordinates in the $p_{T,\mathrm{recoil}}^{\mu-\mathrm{max}}$ vs $p_{T,\mathrm{recoil}}^{\mu-\mathrm{min}}$ plane: the quadratic sum scaled by the maximum recoil momentum kinematically reachable R_T, shown in Fig. S3, and the polar angle.

Variables sensitive to the presence of a $\tau^+\tau^-$ pair in the final state.

- Topological variables such as the thrust value [2, 3] (see Fig. S4), the first Fox-Wolfram moment shape variable [4] (see Fig. S5), the angles between the thrust direction and the directions of each of the two tagging muons.

- Variables built exploiting the information of neutral pions, which are abundant in τ decays and expected both for the signal and for the background. The angles between each π^0 momentum and the direction of each of the two tagging muons are used to define two cases: the first case corresponds to topologies in which a π^0 is close or opposite to the directions of both tagging muons, the second case includes all the other topologies. For each case, we consider the sum of the energies of all neutral pions as discriminant variable. Background events contribute mostly to the first case, while signal is more uniformly distributed.

- Variables built exploiting the rest-of-event (ROE), which is the system made of all charged and neutral particles except the two tagging muons: the difference between $M(4 \text{ tracks})$ and the sum of the ROE invariant mass, computed assuming the pion mass hypothesis for tracks and zero mass for neutrals, and the invariant mass of the two tagging muons; the difference between the total energy of the four-track system and the sum of the ROE energy and the total energy of the two tagging muons.

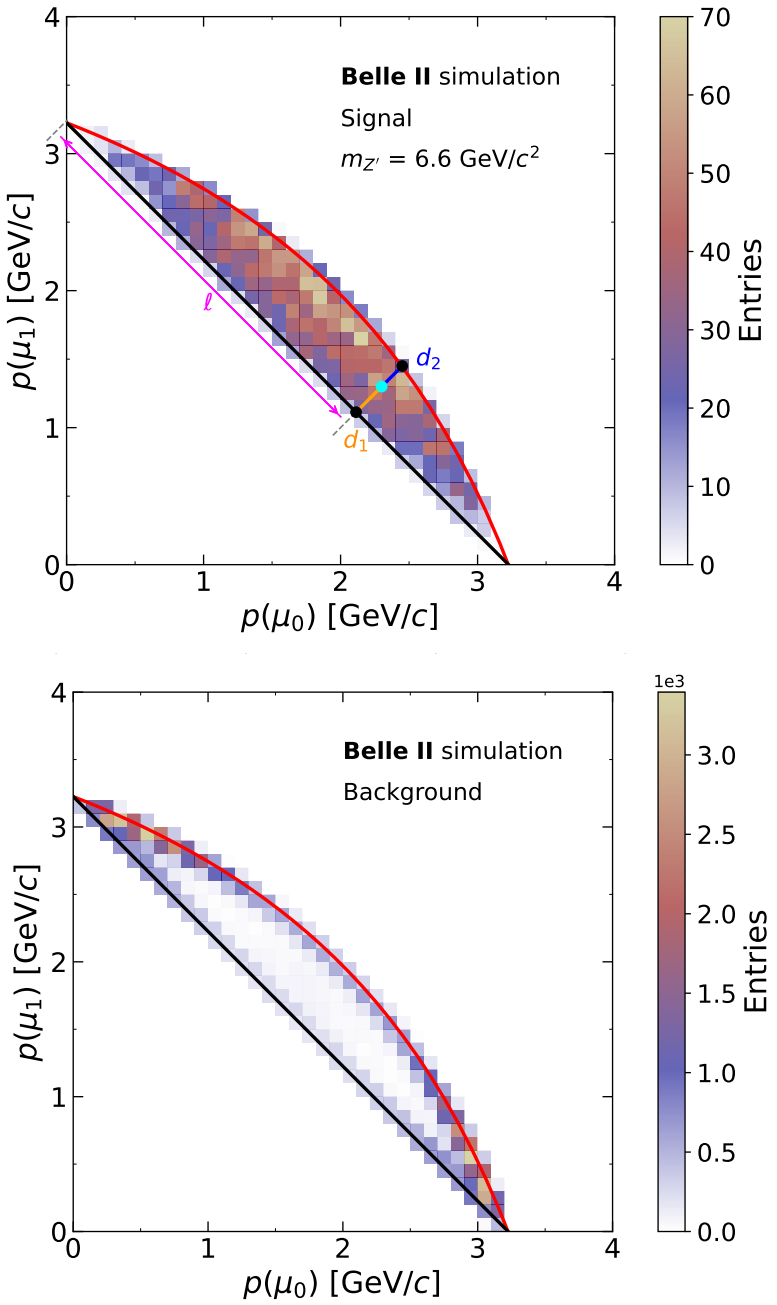

Figure S1: Distribution of the magnitudes of the momenta of the two tagging muons for simulated signal (*top*) and background (*bottom*) events. Signal is generated with a mass $m_{Z'} = 6.6$ GeV/c^2. The reconstructed mass for signal and background is required to be in the interval 6.60 ± 0.05 GeV/c^2, that is within five times the mass resolution. The cyan point is a generic point of the distribution. The black points are the intersections of a straight line, perpendicular to the black line and passing through the cyan point, with the black line and the red hyperbola. The orange and blue segments are the distances d_1 and d_2, respectively. The magenta segment l is the coordinate of the cyan point along the black line.

The following figures show the four most discriminating variables before the MLP selection for data and simulation. In all the cases, signal is generated with a mass $m_{Z'} = 6.6$ GeV/c^2, and the reconstructed mass for signal and background is required to be in the interval 6.60 ± 0.05 GeV/c^2, that is within five times the mass resolution.

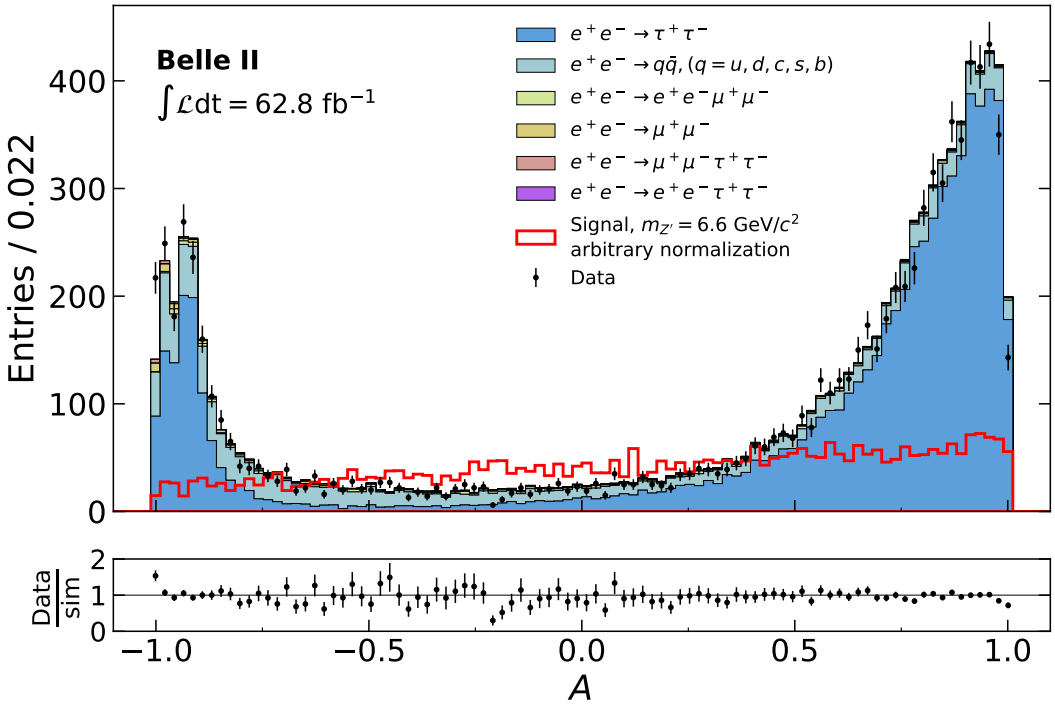

Figure S2: Distribution of the variable A in the c.m. frame, for data and simulation. A is an asymmetry-like variable calculated from the momenta of the two tagging muons. Contributions from the various simulated background processes are stacked. The simulation is normalized to the data luminosity, while the normalization of the signal is arbitrary.

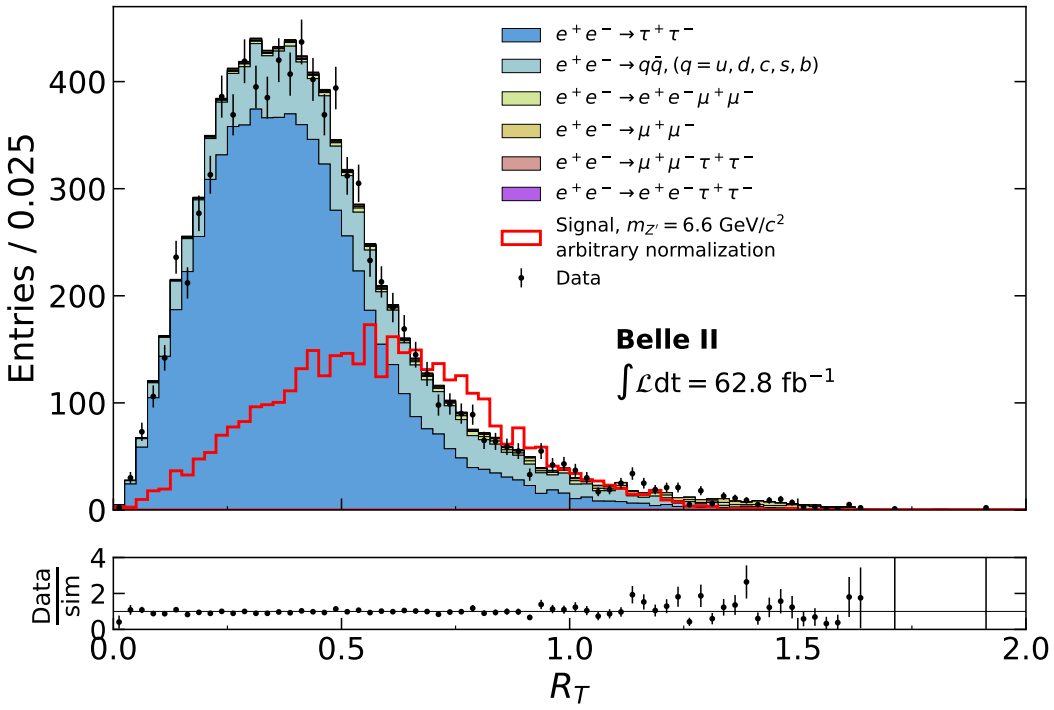

Figure S3: Distribution of the quadratic sum of the components of the recoil momentum transverse to the momentum direction of each of the two tagging muons in the c.m. frame, divided by the maximum recoil momentum kinematically reachable, for data and simulation. Contributions from the various simulated background processes are stacked. The simulation is normalized to the data luminosity, while the normalization of the signal is arbitrary.

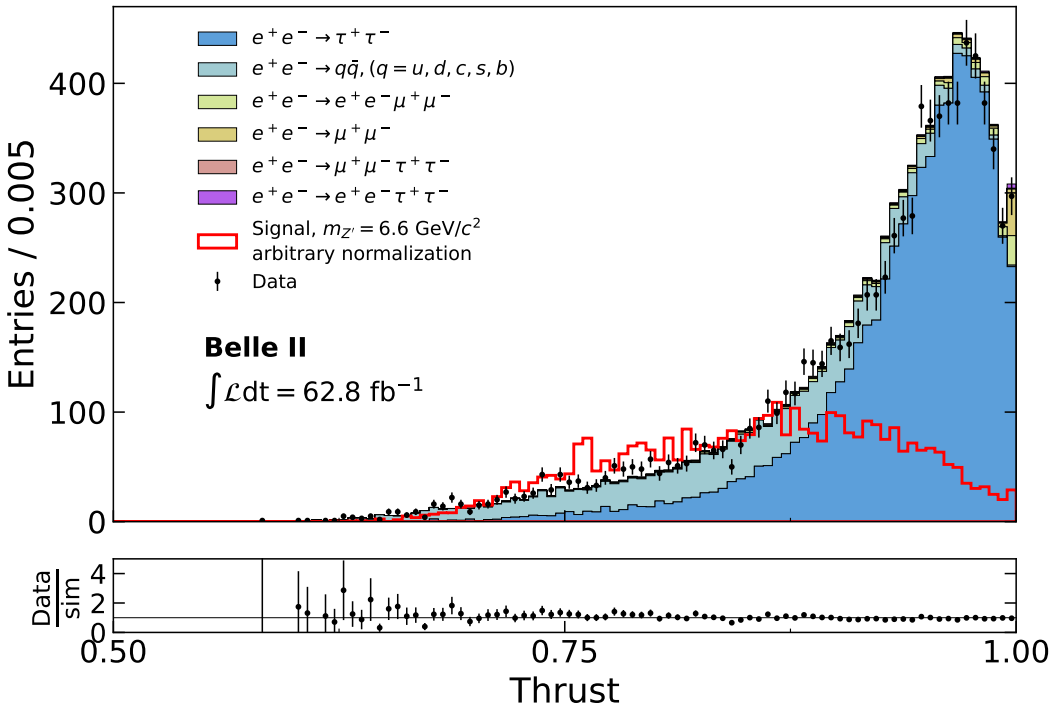

Figure S4: Thrust value distribution computed in the rest frame of the system recoiling against the two tagging muons, for data and simulation. Contributions from the various simulated background processes are stacked. The simulation is normalized to the data luminosity, while the normalization of the signal is arbitrary.

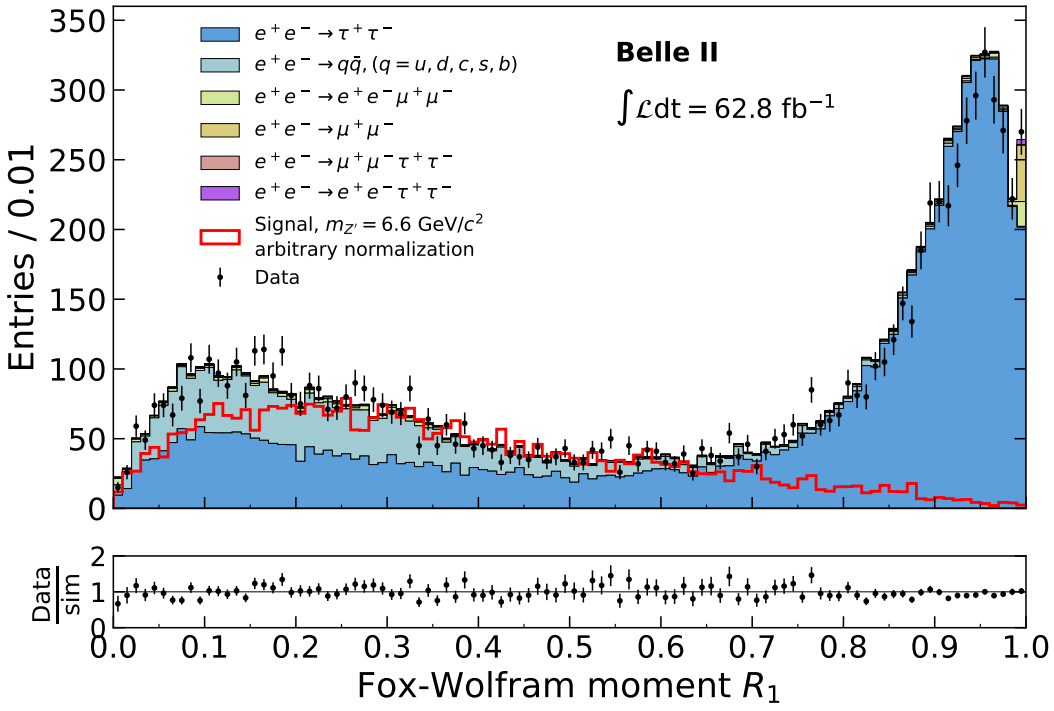

Figure S5: First Fox-Wolfram shape variable distribution computed in the rest frame of the system recoiling against the two tagging muons, for data and simulation. Contributions from the various simulated background processes are stacked. The simulation is normalized to the data luminosity, while the normalization of the signal is arbitrary.

Additional Figures

Fig. S6 shows the distribution of $M(4\text{ tracks})$ before the MLP selection, for $M(\mu\mu) > 2$ GeV/c^2. Fig. S7 shows the distribution of $M(4\text{ tracks})$ in the pion-tagged control sample after the MLP selection, for $M(\pi\pi) > 2$ GeV/c^2. These two distributions are used to understand the origin of the data-simulation discrepancies, as explained in the paper.

Fig. S8 shows the signal efficiency as a function of the the Z' mass after applying all the analysis selections.

Fig. S9 shows the fit corresponding to the highest significance case.

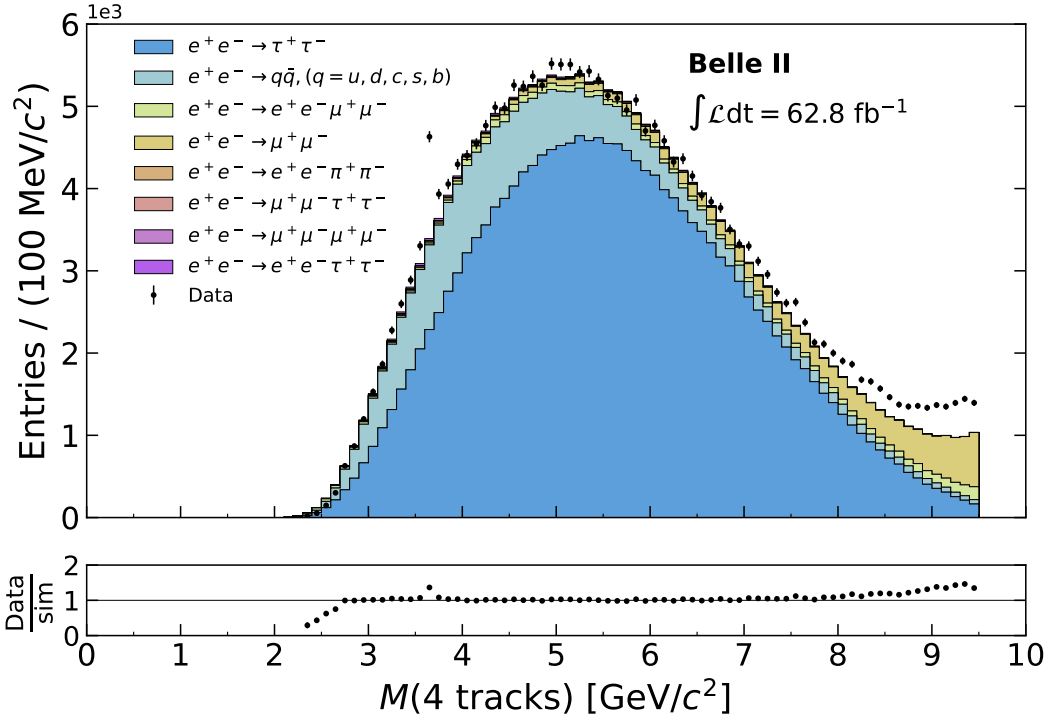

Figure S6: Observed four-track mass distribution, compared to the expectations of the simulation, before the MLP selection and for $M(\mu\mu) > 2$ GeV/c^2. Contributions from the various simulated processes are stacked. Also visible in data is the $\psi(2S)$ resonance at about 3.7 GeV/c^2 through the decay chain $\psi(2S) \to J/\psi\ \pi^+\pi^-$ with $J/\psi \to \mu^+\mu^-$, not present in the simulation. The large data-simulation discrepancy above 8 GeV/c^2 is due to un-modelled ISR in simulation of four-lepton processes.

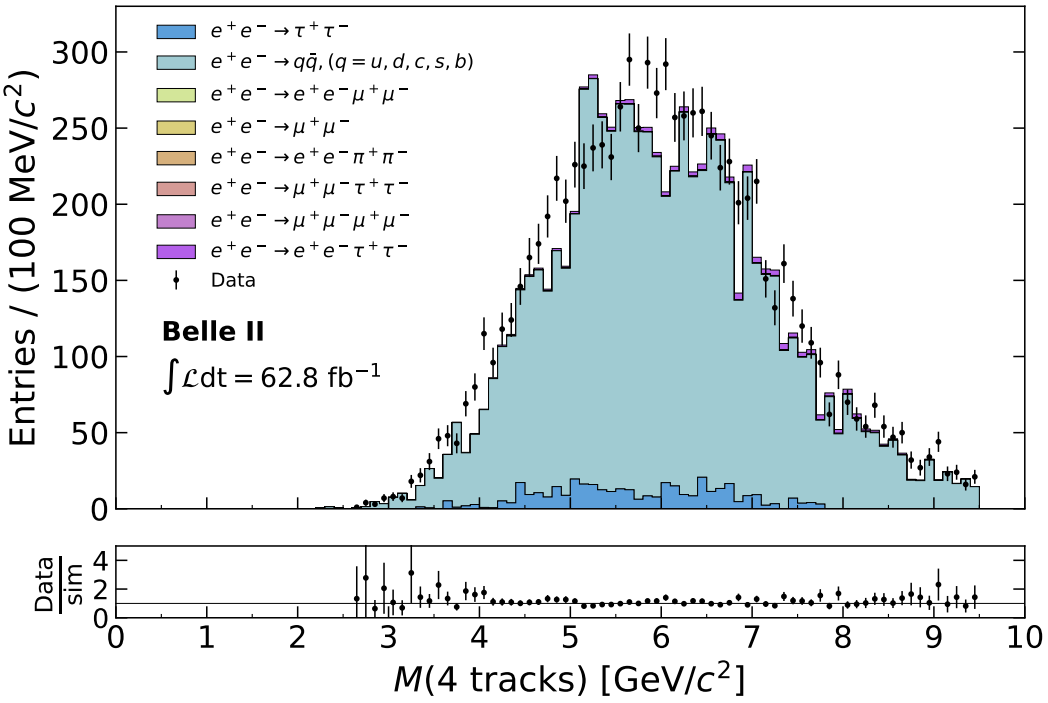

Figure S7: Observed four-track mass distribution, compared to the expectations of the simulation, in the pion-tagged control sample after all the selections and for $M(\pi\pi) > 2$ GeV/c^2. Contributions from the various simulated processes are stacked.

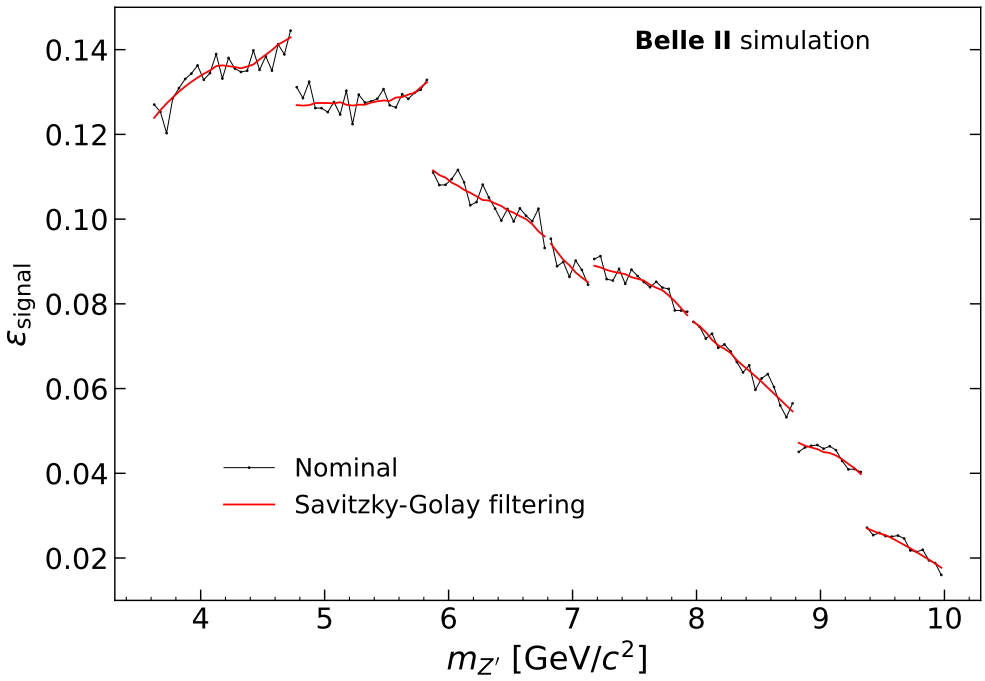

Figure S8: Signal efficiency as a function of the Z' mass. In the analysis we use smoothed values obtained through the application of a Savitzky-Golay filter [5] (red line).

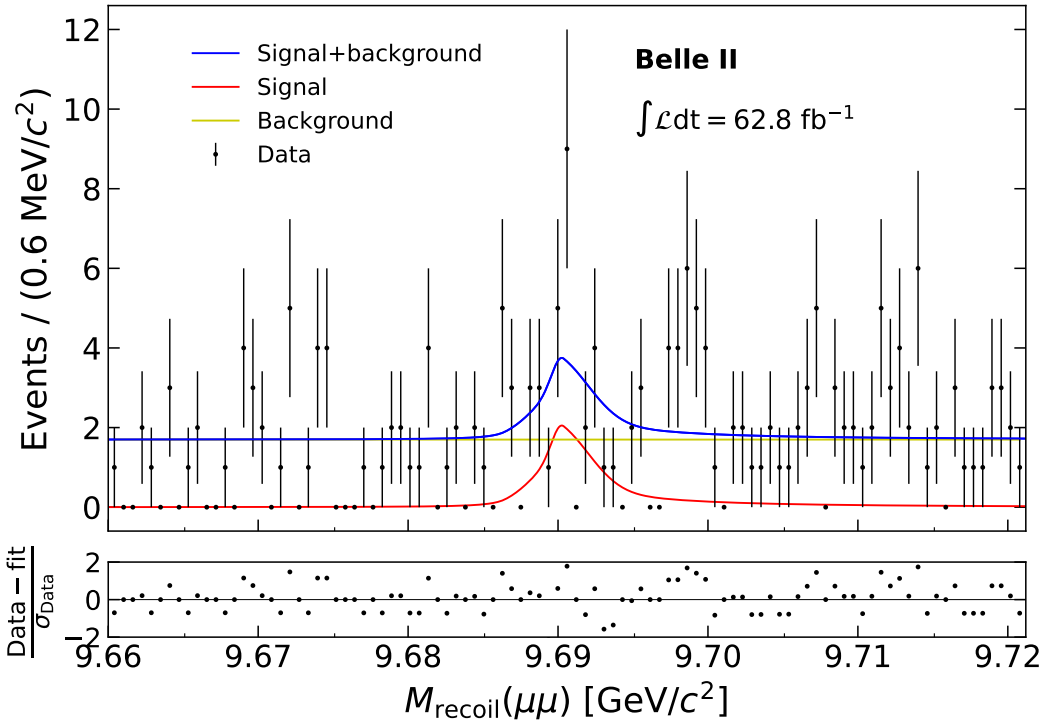

Figure S9: Fit for a Z' mass hypothesis of 9.696 GeV/c^2, for which we obtain the maximum local significance $3.0\,\sigma$ (see text for definition). The bottom panel shows the difference between the observed and fitted events, divided by the statistical uncertainty of the former.

Upper limits to the $L_\mu - L_\tau$ model

Upper limits on the coupling constants of the models are obtained from the upper limits on the cross sections, making use of the quadratic dependence. As an example, for the case of the $L_\mu - L_\tau$ model,

$$\mathrm{UL}(g')_{90\%\mathrm{CL}} = \sqrt{\frac{g'^2_{\mathrm{ref}} \cdot \mathrm{UL}(\sigma)_{90\%\mathrm{CL}}}{\sigma_{\mathrm{ref}}}}, \tag{1}$$

where g'_{ref} is a reference coupling constant used in the MadGraph5@NLO generator to compute a reference cross section (σ_{ref}).

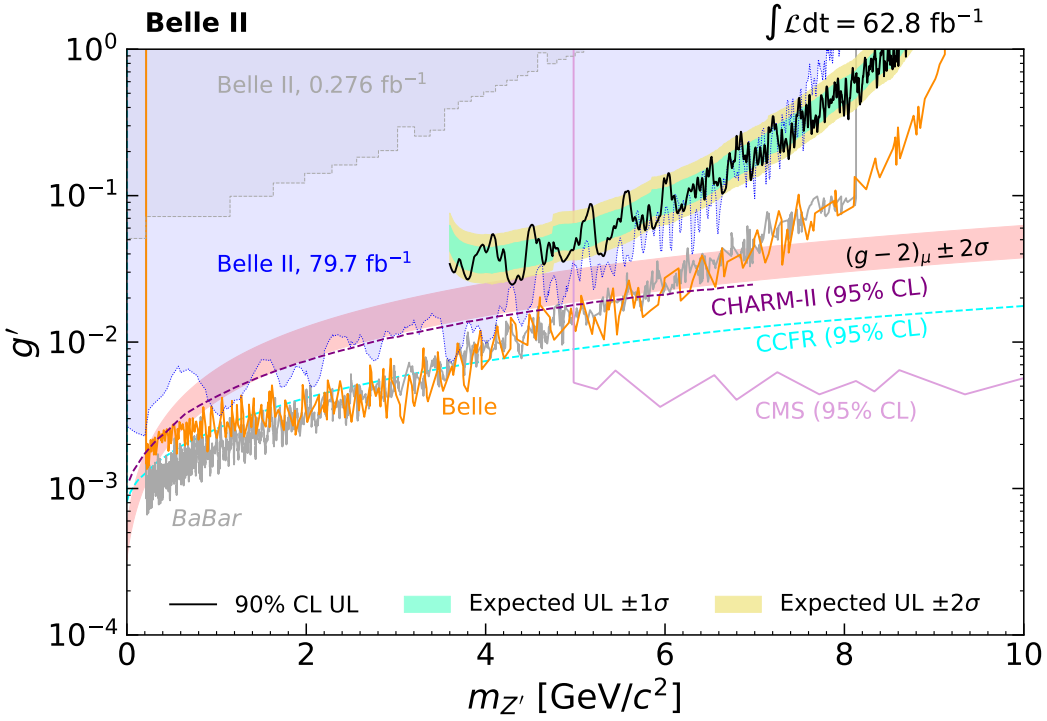

Figure S10: Observed 90% CL upper limits on the g' coupling of the $L_\mu - L_\tau$ model as a function of the Z' mass. Also shown are constraints from Belle II [1, 6] for invisible Z' decays, and from *BABAR* [7], Belle [8], and CMS [9] (95% CL) searches for Z' decays to muons, along with constraints (95% CL) derived from the trident production in neutrino experiments [10–12]. The red band shows the region that could explain the observed value (within two standard deviations) of the muon anomalous magnetic moment [13].

[1] I. Adachi et al. (Belle II Collaboration), Phys. Rev. Lett. **124**, 141801 (2020).
[2] E. Farhi, Phys. Rev. Lett. **39**, 1587 (1977).
[3] S. Brandt, C. Peyrou, R. Sosnowski, and A. Wroblewski, Phys. Lett. **12**, 57 (1964).
[4] G. C. Fox and S. Wolfram, Phys. Rev. Lett. **41**, 1581 (1978).
[5] W. H. Press and S. A. Teukolsky, Computers in Physics **4**, 669 (1990).
[6] I. Adachi et al. (Belle II Collaboration), Phys. Rev. Lett. **130**, 231801 (2023).
[7] J. P. Lees et al. (*BABAR* Collaboration), Phys. Rev. D **94**, 011102 (2016).
[8] T. Czank et al. (Belle Collaboration), Phys. Rev. D **106**, 012003 (2022).
[9] A. M. Sirunyan et al. (CMS Collaboration), Phys. Lett. B **792**, 345 (2019).
[10] W. Altmannshofer, S. Gori, M. Pospelov, and I. Yavin, Phys. Rev. Lett. **113**, 091801 (2014).
[11] G. Bellini et al. (Borexino Collaboration), Phys. Rev. Lett. **107**, 141302 (2011).
[12] A. Kamada, K. Kaneta, K. Yanagi, and H.-B. Yu, J. High Energy Phys. **2018**, 1 (2018).
[13] B. Abi et al. (Muon $g-2$ Collaboration), Phys. Rev. Lett. **126**, 141801 (2021).