LHCb-CONF-2024-001
April 2, 2024

Observation of the rare decay $J/\psi \to \mu^+\mu^-\mu^+\mu^-$

LHCb collaboration[†]

Abstract

The rare electromagnetic decay $J/\psi \to \mu^+\mu^-\mu^+\mu^-$ is observed with significance greatly exceeding the discovery threshold, using the proton-proton collision data collected by the LHCb experiment in 2016-2018 at a center-of-mass energy of 13 TeV, corresponding to an integrated luminosity of 5.4 fb^{-1}. The rate of this decay is measured relative to that of $J/\psi \to \mu^+\mu^-$ mode. Its branching fraction is determined to be

$$\mathcal{B}(J/\psi \to \mu^+\mu^-\mu^+\mu^-) = (1.13 \pm 0.10 \pm 0.05 \pm 0.01) \times 10^{-6},$$

where the uncertainties are statistical, systematic, and that due to the uncertainty on the branching fraction of the $J/\psi \to \mu^+\mu^-$ decay.

[†]Conference report prepared for the 58$^{\text{th}}$ Rencontres de Moriond session devoted to QCD and high energy interactions, La Thuile, Italy, 31 March – 7 April 2024. Contact authors: Vitalii Lisovskyi, lisovskyi@cppm.in2p3.fr and Lesya Shchutska, lesya.shchutska@epfl.ch.

1 Introduction

Decays of heavy-flavour mesons into final states containing more than two leptons are an interesting but less explored probe of the Standard Model (SM). The SM rates of such decays can be computed quite precisely. In addition, possible New Physics particles coupling to leptons can contribute to the rate of these decay processes.

Four-lepton decays are known for the light scalar π^0 and η mesons [1], where they proceed through the diphoton diagram. Recently, the BESIII collaboration reported the observation of $\eta' \to e^+e^-e^+e^-$ decay [2], and the CMS collaboration has observed the $\eta \to \mu^+\mu^-\mu^+\mu^-$ decay [3]. Besides these low-mass particles, four-lepton final states are actively explored for electroweak and Higgs boson measurements. However, four-lepton decays of heavy-quark hadrons are not well studied. The LHCb experiment has searched for the four-muon decay of $K^0_{\rm S}$ and $K^0_{\rm L}$ mesons [4], as well as B^0 and B^0_s mesons [5], that proceed through loop-level electroweak diagrams, but no signals were observed.

Recently, there has been a theoretical interest expressed for searching for the electromagnetic four-lepton decays of the vector quarkonia states [6]. Of a special interest is the J/ψ meson, a charmonium $c\bar{c}$ state with the quantum numbers $J^{PC} = 1^{--}$, that has large production cross-section rates in various experimental environments. It has a large electromagnetic decay width compared to its strong decay width, owing to the suppression of its strong decay modes due to Okubo–Zweig–Iizuka rule [7, 8]. This makes the four-lepton decay of the J/ψ an excellent tool for probing rare electromagnetic processes and the theory of quantum electrodynamics (QED) at the LHC.

The predicted decay rates in Ref. [6] are $(5.288\pm0.028)\times10^{-5}$, $(3.673\pm0.020)\times10^{-5}$ and $(0.0974 \pm 0.0005) \times 10^{-5}$ for $J/\psi \to e^+e^-e^+e^-$, $J/\psi \to e^+e^-\mu^+\mu^-$ and $J/\psi \to \mu^+\mu^-\mu^+\mu^-$, respectively, where the uncertainty comes only from the uncertainty on the known value of $\mathcal{B}(J/\psi \to \mu^+\mu^-)$. The difference in rates between decays to leptons of different flavours is due to the different lepton masses.

The BESIII experiment has searched for the four-lepton decays of the J/ψ meson [9], observing its decays to the $e^+e^-e^+e^-$ and $e^+e^-\mu^+\mu^-$ final states, with measured branching fractions being consistent with the theoretical predictions. The four-muon decay was not observed owing to its smaller rate, and only an upper limit was set on its branching fraction, $< 0.16 \times 10^{-5}$ at 90% confidence level. Recently, the CMS experiment has observed the $J/\psi \to \mu^+\mu^-\mu^+\mu^-$ decay, reporting a signal yield of $11.6^{+3.8}_{-3.1}$. The corresponding branching fraction was measured to be $[10.1^{+3.3}_{-2.7}\,(\text{stat}) \pm 0.4\,(\text{syst})] \times 10^{-7}$, assuming a phase-space J/ψ decay model for efficiency estimation [10].

Large production cross-sections of vector quarkonia states at the LHCb experiment [11, 12] make it a natural place for precision measurements of four-lepton quarkonia decays. However, most of the $c\bar{c}$, and all $b\bar{b}$ mesons are produced promptly in the proton-proton (pp) collision. The prompt production mechanism is prone to huge hadronic backgrounds, making decays of these mesons hard to measure. The alternative mechanism of $c\bar{c}$ production is that in decays of b hadrons, denoted secondary production, in which $c\bar{c}$ mesons are produced displaced from the pp interaction point. Electrons are more difficult to reconstruct and identify than muons, resulting in a worse mass resolution in decay modes with electrons. Therefore, the decay mode $J/\psi \to \mu^+\mu^-\mu^+\mu^-$ is the ideal candidate for a first study of four-lepton quarkonia decays at LHCb. As this decay is comparatively abundant, its study would also allow validation of the theoretical description of the final-state radiation, which can then be applied to much rarer four-lepton decays of B

mesons [5, 13].

The decay $J/\psi \to \mu^+\mu^-\mu^+\mu^-$ proceeds dominantly through the final-state radiation (FSR) of a virtual photon, γ^*, as shown in Fig. 1. The analogous initial-state radiation process is not expected to contribute significantly, as the process $J/\psi \to \gamma\gamma$ is forbidden by the requirement of C-parity conservation in electromagnetic processes. The rate of the FSR process is suppressed by two factors of the electromagnetic coupling constant compared to the dimuon decay without FSR. Further suppression comes from the relatively large muon mass: the equivalent process with a virtual photon transition into two electrons has a higher rate due to smaller electron mass. Finally, presence of identical muons in the final state contributes with interference terms between possible diagrams. It should be noted that most of the processes with $J/\psi \to VV \to \mu^+\mu^-\mu^+\mu^-$, where V is a vector resonance, that could lead to the same final state, are forbidden by C-parity conservation in strong and electromagnetic interactions.

Assuming that the SM rate is dominated by the FSR process, it is possible to separate the hadronic physics of the J/ψ annihilation into a virtual photon (identical for the $J/\psi \to \mu^+\mu^-\mu^+\mu^-$ and $J/\psi \to \mu^+\mu^-$ decays), and the electromagnetic process of FSR (only appearing in the $J/\psi \to \mu^+\mu^-\mu^+\mu^-$). Therefore, the ratio of branching fractions of $J/\psi \to \mu^+\mu^-\mu^+\mu^-$ and $J/\psi \to \mu^+\mu^-$ decays is an observable that can be theoretically predicted to a rather high precision. This observable is also convenient experimentally, allowing the cancellation of numerous sources of systematic uncertainties. The ratio of branching fractions is therefore defined as

$$R_{\mathcal{B}} \equiv \frac{\mathcal{B}(J/\psi \to \mu^+\mu^-\mu^+\mu^-)}{\mathcal{B}(J/\psi \to \mu^+\mu^-)}. \tag{1}$$

Throughout this document, the $J/\psi \to \mu^+\mu^-\mu^+\mu^-$ will be referred to as the "signal", and $J/\psi \to \mu^+\mu^-$ as the "normalisation" mode.

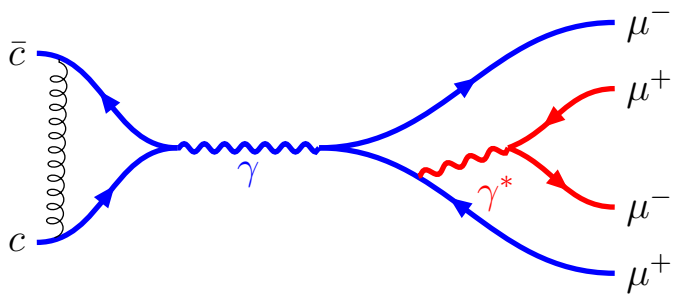

Figure 1: Tree-level diagram for the four-muon decay of J/ψ mesons. The virtual photon, γ^*, can be emitted from either muon leg. The blue part of the diagram corresponds to the $J/\psi \to \mu^+\mu^-$ decay.

The rest of this document is organised as follows: after a short introduction to the LHCb detector, data processing chain and simulation in Sec. 2, the candidate selection and background suppression is described in Sec. 3. The efficiency modelling is described in Sec. 4, invariant mass fits used to extract final results in Sec. 5, and systematic uncertainties in Sec. 6. This paper concludes with a discussion of the results in Sec. 7.

2 LHCb detector, datasets and simulation

The LHCb detector [14, 15] is a single-arm forward spectrometer covering the pseudorapidity range $2 < \eta < 5$, designed for the study of particles containing b or c quarks. The detector includes a high-precision tracking system consisting of a silicon-strip vertex detector surrounding the pp interaction region, a large-area silicon-strip detector located upstream of a dipole magnet with a bending power of about $4\,\mathrm{T\,m}$, and three stations of silicon-strip detectors and straw drift tubes placed downstream of the magnet. The tracking system provides a measurement of the momentum, p, of charged particles with a relative uncertainty that varies from 0.5% at low momentum to 1.0% at $200\,\mathrm{GeV}/c$. The minimum distance of a track to a primary pp collision vertex (PV), the impact parameter, is measured with a resolution of $(15 + 29/p_{\mathrm{T}})\,\mu\mathrm{m}$, where p_{T} is the component of the momentum transverse to the beam, in GeV/c. The particle identification is provided by several subdetectors. Two ring-imaging Cherenkov detectors mainly aim at distinguishing different types of charged hadrons, but also contribute to muon and electron identification at low momenta. A calorimeter system consisting of scintillating-pad and preshower detectors, an electromagnetic and a hadronic calorimeter, allows to distinguish photons, electrons, hadrons and muons. Muons are identified by a system composed of alternating layers of iron and multiwire proportional chambers. This system returns a binary value for positive muon identification if the muon candidate produces signal in a predefined number of stations that depends on its momentum [16]. Furthermore, a multivariate classifier combines information from each detector system, returning a single probability value, denoted here as P_μ, for the muon hypothesis [17].

This measurement relies on the data collected by the LHCb detector in 2016-2018, in the proton-proton collisions at a center-of-mass energy of 13 TeV. The online event selection is performed by a trigger [18], which consists of a hardware stage, based on information from the calorimeter and muon systems, followed by a software stage, which applies a full event reconstruction.

Simulation of the decays of interest is required to model the effects of the detector acceptance and the imposed selection requirements. In the simulation, pp collisions are generated using Pythia [19] with a specific LHCb configuration [20]. Decays of unstable particles are described by EvtGen [21], where final-state radiation is generated using Photos [22]. The interaction of the generated particles with the detector, and its response, are implemented using the Geant4 toolkit [23] as described in Ref. [24].

The decay $J/\psi \to \mu^+\mu^-$ is simulated using the EvtGen model for vector meson decays into two leptons. For the decay $J/\psi \to \mu^+\mu^-\mu^+\mu^-$, the simulation is generated according to an EvtGen model based on the analytical leading-order calculation of the QED decay rate, which assumes unpolarised J/ψ production. It is adapted from the model used by the BESIII experiment in Ref. [9].

The simulation of $J/\psi \to \mu^+\mu^-$ and $J/\psi \to \mu^+\mu^-\mu^+\mu^-$ decays undergoes several offline calibration steps to improve its agreement with data. Muon identification is calibrated, resampling the distributions of the P_μ for each muon, using large tag-and-probe calibration datasets of $J/\psi \to \mu^+\mu^-$ decays, collected in parallel to the regular data taking [17]. Additional $J/\psi \to \mu^+\mu^-$ datasets are used to calibrate the muon tracking efficiency and apply the correction weights per track to simulated events [25]. The hardware trigger response is also calibrated with the $J/\psi \to \mu^+\mu^-$ datasets using the tag-and-probe technique where the trigger decision was taken using non-signal particles

in the event. As mentioned above, the two dominant production mechanisms of J/ψ mesons are prompt production in the pp collision, or secondary production in decays of b hadrons. The proportion of J/ψ candidates produced through the prompt and secondary mechanisms in simulation is reweighted to the value measured at LHCb [11]. Finally, the distributions of p_{T} and η for the J/ψ candidates are weighted in simulation, separately for prompt and secondary production mechanisms, to match those of the $J/\psi \to \mu^+\mu^-$ data; this weighting is also propagated to the four-muon datasets.

3 Candidate selection

The data are analysed separately in two orthogonal categories, prompt and secondary. The prompt category contains candidates whose decay-length significance (DLS) with respect to the PV does not exceed three standard deviations, while candidates with DLS above this threshold are classified to the secondary category. The prompt category is characterised by an immense background from other particles, mostly hadrons, produced directly in the pp collisions. They can be misidentified as muons due to their decay in flight; in addition, energetic hadrons are also able to cross the LHCb calorimeter and reach the muon system. The secondary category offers less challenging background conditions, therefore an equivalent signal purity can be achieved with a looser selection. It also benefits from the trigger system of the LHCb experiment, optimised towards b-hadron decays.

At the hardware trigger stage, for both the signal and normalisation decays, one of the muons is required to have a transverse momentum above 1.3-1.8 GeV/c, depending on the data-taking period. At the first software trigger stage, at least one dimuon candidate is required in the prompt dataset; at least one dimuon or a displaced muon is required in the secondary dataset. At the second software trigger stage, a four-muon candidate is required in the prompt signal dataset; a displaced two-, three- or four-muon secondary vertex with a significant displacement from any primary pp interaction vertex is required in the secondary signal dataset. For the normalisation dataset, a dimuon consistent with a $J/\psi \to \mu^+\mu^-$ candidate is required at the second software trigger stage.

All muon track candidates used in this analysis are required to be positively identified as muons. This implies that their momentum has to exceed 3 GeV/c – the threshold needed to cross the calorimeter and reach the muon system.

The prompt four-muon dataset is constructed by forming a good vertex from two dimuon candidates, each with a good vertex quality. Both the dimuon and four-muon candidates are required to have DLS less than three standard deviations, in order to prevent overlap with the secondary dataset. Each muon candidate is required to have p_{T} greater than 500 MeV/c and p greater than 10 GeV/c.

The secondary four-muon candidates are reconstructed in a different manner, by directly forming a good-quality four-track vertex, displaced from the PV by more than three standard deviations. This avoids the double counting present in the prompt case. At least one muon candidate must have p_{T} greater than 500 MeV/c, and at least one other must have p_{T} greater than 300 MeV/c.

The prompt $J/\psi \to \mu^+\mu^-$ candidates are required to have two positively identified oppositely charged muons with p_{T} above 650 MeV/c and p above 10 GeV/c, forming a good vertex. The transverse momentum of the dimuon must exceed 3 GeV/c. The secondary

$J/\psi \to \mu^+\mu^-$ candidates are constructed from two positively identified oppositely charged muons with $p_{\rm T}$ above 500 MeV/c, forming a good vertex, displaced from the PV.

To reject fake tracks sharing track segments with real muons, the angle between each dimuon pair in the laboratory frame is required to exceed 0.5 mrad. All opposite-sign dimuon combinations are required to have invariant mass below 2900 MeV/c^2 in order to suppress contributions from $J/\psi \to \mu^+\mu^-$ decays combined with two other muons. For the $J/\psi \to \mu^+\mu^-$ decay, no further selection is required. Fewer than one per mille dimuon events have more than one $J/\psi \to \mu^+\mu^-$ candidate, and all candidates are retained, as pair production of J/ψ mesons is not infrequent [26].

In the four-muon dataset, the dominant remaining background source is the random combination of four tracks: muons or misidentified hadrons. A multivariate classifier is used to suppress this background. It is based on the boosted decision tree (BDT) classifier with gradient boost [27, 28], as implemented in the TMVA toolkit [29], and is trained separately for prompt and secondary $J/\psi \to \mu^+\mu^-\mu^+\mu^-$ decays. The training uses the following information: $p_{\rm T}$, η and the vertex quality of the four-muon candidate, as well as the minimum and maximum P_μ values for the muon hypothesis amongst the four muon candidates. In the secondary J/ψ dataset, in addition, the angle between the momentum vector of the J/ψ candidate and the vector joining the PV and the J/ψ decay vertex, is used. As a signal proxy, simulated $J/\psi \to \mu^+\mu^-\mu^+\mu^-$ candidates are used, with all calibrations applied, as discussed above. As a background proxy, four-muon candidates in data are selected in the invariant-mass sidebands, $2700 < m(\mu^+\mu^-\mu^+\mu^-) < 2900$ MeV/c^2 and $3200 < m(\mu^+\mu^-\mu^+\mu^-) < 3400$ MeV/c^2. The requirement on the BDT output value is optimised by maximising the quantity $S/\sqrt{S+B}$. Here, S is the expected $J/\psi \to \mu^+\mu^-\mu^+\mu^-$ yield passing a certain BDT requirement, estimated assuming the predicted branching fraction value from Ref. [6], and B is the background yield in the signal region.

A possible contamination due to misidentified hadrons in four-body decays of quarkonia, such as the $J/\psi \to \pi^+\pi^-\pi^+\pi^-$ decay or the as yet unobserved decay $J/\psi \to \pi^+\pi^-\mu^+\mu^-$, is tested by requiring a positive pion or kaon identification on at least two muons, as well as by loosening the muon identification criteria. No hint of such backgrounds is seen.

As the $J/\psi \to \mu^+\mu^-\mu^+\mu^-$ decay is rare, it is unlikely to occur twice in the same event even in case of the pair J/ψ production. As the prompt four-muon candidates are built from two dimuon candidates, constructed independently, each four-muon candidate is reconstructed twice, unless one of the combinations fails the dimuon vertex-quality requirement. Therefore, only one candidate per event is preserved randomly, both in data and simulation, after the full selection is applied. In the secondary four-muon candidates, no duplication occurs at the reconstruction stage, and less than 1% of events contain multiple candidates. Where multiple candidates exist, one is selected randomly.

4 Efficiency estimation

In order to measure the value of $R_{\mathcal{B}}$, the ratio of the observed $J/\psi \to \mu^+\mu^-\mu^+\mu^-$ and $J/\psi \to \mu^+\mu^-$ decays needs to be corrected for the total detection efficiency for each of the two decay modes. The efficiency estimation is performed using a simulated dataset of inclusive J/ψ decays, *i.e.* with both prompt and secondary components. This aims for a consistent treatment of simulation and data: although in simulation it is possible to fully distinguish the prompt and nonprompt components using truth-level information,

this cannot be done in data. Under this definition, the maximum possible efficiency for the prompt (secondary) J/ψ equals to the fraction of prompt (secondary) J/ψ mesons in reweighted simulation.

The total efficiency accounts for the LHCb acceptance (common for prompt and secondary datasets), trigger, reconstruction and selection requirements. The values of total efficiency are shown in Table 1, where the quoted precision corresponds to the uncertainty due to the size of simulated samples, discussed later. The efficiency ratio is much lower in the prompt dataset, due to the tight requirements imposed in the trigger and offline selection. It is worth pointing out that the reweighting of the J/ψ kinematics changes the efficiency ratio by about 10% (30%) in the secondary (prompt) datasets. This is because simulation underestimates the fraction of low-momentum J/ψ mesons, which typically have low efficiency, especially in the prompt case. A large shift in efficiencies, exceeding 30%, is also observed when comparing the simulation based on the QED calculation with less motivated models, such as those with kinematics uniformly distributed across the phase space. This is due to the fact that the decay model choice impacts the simulated muon kinematics, and, therefore, the efficiency of the selection requirements on the muons. Therefore, a simultaneous analysis of the two datasets is a powerful tool to cross-check internally the obtained results.

Table 1: Total efficiency for the $J/\psi \to \mu^+\mu^-\mu^+\mu^-$ and $J/\psi \to \mu^+\mu^-$ decays, as well as their ratio.

Category	$J/\psi \to \mu^+\mu^-\mu^+\mu^-$	$J/\psi \to \mu^+\mu^-$	Ratio
Prompt	0.052%	1.992%	0.0262
Secondary	0.085%	0.518%	0.1641

5 Mass modelling

Signals are separated from the residual background with fits to the four-muon (dimuon) invariant mass for signal (normalisation) datasets. The maximum-likelihood fit is performed separately for each studied dataset. It is implemented in the RooFit toolkit [30] within the ROOT framework [31].

First, fits to the invariant-mass distributions of the selected dimuon candidates are performed. The signal shape is parameterised with a linear combination of two double-sided Crystal Ball distributions [32] sharing the same mean, and one Gaussian distribution. The shape parameters are determined by fitting the simulated datasets. In the fit to data, all the tail parameters of the signal shape, as well as the relative fractions of its components, are fixed to the values obtained from simulation. One mean and one resolution parameter are allowed to vary, as simulation slightly underestimates the resolution. The difference between the means of the Gaussian and Crystal Ball distributions is fixed from simulation, as are the ratios of the width parameters of the three sub-components. A combinatorial-background shape is added, parameterised by a linear distribution with a slope parameter that is allowed to vary. Due to large size of the dimuon datasets, binned maximum-likelihood fits are performed with 10 000 bins. The resulting invariant-mass distributions in the prompt and secondary $J/\psi \to \mu^+\mu^-$ datasets, as well as the fit results,

are shown in Fig. 2: $341\,880\,000 \pm 21\,000$ $J/\psi \to \mu^+\mu^-$ decays are observed in the prompt dataset, and $91\,703\,800 \pm 11\,400$ in the secondary dataset.

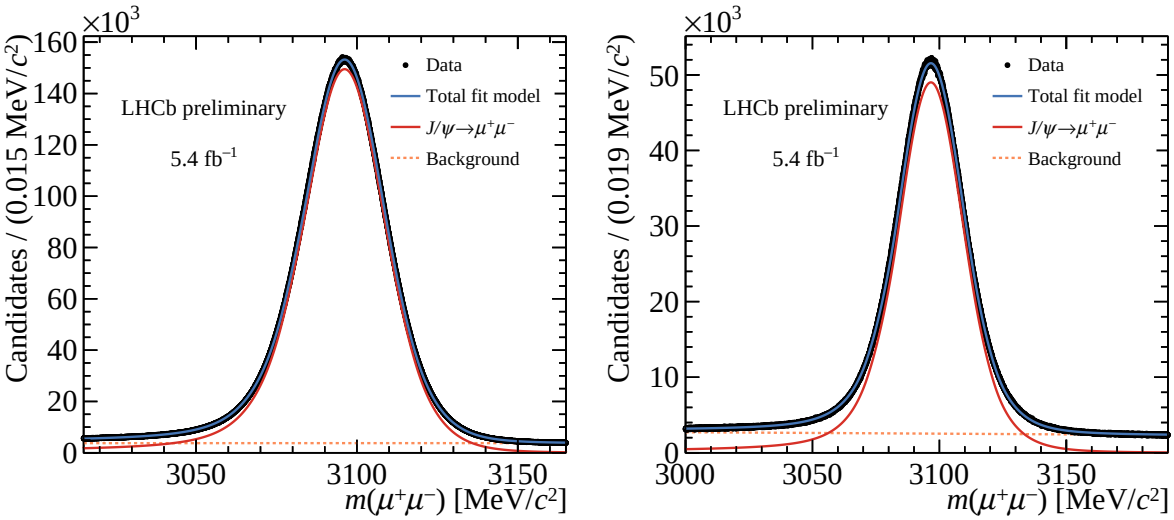

Figure 2: Invariant mass distributions for (left) prompt $J/\psi \to \mu^+\mu^-$ candidates and (right) secondary $J/\psi \to \mu^+\mu^-$ candidates. The results of the fit are overlaid.

Unbinned maximum-likelihood fits are used to describe the $J/\psi \to \mu^+\mu^-\mu^+\mu^-$ datasets. The signal shape is parameterised by a linear combination of two double-sided Crystal Ball distributions sharing the same mean. As previously, the shape parameters are determined from simulation and fixed in the fit to data, except for one mean and one resolution parameter that are allowed to vary. The background distribution is parameterised by an exponential function with a slope parameter, which is allowed to vary. In the prompt dataset, the J/ψ candidate is constrained to originate from the PV, which slightly improves the resolution on the four-muon invariant mass. The $J/\psi \to \mu^+\mu^-\mu^+\mu^-$ yield is expressed as a product of the ratio of branching fractions of $J/\psi \to \mu^+\mu^-\mu^+\mu^-$ and $J/\psi \to \mu^+\mu^-$ decays ($R_{\mathcal{B}}$), the ratio of efficiencies between these two decay modes, and the previously measured $J/\psi \to \mu^+\mu^-$ yield. The quantity $R_{\mathcal{B}}$ is therefore a parameter extracted directly from the fit. The resulting four-muon invariant-mass distributions, with the fits overlaid, are shown in Fig. 3. A clear peak of $J/\psi \to \mu^+\mu^-\mu^+\mu^-$ decays is observed in both datasets, with the significance markedly exceeding the discovery threshold. The measured values of $R_{\mathcal{B}}$ are $(1.86 \pm 0.30) \times 10^{-5}$ (taking into account the fit bias correction, discussed below) and $(1.90 \pm 0.20) \times 10^{-5}$ in the prompt and secondary datasets, respectively, where the quoted uncertainty is statistical only. The two categories show a good agreement.

If the fit is repeated with the $J/\psi \to \mu^+\mu^-\mu^+\mu^-$ signal yield as a free parameter, yields of 166 ± 27 and 286 ± 30 decays are measured in the prompt and secondary datasets, respectively. One can note that, unlike in the dimuon case, the yield of prompt $J/\psi \to \mu^+\mu^-\mu^+\mu^-$ decays is lower compared to the secondary one, which is explained by the lower selection efficiency in the prompt case.

Several cross-checks are performed to ensure the accuracy of the obtained results. The behaviour of each fit model is validated using a large number of pseudoexperiments. While good behaviour is found in the secondary dataset, the lower signal purity in the prompt dataset leads to an underestimation of the uncertainty of about 4%. A bias

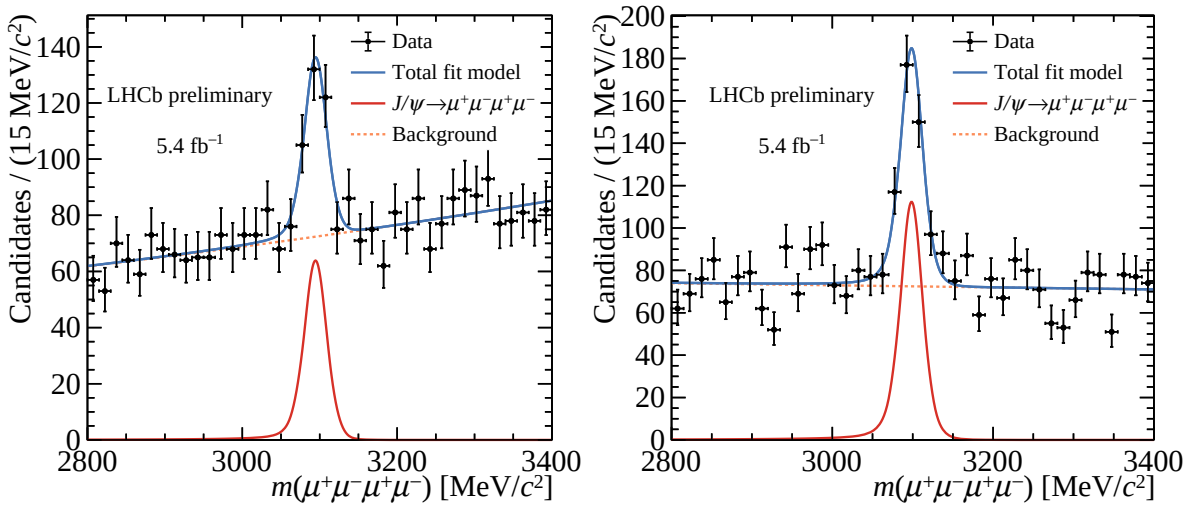

Figure 3: Invariant mass distributions for (left) prompt $J/\psi\to\mu^+\mu^-\mu^+\mu^-$ candidates and (right) secondary $J/\psi\to\mu^+\mu^-\mu^+\mu^-$ candidates. The results of the fit are overlaid.

correction of 4% is therefore applied to the uncertainty returned by the fit in the prompt case. Fits are repeated in a narrower mass range, or by fixing the resolution parameter under the assumption that the resolution differs between the data and simulation by the same amount in the dimuon and four-muon datasets: consistent results are obtained. The results are also extracted separately for each data-taking period and are found to agree within the expected statistical scatter.

Agreement between the background-subtracted data and calibrated simulation is examined in the kinematic observables of interest in the $J/\psi\to\mu^+\mu^-\mu^+\mu^-$ decays. The background subtraction is performed using the *sPlot* method [33]. As shown in Fig. 4, the dimuon invariant-mass distributions match well within the statistical uncertainties, showing that the QED model describes the data accurately at the current precision level. In contrast, the model with kinematics uniformly distributed across the phase space does not match the data.

6 Systematic uncertainties

A number of systematic effects, affecting both efficiency estimation and the mass modelling, are studied.

The uncertainty due to limited amount of simulated sample is propagated to the final result. In simulation, a small fraction of signal candidates fail the truth-matching algorithm. Studies show that this fraction is larger for the four-muon dataset compared to the dimuon one and an uncertainty is assigned for the non-cancellation in the ratio. Furthermore, the proportion of prompt and secondary J/ψ mesons in simulation is altered within the uncertainty on its value from Ref. [11].

For each correction applied to the signal simulation, a respective uncertainty is estimated. The calibrations of the kinematics of the J/ψ meson, as well as the hardware trigger response, are reevaluated with alternative binning schemes. An alternative dataset is also used to validate the tag-and-probe efficiency measurement for the hardware trigger.

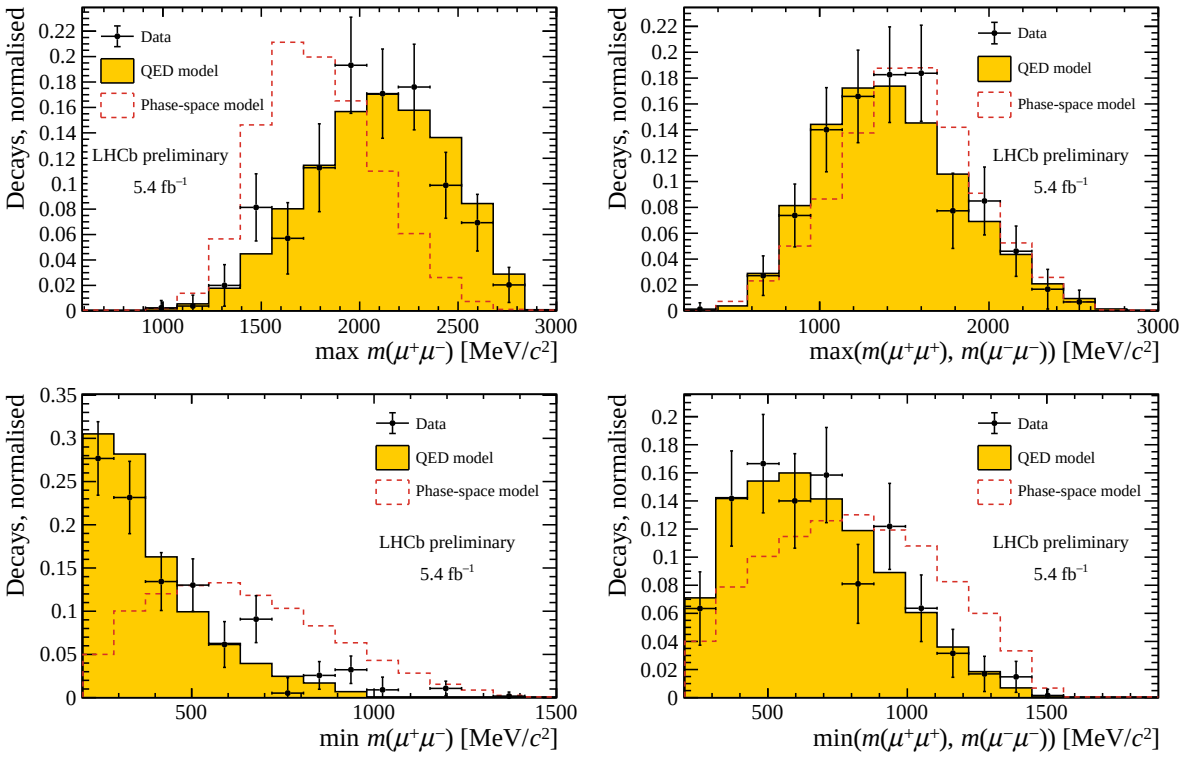

Figure 4: Comparison of background-subtracted data and calibrated simulation relying on the QED model in the secondary $J/\psi \to \mu^+\mu^-\mu^+\mu^-$ dataset. The simulated phase-space model is added for comparison. The secondary dataset is chosen here, as it provides the best statistical power. For the data, only statistical uncertainties are shown.

Similar studies are performed for the software trigger response. The uncertainties on the tracking efficiency calibration procedure, both of statistical and systematic nature, are propagated to the final result.

Detailed studies are performed to estimate the accuracy of the muon identification calibration. In the resampling procedure of the P_μ distributions, the kernel size is modified and the resampling is repeated. This allows the assessment of the uncertainty due to nonuniform statistical coverage of the calibration sample. Furthermore, the bias coming from the inaccurate background subtraction procedure in the calibration samples is studied. This effect is pronounced at low muon p_{T} and therefore affects the four-muon datasets more than the dimuon ones. The associated systematic uncertainty is estimated by performing dedicated fits to the calibration data in intervals of p_{T} and measuring the difference in muon efficiency as compared to the background-subtracted data.

Alternative models are studied to parameterise the signal and background shapes in the invariant-mass fits. For the signal shapes in each dataset, a Hypatia distribution [34] is used. For the background shapes, the second-degree polynomial distribution is used in the four-muon dataset, with the parameters of the polynomial extracted from a fit to a wider mass range. In the dimuon datasets, an exponential shape is used. Each alternative model is used to generate a large number of pseudodatasets, and the fit with the default model is repeated to estimate the average bias. As expected, in the dimuon datasets this systematic uncertainty is larger compared to the statistical uncertainty, however, it is still negligible compared to the statistical precision of the $J/\psi \to \mu^+\mu^-\mu^+\mu^-$ yield.

The procedure to remove multiple candidates is also scrutinised. In the dimuon datasets, where all candidates are preserved in the default approach, only one randomly chosen candidate is kept. In the secondary four-muon dataset, where the fraction of the multiple candidates is about 1% but they are removed in the default approach, all candidates are kept. Finally, in the prompt four-muon dataset, where the reconstruction procedure leads to the presence of duplicate candidates for most events, it is not meaningful to retain all candidates. Instead, the criterion for selecting the one candidate to be kept is modified by selecting the candidates that are rejected by the default procedure.

A summary of all sources of systematic uncertainty on $R_{\mathcal{B}}$ is presented in Table 2. Each source of uncertainty is assumed to be either fully correlated or fully uncorrelated between the two analysis categories.

Table 2: Summary of systematic uncertainties considered in this analysis. All values are quoted in percent of $R_{\mathcal{B}}$. The relative statistical uncertainty is also quoted for reference.

Source	Secondary	Prompt	Correlated?
Size of simulated samples	0.6	0.8	no
Treatment of simulation	1.1	0.8	yes
Fraction of secondary J/ψ component	0.3	0.2	no
Kinematic calibration	0.5	0.2	no
Hardware trigger calibration	0.3	0.4	yes
Software trigger calibration	1.0	0.4	no
Tracking calibration	2.3	1.8	yes
Muon ID calibration	3.1	5.8	yes
Mass modelling	0.6	0.7	no
Multiple candidate selection	0.6	0.8	no
Total uncorrelated	1.5	1.5	–
Total correlated	4.1	6.2	–
Total systematic uncertainty	4.4	6.3	–
Statistical uncertainty	10.5	16.2	no

Finally, the known value of $\mathcal{B}(J/\psi \to \mu^+\mu^-) = 5.961 \pm 0.033\%$ [1] is used to determine the absolute value of $\mathcal{B}(J/\psi \to \mu^+\mu^-\mu^+\mu^-)$. The uncertainty on this value is propagated to the final result as a separate term.

7 Results and conclusion

In the dataset of secondary J/ψ candidates, the ratio of branching fractions of the $J/\psi \to \mu^+\mu^-\mu^+\mu^-$ and $J/\psi \to \mu^+\mu^-$ decays is found to be $R_{\mathcal{B}} = (1.90 \pm 0.20 \pm 0.08) \times 10^{-5}$, where the first uncertainty is statistical and second systematic. In the prompt J/ψ dataset, the measured value is $R_{\mathcal{B}} = (1.86 \pm 0.30 \pm 0.12) \times 10^{-5}$. These two results are in excellent agreement and show stability with respect to splits by data-taking periods.

The results from the two analysis categories are averaged. Statistical uncertainties are treated as uncorrelated. Systematic uncertainties are treated as either fully correlated or uncorrelated, as described in Table 2. This results in the value

$$R_{\mathcal{B}} = (1.89 \pm 0.17 \pm 0.09) \times 10^{-5}, \tag{2}$$

which agrees within 1.4 standard deviations with the leading-order QED calculation of $R_{\mathcal{B}} = 0.163 \times 10^{-4}$ from Ref. [6].

Multiplying the measured value of $R_{\mathcal{B}}$ by the known value of the dimuon branching fraction of the J/ψ meson [1], one obtains

$$\mathcal{B}(J/\psi \to \mu^+\mu^-\mu^+\mu^-) = (1.13 \pm 0.10 \pm 0.05 \pm 0.01) \times 10^{-6}, \tag{3}$$

where the first uncertainty is statistical, the second is systematic, and the third is due to the uncertainty on $\mathcal{B}(J/\psi \to \mu^+\mu^-)$. This result relies on the simulated decay model based on the QED calculation. Indeed, the dimuon invariant-mass distributions in the background-subtracted $J/\psi \to \mu^+\mu^-\mu^+\mu^-$ data agree, within available statistical precision, with the expectations from the QED calculation; no significant resonant structures are observed.

To conclude, this analysis presents a first measurement of a four-lepton decay of quarkonium at the LHCb experiment. The $J/\psi \to \mu^+\mu^-\mu^+\mu^-$ decay is observed with significance far above the discovery threshold, and the most precise measurement to date of its branching fraction is performed. This result provides a complementary input to the study of $J/\psi \to e^+e^-e^+e^-$ and $J/\psi \to e^+e^-\mu^+\mu^-$ decays by the BESIII experiment [9]. It also is consistent with the recent measurement by the CMS collaboration [10]. To prove the robustness of the analysis procedure, the data are divided into two disjoint samples, corresponding to prompt and secondary J/ψ meson production, and analysed separately. The $J/\psi \to \mu^+\mu^-$ decay is used as a normalisation mode, also split by the production mechanism. The kinematic distributions in the $J/\psi \to \mu^+\mu^-\mu^+\mu^-$ decay are found to agree with the QED model, but differ significantly from the phase-space model. This observation of the $J/\psi \to \mu^+\mu^-\mu^+\mu^-$ decay paves the way for study of other multilepton decays at LHCb.

The authors would like to thank the BESIII collaboration and Jianping Dai for providing us with the EvtGen model for the decay of interest.

Appendix

A summary of existing measurements of the $J/\psi \to \mu^+\mu^-\mu^+\mu^-$ branching fraction is shown in Fig. 5.

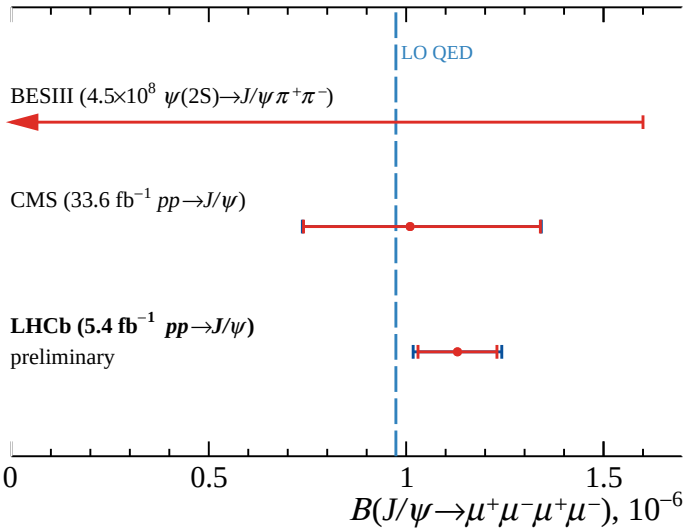

Figure 5: Summary of the measurements of the $J/\psi \to \mu^+\mu^-\mu^+\mu^-$ branching fraction by the BESIII [9] (upper limit at 90% CL) and CMS [10] experiments, compared with the result from this document. The statistical uncertainties are shown in red, while the total uncertainties in blue. The leading-order QED prediction from Ref. [6] is shown as a dashed line.

References

[1] Particle Data Group, R. L. Workman *et al.*, *Review of particle physics*, Prog. Theor. Exp. Phys. **2022** (2022) 083C01.

[2] BESIII collaboration, M. Ablikim *et al.*, *Observation of the double Dalitz decay* $\eta' \to e^+e^-e^+e^-$, Phys. Rev. **D105** (2022) 112010, arXiv:2203.12229.

[3] CMS collaboration, A. Hayrapetyan *et al.*, *Observation of the rare decay of the* η *meson to four muons*, Phys. Rev. Lett. **131** (2023) 091903, arXiv:2305.04904.

[4] LHCb collaboration, R. Aaij *et al.*, *Search for* $K^0_{S(L)} \to \mu^+\mu^-\mu^+\mu^-$ *decays at LHCb*, Phys. Rev. **D108** (2023) L031102, arXiv:2212.04977.

[5] LHCb collaboration, R. Aaij *et al.*, *Search for rare* $B^0_{(s)} \to \mu^+\mu^-\mu^+\mu^-$ *decays*, JHEP **03** (2022) 109, arXiv:2111.11339.

[6] W. Chen *et al.*, *Four-lepton decays of neutral vector mesons*, Phys. Rev. **D104** (2021) 094023, arXiv:2009.12363.

[7] L. Kopke and N. Wermes, J/ψ *decays*, Phys. Rept. **174** (1989) 67.

[8] M. B. Voloshin, *Charmonium*, Prog. Part. Nucl. Phys. **61** (2008) 455, arXiv:0711.4556.

[9] BESIII collaboration, M. Ablikim *et al.*, *Observation of* J/ψ *decays to* $e^+e^-e^+e^-$ *and* $e^+e^-\mu^+\mu^-$, Phys. Rev. D **109** (2024) 052006, arXiv:2111.13881.

[10] CMS collaboration, A. Hayrapetyan *et al.*, *Observation of the* $J/\psi \to \mu^+\mu^-\mu^+\mu^-$ *decay in proton-proton collisions at* $\sqrt{s} = 13$ *TeV*, arXiv:2403.11352.

[11] LHCb collaboration, R. Aaij *et al.*, *Measurement of forward* J/ψ *production cross-sections in pp collisions at* $\sqrt{s} = 13$ *TeV*, JHEP **10** (2015) 172, Erratum ibid. **05** (2017) 063, arXiv:1509.00771.

[12] LHCb collaboration, R. Aaij *et al.*, *Measurement of* Υ *production cross-section in pp collisions at* $\sqrt{s} = 13$ *TeV*, JHEP **07** (2018) 134, arXiv:1804.09214.

[13] A. V. Danilina and N. V. Nikitin, *Four-leptonic decays of charged and neutral B mesons within the Standard Model*, Phys. Atom. Nucl. **81** (2018) 347.

[14] LHCb collaboration, A. A. Alves Jr. *et al.*, *The LHCb detector at the LHC*, JINST **3** (2008) S08005.

[15] LHCb collaboration, R. Aaij *et al.*, *LHCb detector performance*, Int. J. Mod. Phys. **A30** (2015) 1530022, arXiv:1412.6352.

[16] F. Archilli *et al.*, *Performance of the muon identification at LHCb*, JINST **8** (2013) P10020, arXiv:1306.0249.

[17] R. Aaij *et al.*, *Selection and processing of calibration samples to measure the particle identification performance of the LHCb experiment in Run 2*, Eur. Phys. J. Tech. Instr. **6** (2019) 1, arXiv:1803.00824.

[18] R. Aaij *et al.*, *The LHCb trigger and its performance in 2011*, JINST **8** (2013) P04022, arXiv:1211.3055.

[19] T. Sjöstrand, S. Mrenna, and P. Skands, *A brief introduction to PYTHIA 8.1*, Comput. Phys. Commun. **178** (2008) 852, arXiv:0710.3820; T. Sjöstrand, S. Mrenna, and P. Skands, *PYTHIA 6.4 physics and manual*, JHEP **05** (2006) 026, arXiv:hep-ph/0603175.

[20] I. Belyaev *et al.*, *Handling of the generation of primary events in Gauss, the LHCb simulation framework*, J. Phys. Conf. Ser. **331** (2011) 032047.

[21] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. **A462** (2001) 152.

[22] N. Davidson, T. Przedzinski, and Z. Was, *PHOTOS interface in C++: Technical and physics documentation*, Comp. Phys. Comm. **199** (2016) 86, arXiv:1011.0937.

[23] Geant4 collaboration, J. Allison *et al.*, *Geant4 developments and applications*, IEEE Trans. Nucl. Sci. **53** (2006) 270; Geant4 collaboration, S. Agostinelli *et al.*, *Geant4: A simulation toolkit*, Nucl. Instrum. Meth. **A506** (2003) 250.

[24] M. Clemencic *et al.*, *The LHCb simulation application, Gauss: Design, evolution and experience*, J. Phys. Conf. Ser. **331** (2011) 032023.

[25] LHCb collaboration, R. Aaij *et al.*, *Measurement of the track reconstruction efficiency at LHCb*, JINST **10** (2015) P02007, arXiv:1408.1251.

[26] LHCb collaboration, R. Aaij *et al.*, *Measurement of J/ψ-pair production in pp collisions at $\sqrt{s} = 13$ TeV and study of gluon transverse-momentum dependent PDFs*, arXiv:2311.14085, submitted to JHEP.

[27] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, *Classification and regression trees*, Wadsworth international group, Belmont, California, USA, 1984.

[28] Y. Freund and R. E. Schapire, *A decision-theoretic generalization of on-line learning and an application to boosting*, J. Comput. Syst. Sci. **55** (1997) 119.

[29] H. Voss, A. Hoecker, J. Stelzer, and F. Tegenfeldt, *TMVA - Toolkit for Multivariate Data Analysis with ROOT*, PoS **ACAT** (2007) 040; A. Hoecker *et al.*, *TMVA 4 — Toolkit for Multivariate Data Analysis with ROOT. Users Guide.*, arXiv:physics/0703039.

[30] W. Verkerke and D. P. Kirkby, *The RooFit toolkit for data modeling*, eConf **C0303241** (2003) MOLT007, arXiv:physics/0306116.

[31] R. Brun and F. Rademakers, *ROOT: An object oriented data analysis framework*, Nucl. Instrum. Meth. **A389** (1997) 81.

[32] T. Skwarnicki, *A study of the radiative cascade transitions between the Upsilon-prime and Upsilon resonances*, PhD thesis, Institute of Nuclear Physics, Krakow, 1986, DESY-F31-86-02.

[33] M. Pivk and F. R. Le Diberder, *sPlot: A statistical tool to unfold data distributions*, Nucl. Instrum. Meth. **A555** (2005) 356, arXiv:physics/0402083.

[34] D. Martínez Santos and F. Dupertuis, *Mass distributions marginalized over per-event errors*, Nucl. Instrum. Meth. **A764** (2014) 150, arXiv:1312.5000.