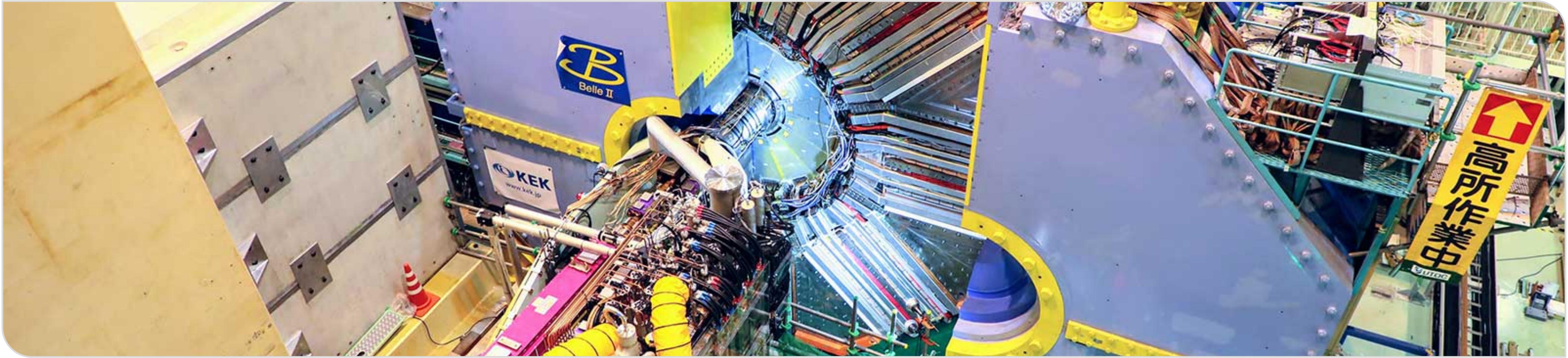
Search for ALPs in $e^+e^- \to \gamma a, a \to \gamma\gamma$ at Belle II

Belle II Germany Meeting, October 1st, 2024

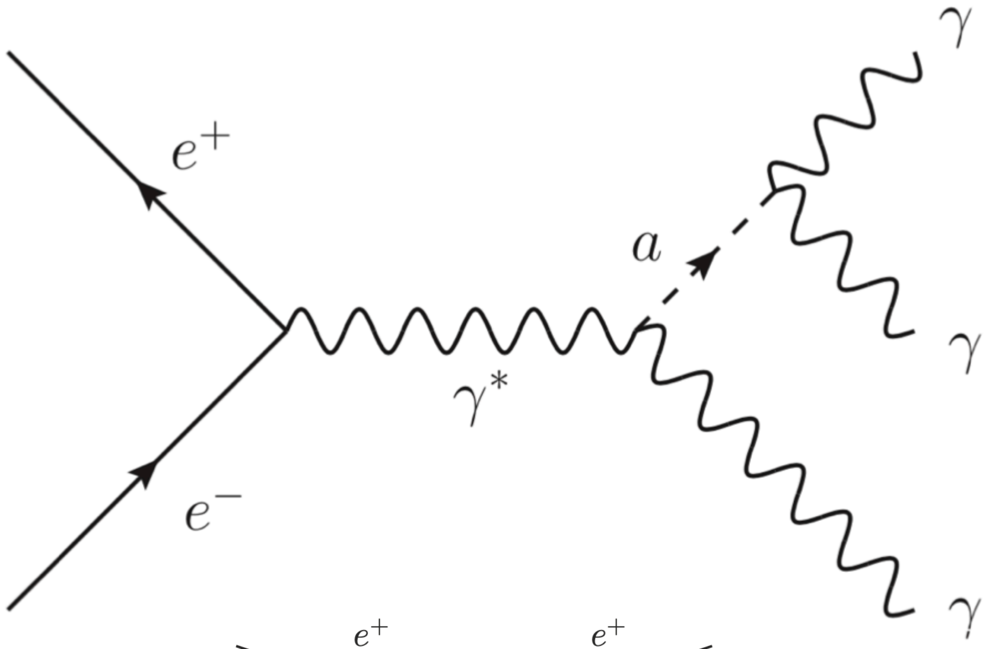
Alexander Heidelbach, Giacomo De Pietro, Torben Ferber, and Pablo Goldenzweig

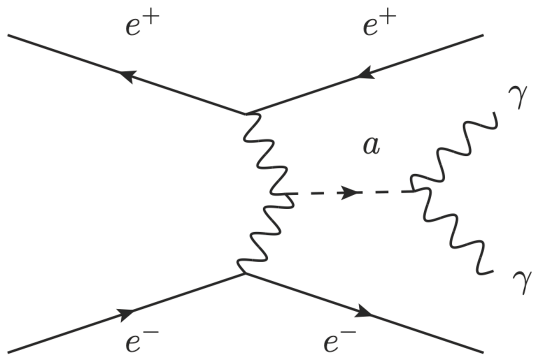
alexander.heidelbach@kit.edu



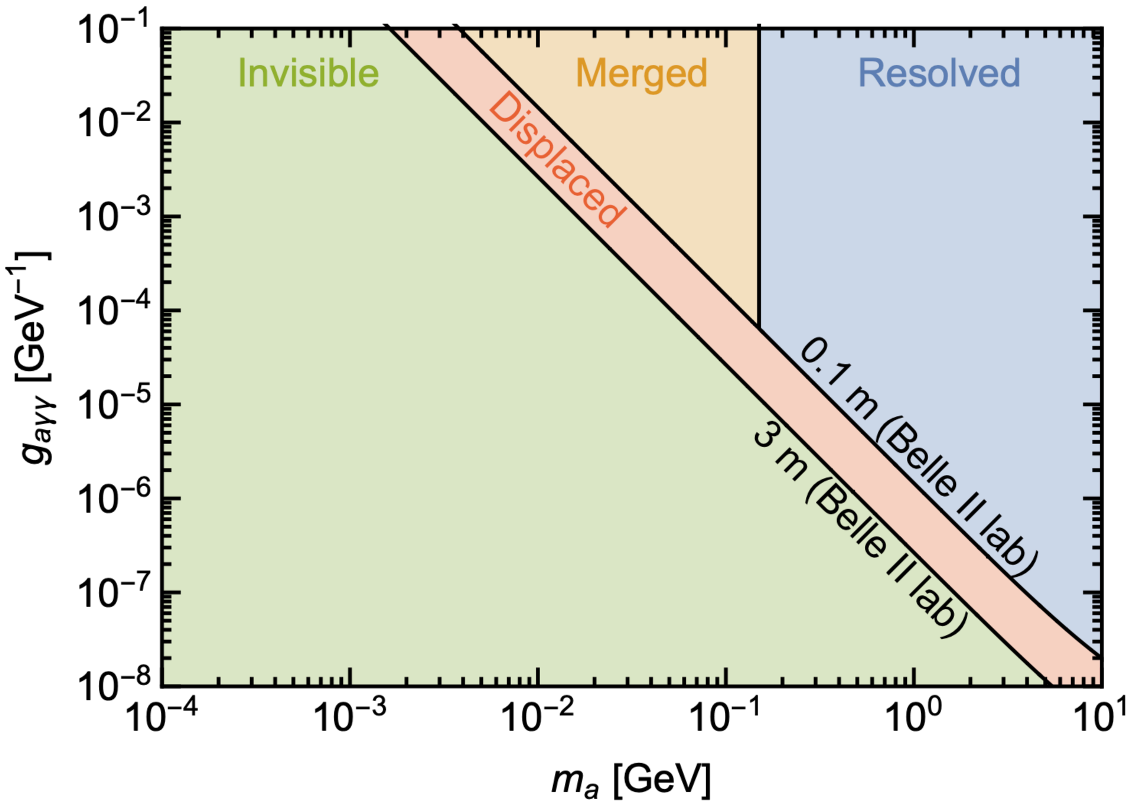
ALPs in e^+e^- decays

- **A**xion **L**ike **P**articles (**ALPs**)
  - Pseudo Goldstone boson of *spontaneously broken* and under SM *anomalous* Peccei-Quinn symmetry
- Our model: dominant coupling to photons:
  - **ALP-Strahlung** $e^+e^- \to \gamma a, a \to \gamma\gamma$
    - This search
  - **Photon-Fusion** $e^+e^- \to e^+e^- a, a \to \gamma\gamma$
    - Experimentally challenging due to the low scattering angle of finale state e^+e^-
- Search for ALP Strahlung performed in 2018 with $0.445\ \mathrm{fb}^{-1}$ (Belle II, Phys. Rev. Lett. 125 (2020) 161806)





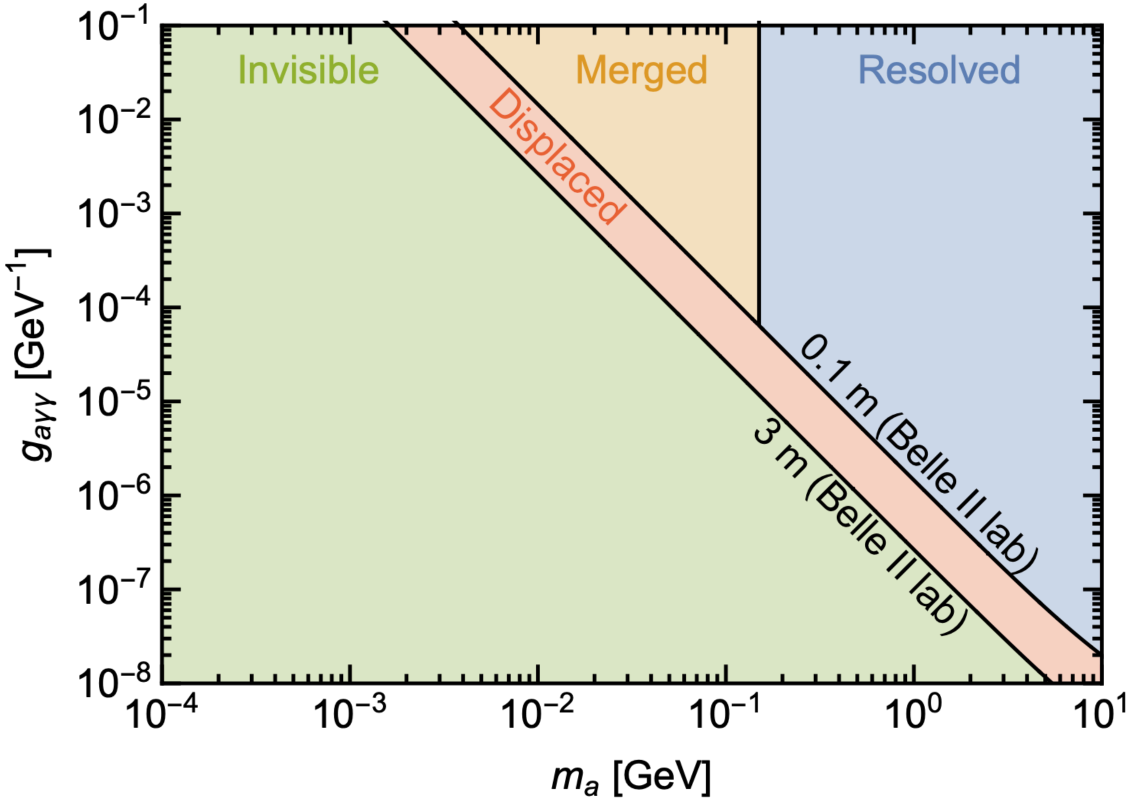
ALP-Strahlung decay at Belle II



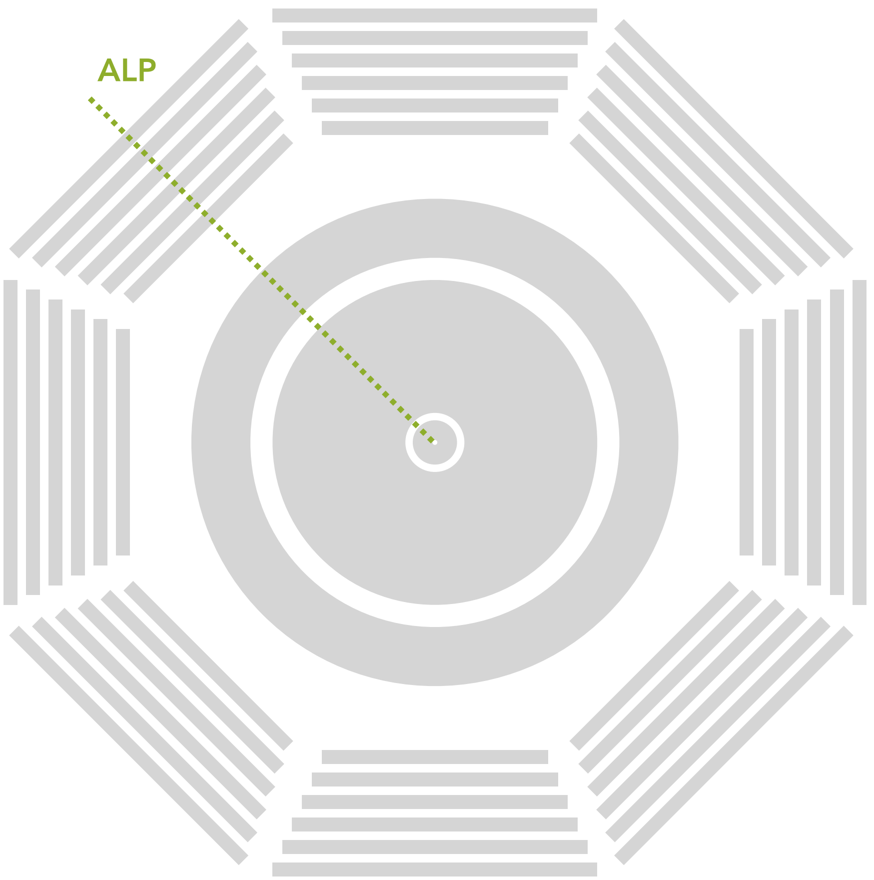
Dolan et al., J. High Energy Phys. 1712 (2017) 094

ALP-Strahlung decay at Belle II

Institute of Experimental Particle Physics

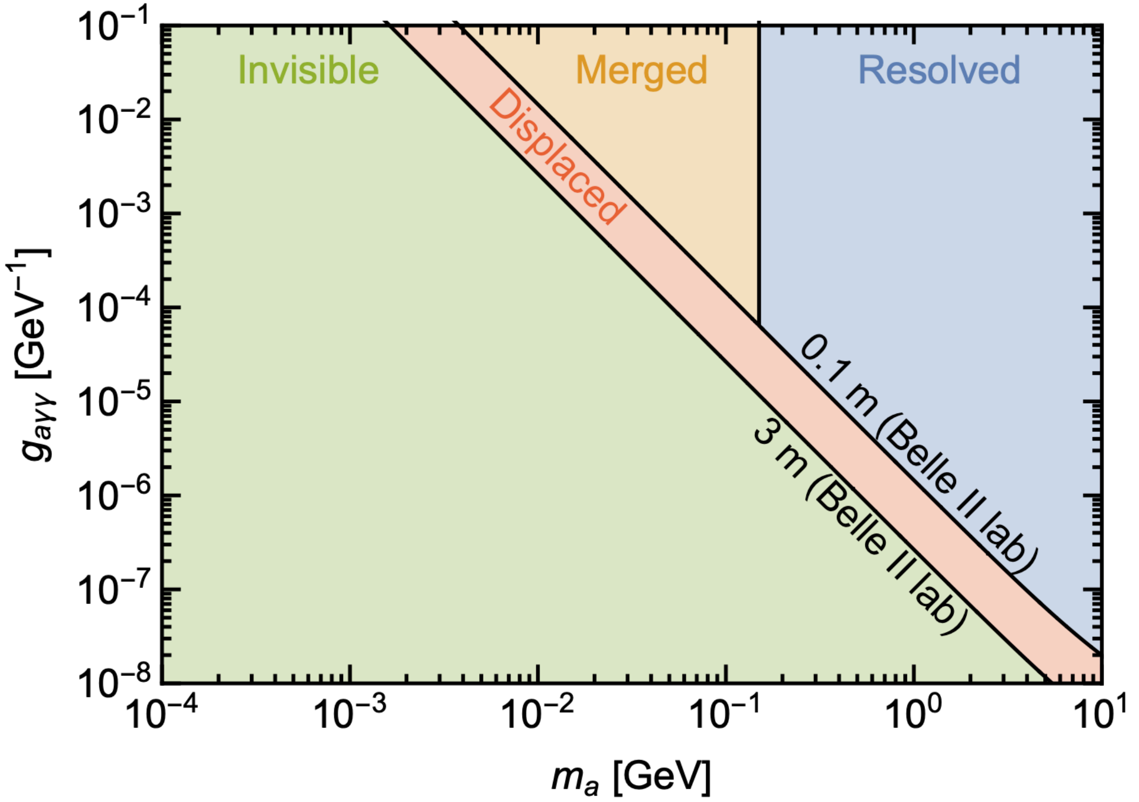


Dolan et al., J. High Energy Phys. 1712 (2017) 094

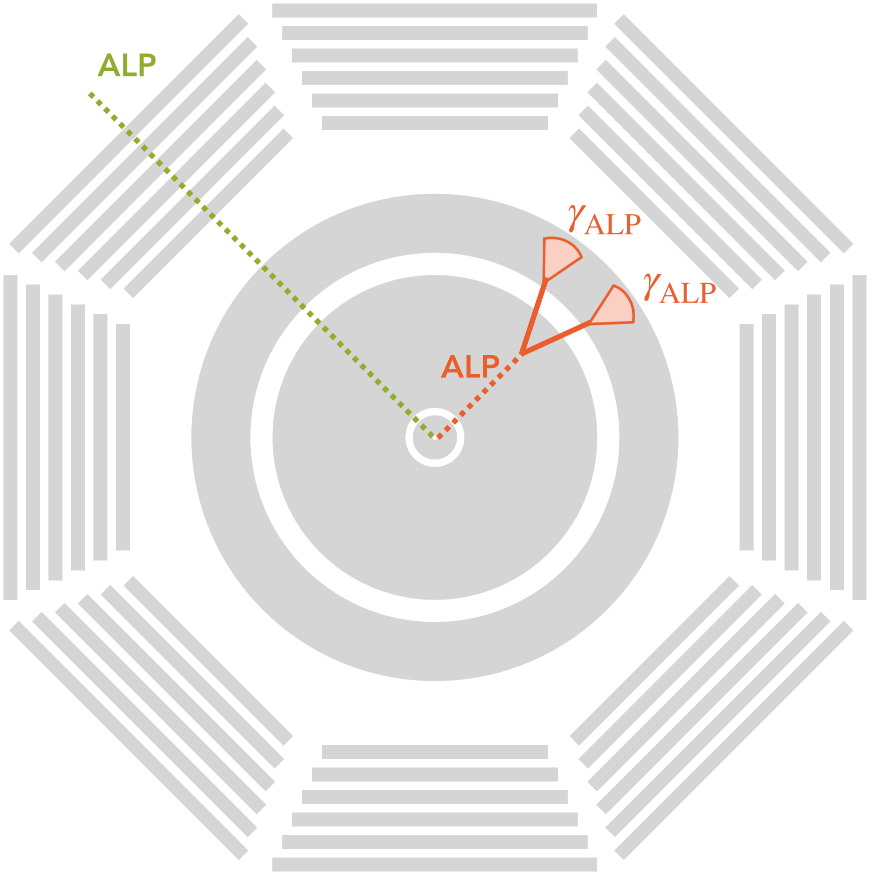


ALP Decay Outside of Detector
→ Invisible Decay

ALP-Strahlung decay at Belle II



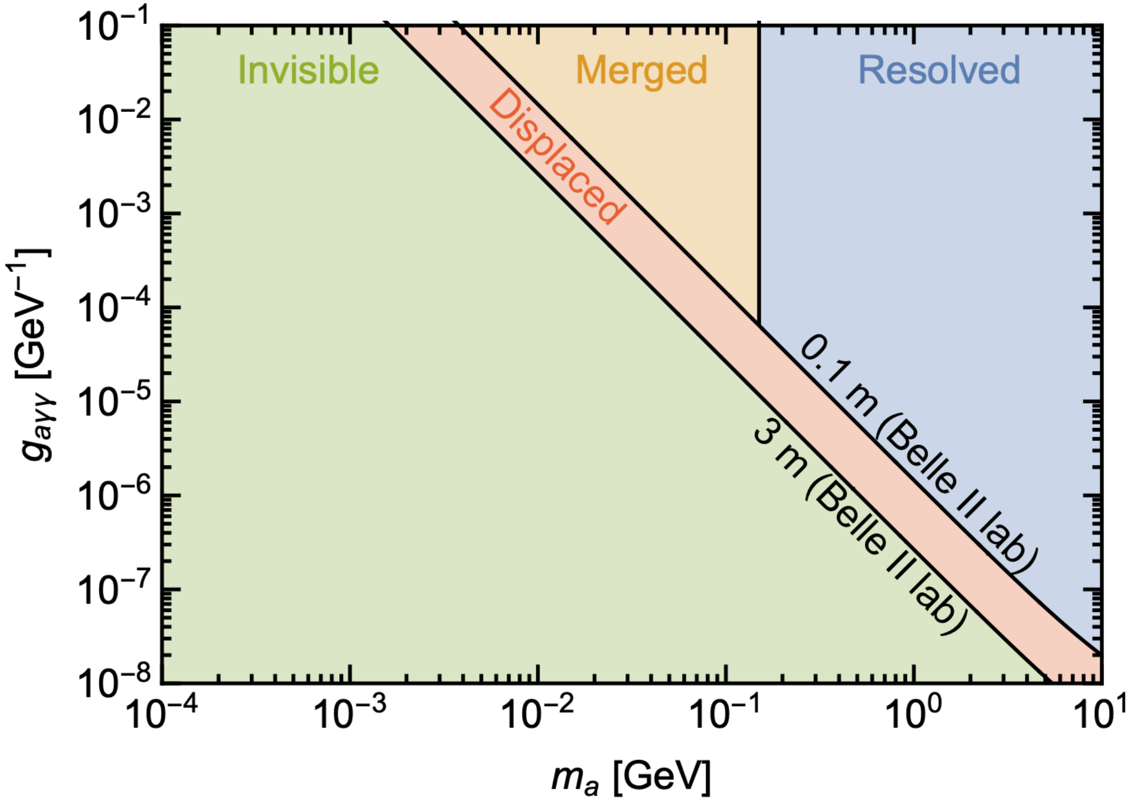
Dolan et al., J. High Energy Phys. 1712 (2017) 094



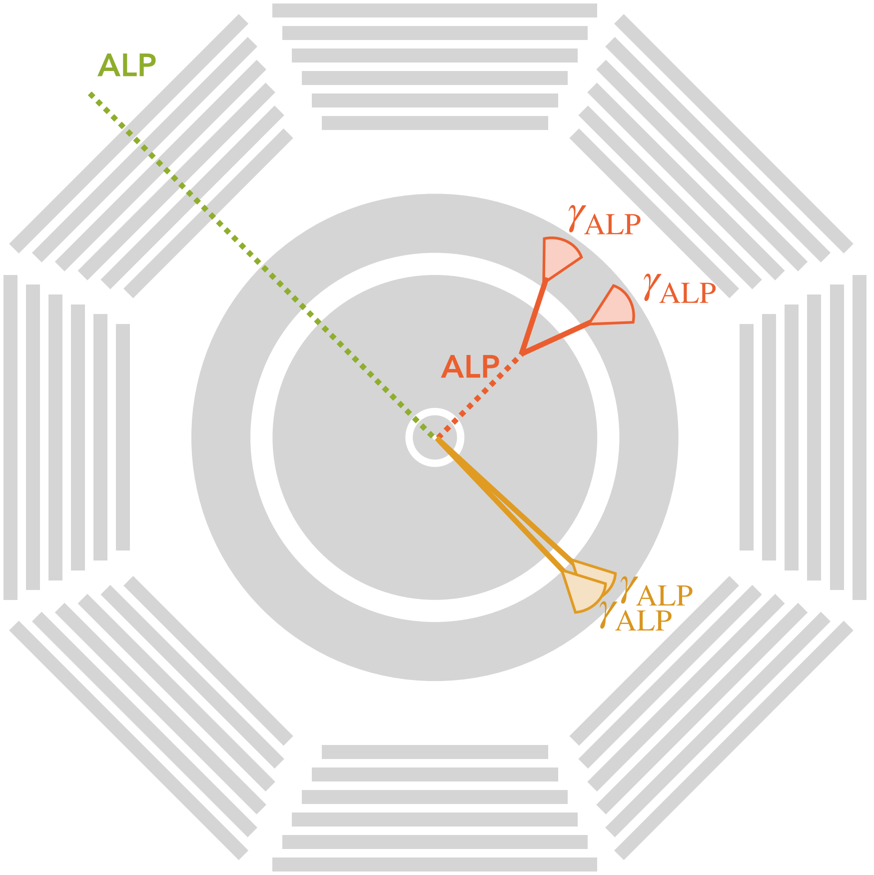
ALP Decay Outside of Detector
→ Invisible Decay

Large ALP Lifetime
→ 1-2 Photons with Incorrect Direction Reconstruction

ALP-Strahlung decay at Belle II



Dolan et al., J. High Energy Phys. 1712 (2017) 094

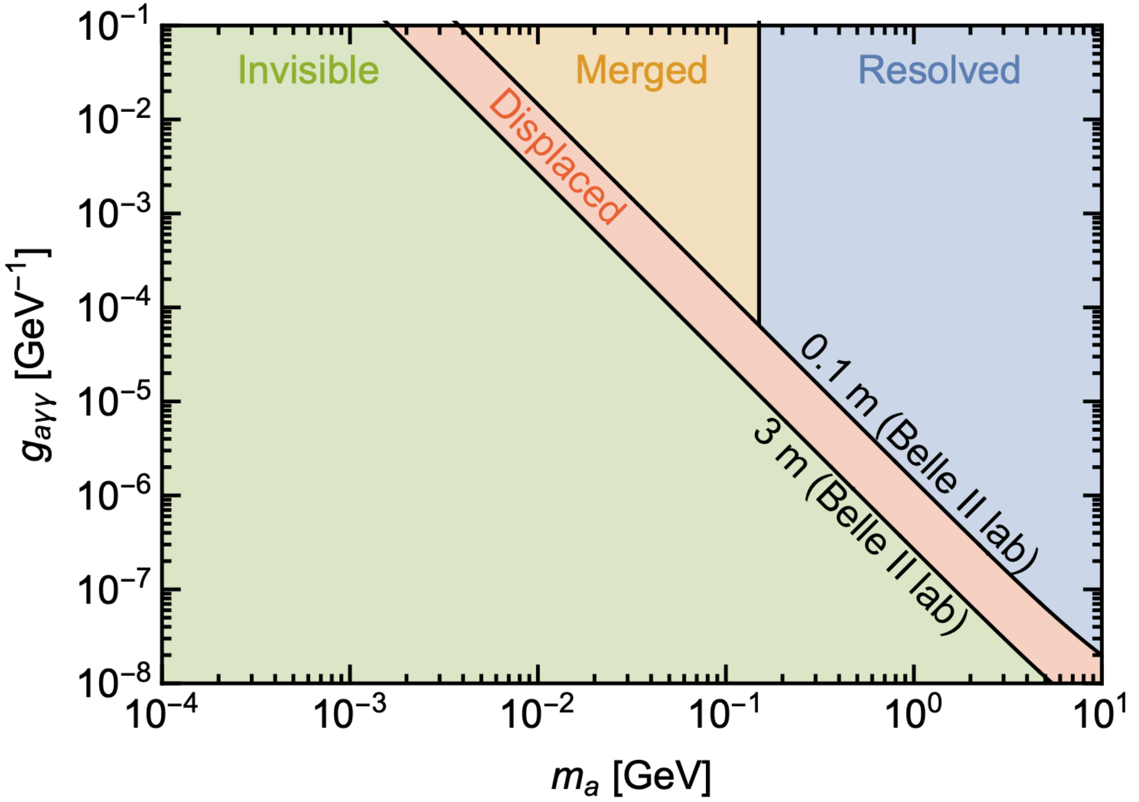


ALP Decay Outside of Detector
→ Invisible Decay

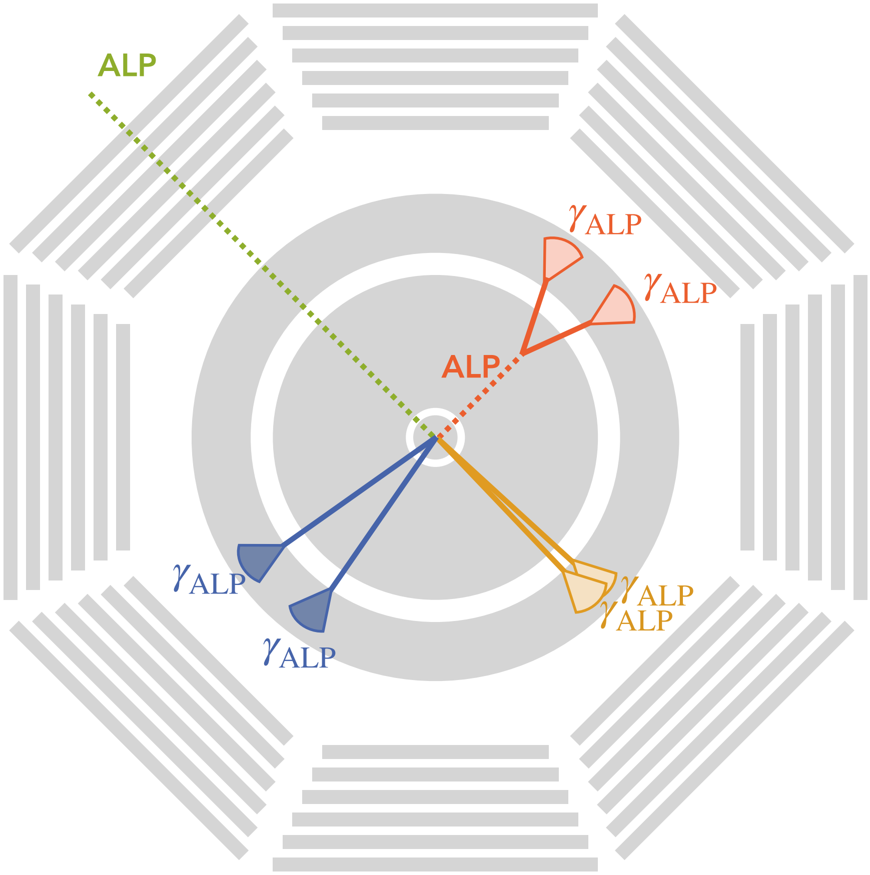
Large ALP Lifetime
→ 1-2 Photons with Incorrect Direction Reconstruction

Light ALP
Heavily Boosted
→ Single cluster of Merged Photons

ALP-Strahlung decay at Belle II



Dolan et al., J. High Energy Phys. 1712 (2017) 094



ALP Decay Outside of Detector
→ Invisible Decay

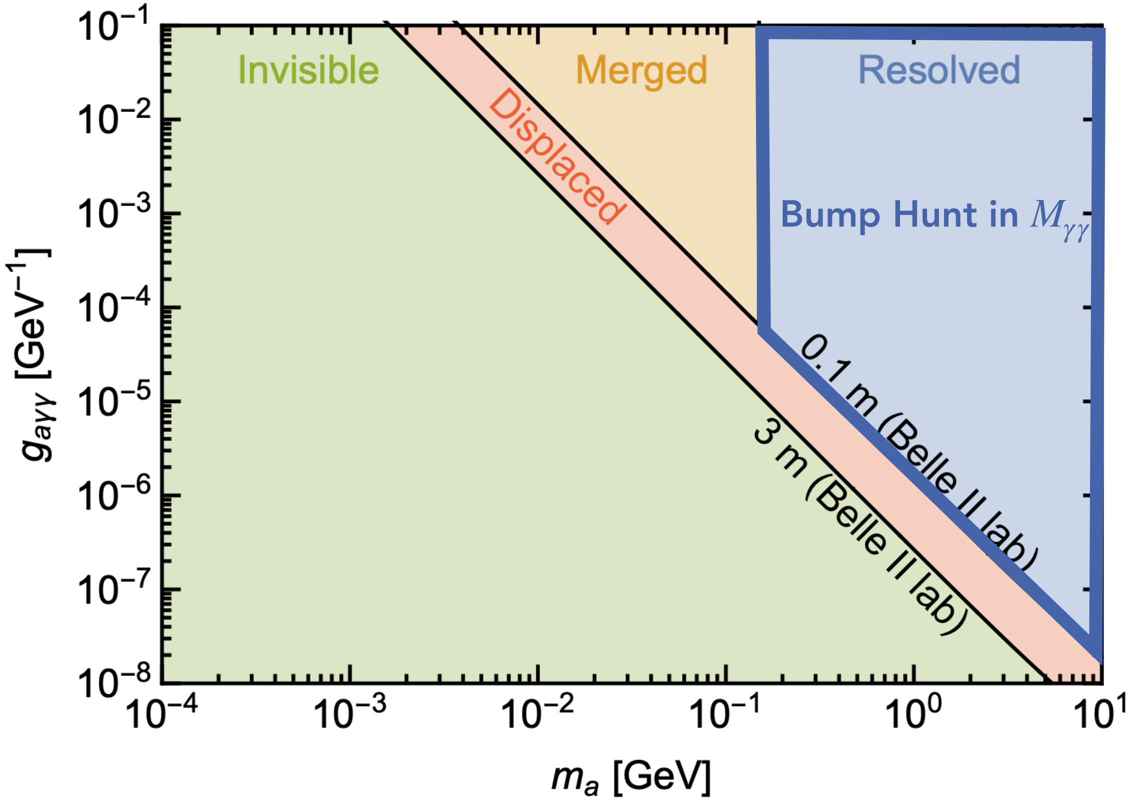
Large ALP Lifetime
→ 1-2 Photons with Incorrect Direction Reconstruction

Light ALP
Heavily Boosted
→ Single cluster of Merged Photons

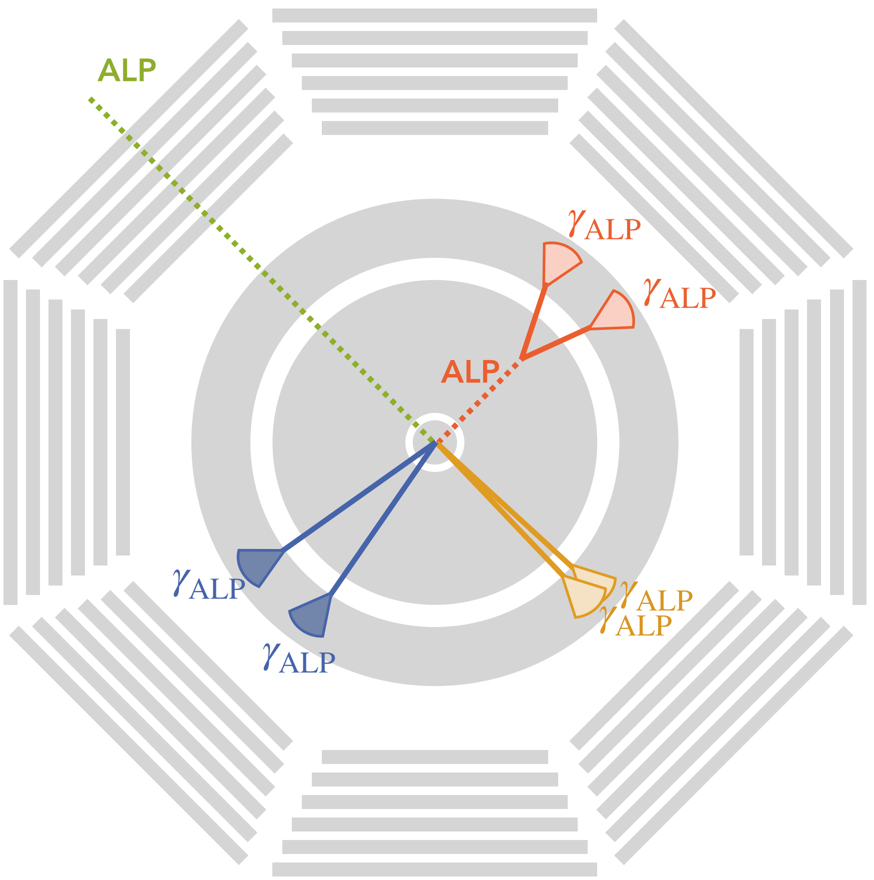
Heavier ALP
Decaying Promptly
→ 2 resolved Photons

ALP-Strahlung decay at Belle II

Institute of Experimental Particle Physics



Dolan et al., J. High Energy Phys. 1712 (2017) 094



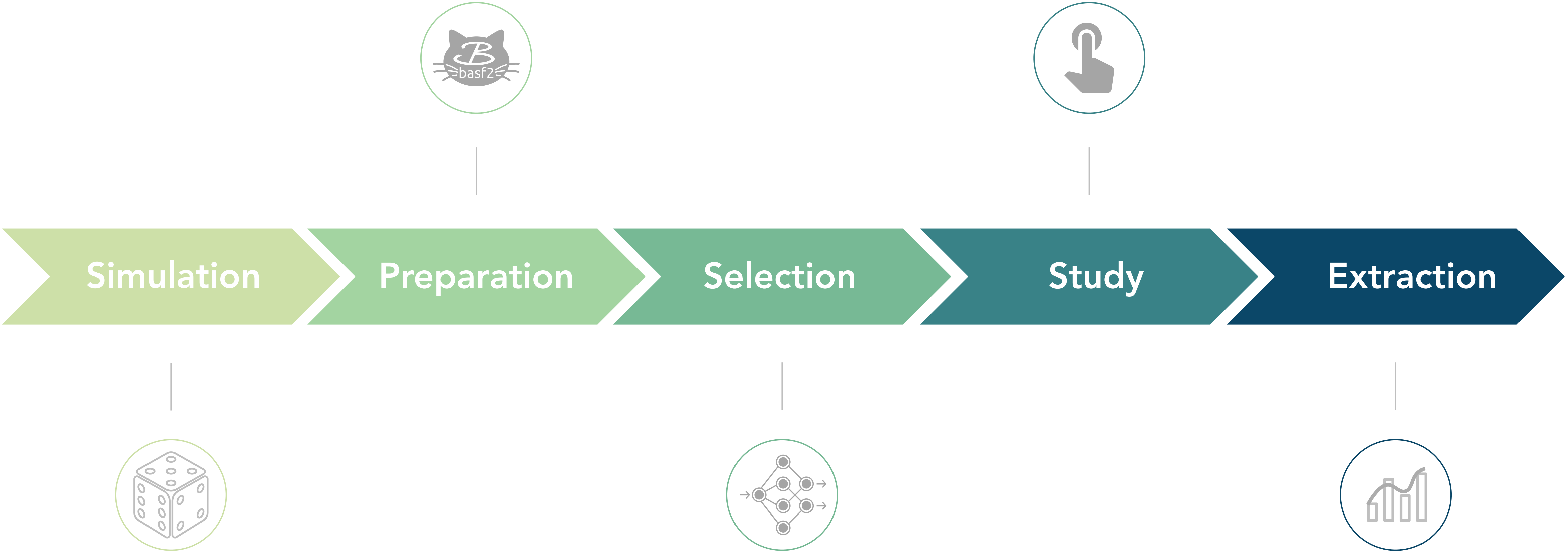
ALP Decay Outside of Detector
→ Invisible Decay

Large ALP Lifetime
→ 1-2 Photons with Incorrect Direction Reconstruction

Light ALP
Heavily Boosted
→ Single cluster of Merged Photons

Heavier ALP
Decaying Promptly
→ 2 resolved Photons

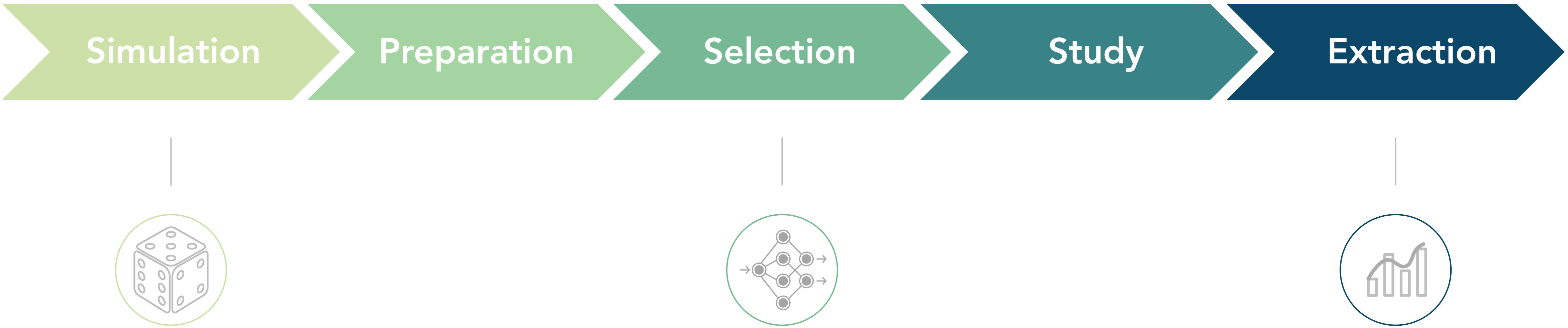
Analysis Strategy



Analysis Strategy

Produce run dependent MC

- **MadGraph5** aMC@NLO for signal with different m_a
- **BABAYAGA.NLO** and **PHOKHARA** for background



Analysis Strategy

Institute of Experimental Particle Physics

Produce run dependent MC

- **MadGraph5** aMC@NLO for signal with different m_a
- **BABAYAGA.NLO** and **PHOKHARA** for background

Simulation → Preparation → Selection → Study → Extraction

Reconstruction

- **Pre-selection** on events
- **Kinematic constraint fit** of three photons to the beam energy
  - **Need**: Photon resolution

Analysis Strategy

Institute of Experimental Particle Physics

Produce run dependent MC

- **MadGraph5** aMC@NLO for signal with different m_a
- **BABAYAGA.NLO** and **PHOKHARA** for background

MVA Based Selection

- Train so-called **Punzi Net**
- Suppress background
- Have **one selection**
  - For all signal masses m_a
  - Maximise sensitivity

Simulation → Preparation → Selection → Study → Extraction

Reconstruction

- **Pre-selection** on events
- **Kinematic constraint fit** of three photons to the beam energy
  - **Need**: Photon resolution

Analysis Strategy

ETP
Institute of Experimental Particle Physics

Simulation → Preparation → Selection → Study → Extraction

Produce run dependent MC
- **MadGraph5** aMC@NLO for signal with different m_a
- **BABAYAGA.NLO** and **PHOKHARA** for background

Reconstruction
- **Pre-selection** on events
- **Kinematic constraint fit** of three photons to the beam energy
  - **Need**: Photon resolution

MVA Based Selection
- Train so-called **Punzi Net**
- Suppress background
- Have **one selection**
  - For all signal masses m_a
  - Maximise sensitivity

Fit Preparation
- Calculate final **signal efficiency**
- Calculate **trigger efficiency**
- Study **MC/data agreement**
- Implement data-driven **corrections** and selections

Analysis Strategy

Institute of Experimental Particle Physics

Simulation → Preparation → Selection → Study → Extraction

Simulation: Produce run dependent MC
- **MadGraph5** aMC@NLO for signal with different m_a
- **BABAYAGA.NLO** and **PHOKHARA** for background

Preparation: Reconstruction
- **Pre-selection** on events
- **Kinematic constraint fit** of three photons to the beam energy
  - **Need**: Photon resolution

Selection: MVA Based Selection
- Train so-called **Punzi Net**
- Suppress background
- Have **one selection**
  - For all signal masses m_a
  - Maximise sensitivity

Study: Fit Preparation
- Calculate final **signal efficiency**
- Calculate **trigger efficiency**
- Study **MC/data agreement**
- Implement data-driven **corrections** and selections

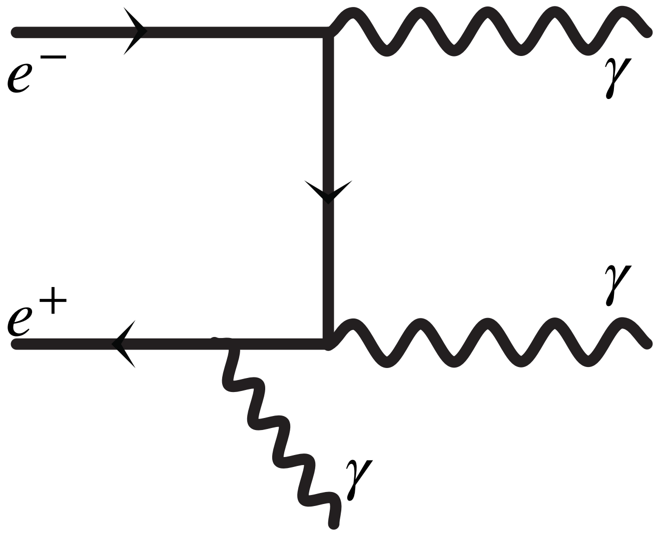
Extraction: Fit Strategy
- Perform **Bump Hunt** in $M_{\gamma\gamma} \in [0.2, 9.7]\,\mathrm{GeV/c^2}$
- Include systematic uncertainties
- Calculate **upper limit** in case of no significant deviation

Background Sources

Background Sources

$e^+e^- \to \gamma\gamma\gamma$

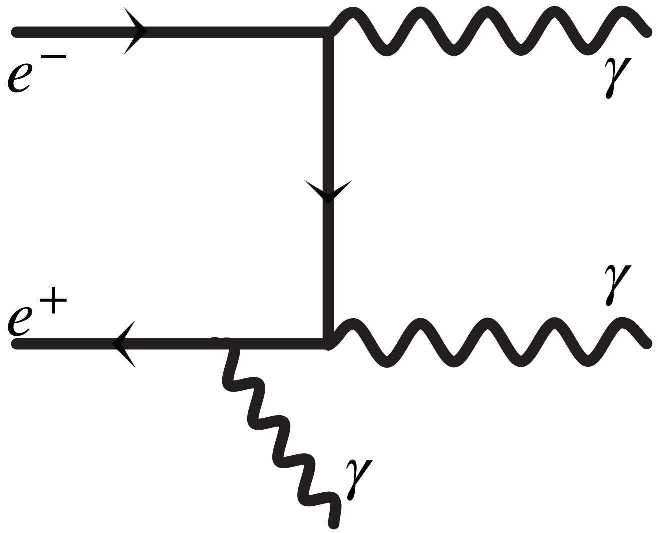
- Most dominant background
- Approximately constant distribution in $M_{\gamma\gamma}$



Background Sources

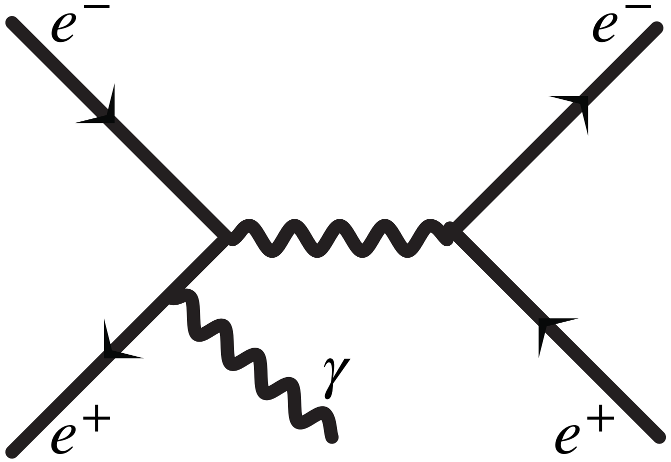
$e^+e^- \to \gamma\gamma\gamma$

- Most dominant background
- Approximately constant distribution in $M_{\gamma\gamma}$



$e^+e^- \to e^+e^-\gamma$

- Most common process
- Need to miss both tracks and reconstructed ECL clusters to be counted as background

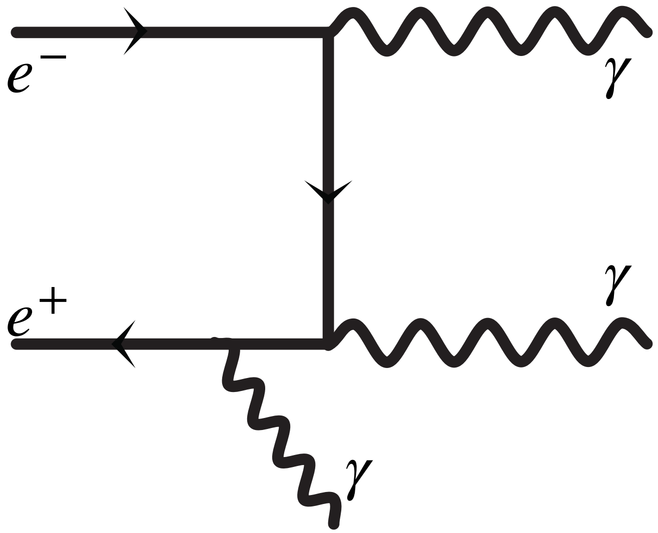


Background Sources

Simulation › Preparation › Selection › Study › Extraction

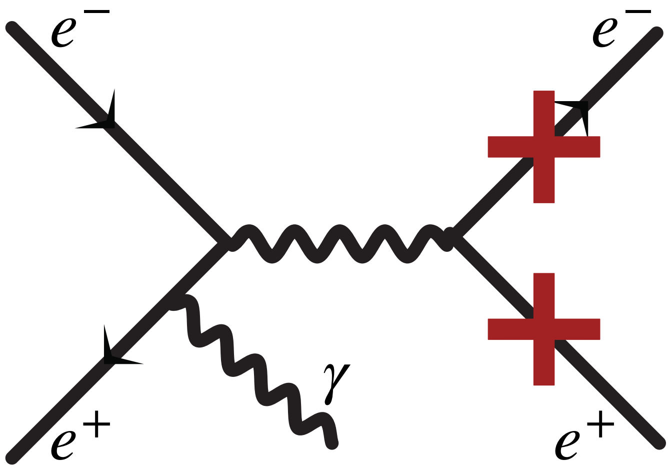
$e^+e^- \to \gamma\gamma\gamma$

- Most dominant background
- Approximately constant distribution in $M_{\gamma\gamma}$



$e^+e^- \to e^+e^-\gamma$

- Most common process
- Need to miss both tracks and reconstructed ECL clusters to be counted as background

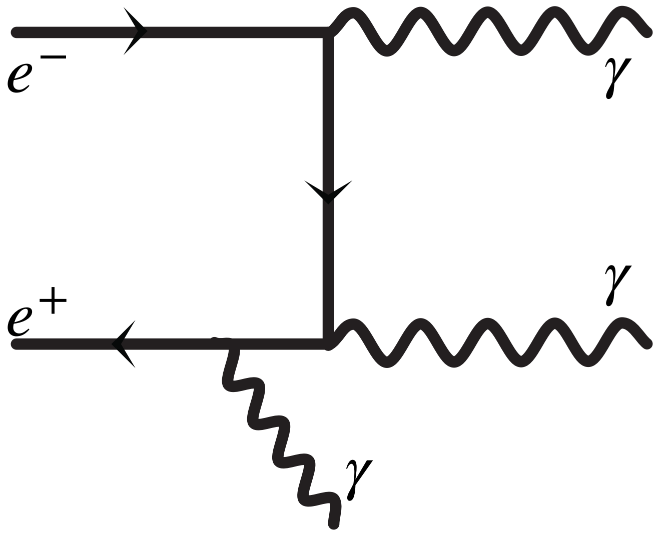


Background Sources

Simulation | Preparation | Selection | Study | Extraction

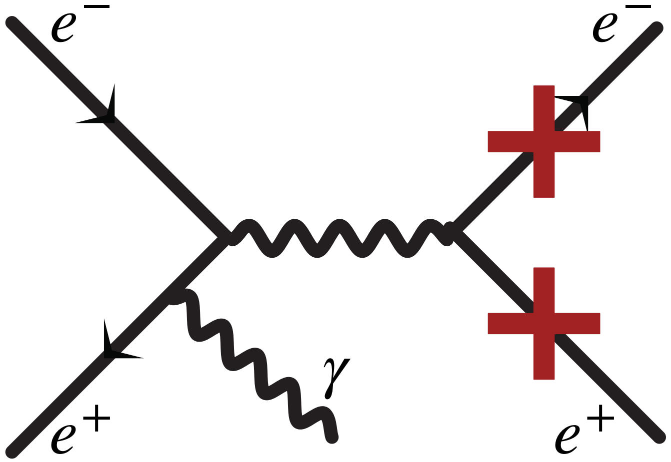
$e^+e^- \to \gamma\gamma\gamma$

- Most dominant background
- Approximately constant distribution in $M_{\gamma\gamma}$



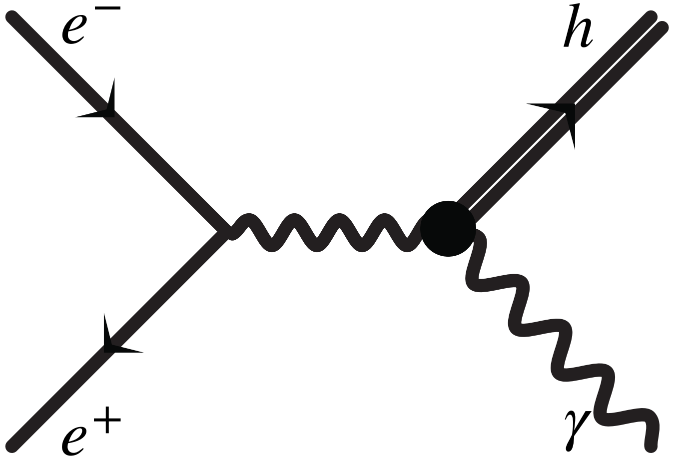
$e^+e^- \to e^+e^-\gamma$

- Most common process
- Need to miss both tracks and reconstructed ECL clusters to be counted as background



$e^+e^- \to h\gamma(\gamma)$

- $h = \pi^0, \eta, \eta', \ldots$
- Irreducible background in $M_{\gamma\gamma}$
- Additional source through next order process $ee \to h\gamma\gamma$

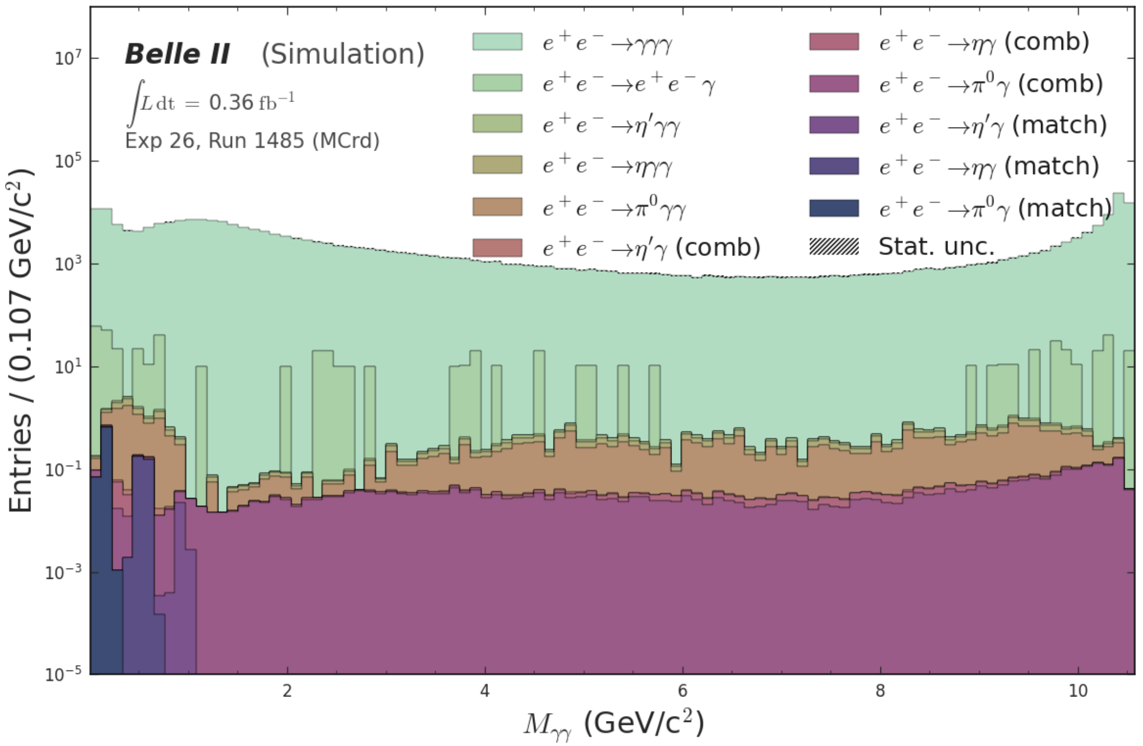


Background Distribution

Simulation | Preparation | Selection | Study | Extraction

Pre-Selection

- *Photon Candidates:*
  - $E_\gamma > 100\,\mathrm{MeV}$
  - θ_γ in CDC acceptance
  - |Cluster Timing|<200ns
  - Choose the three highest energetic candidates
    - The multiplicity of each event is three
- *Event:*
  - $0.8 \cdot E_{\mathrm{cms}} < M_{\gamma\gamma\gamma} < 1.05 \cdot E_{\mathrm{cms}}$
  - Clusters in an event should occur simultaneously
  - Absence of a good track†



†good track $\hat{=}\, \mathrm{d}r < 1$ cm, $|\mathrm{d}z| < 3$ cm, θ in CDC acceptance

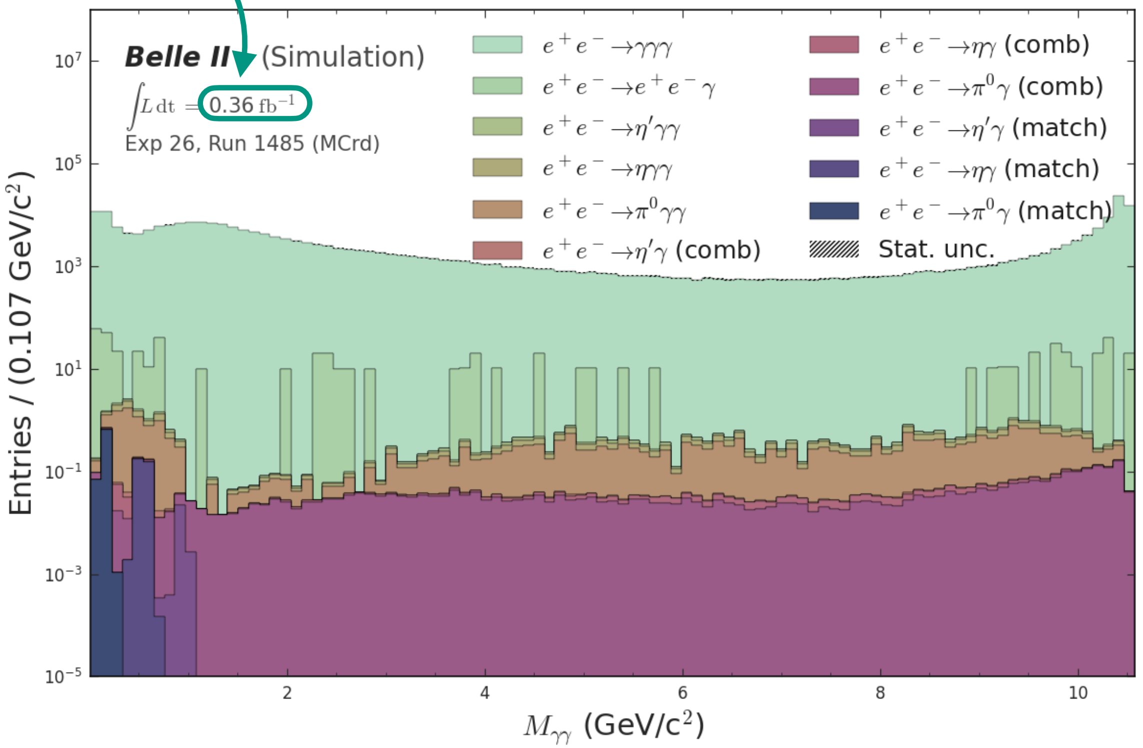
Background Distribution

Simulation | Preparation | Selection | Study | Extraction

Pre-Selection

- *Photon Candidates:*
  - $E_\gamma > 100\,\mathrm{MeV}$
  - θ_γ in CDC acceptance
  - |Cluster Timing|<200ns
  - Choose the three highest energetic candidates
    - The multiplicity of each event is three
- *Event:*
  - $0.8 \cdot E_{\mathrm{cms}} < M_{\gamma\gamma\gamma} < 1.05 \cdot E_{\mathrm{cms}}$
  - Clusters in an event should occur simultaneously
  - Absence of a good track†

About the size of the 2018 dataset

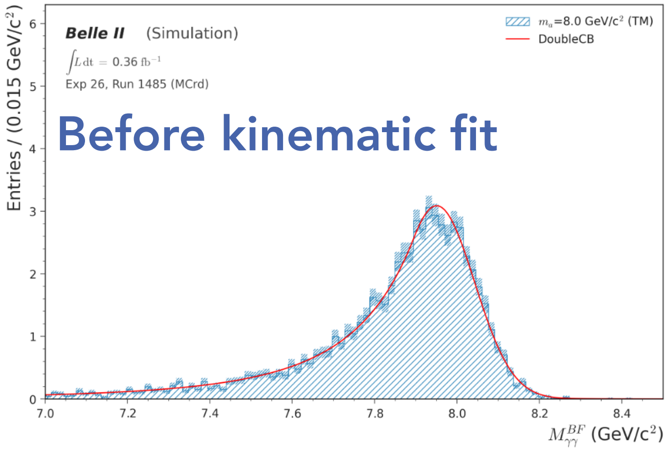


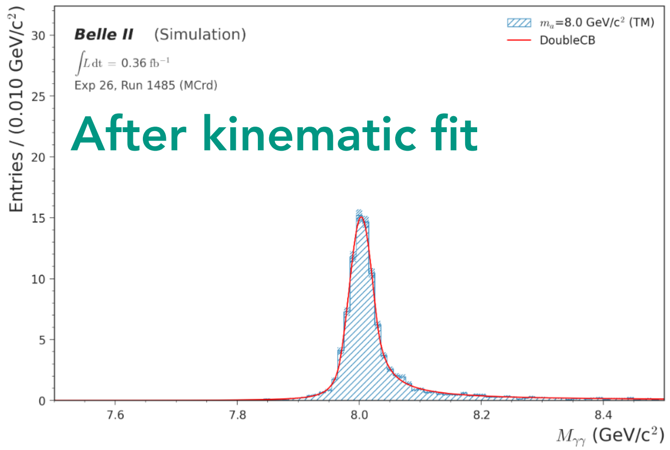
†good track $\hat{=}\,\mathrm{d}r < 1$ cm, $|\mathrm{d}z| < 3$ cm, θ in CDC acceptance

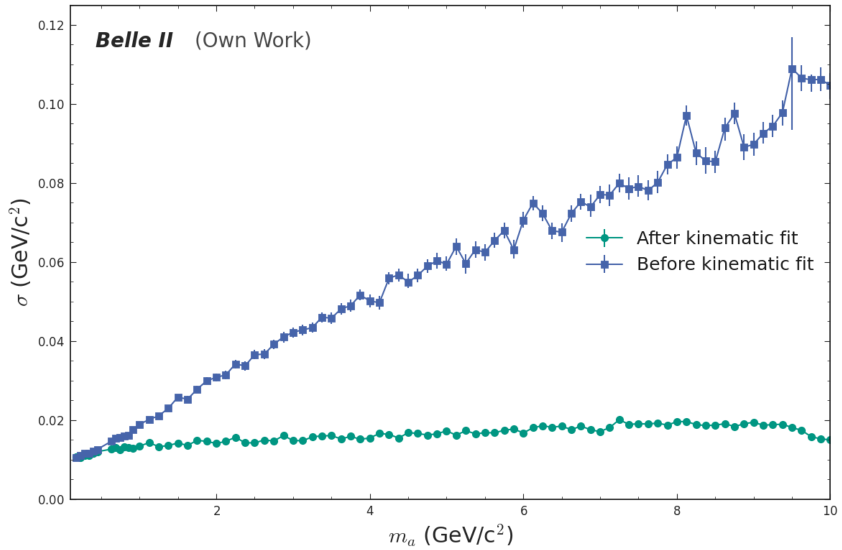
Belle II Group,
ETP, Karlsruhe Institute of Technology

Kinematic Fit

- **Challenge**:
  - Absolute photon energy resolution rises with higher photon energies
- **Solution**:
  - Constrain the final state momenta to a known quantity
  - Here: constrain to very well-known/measured beam properties





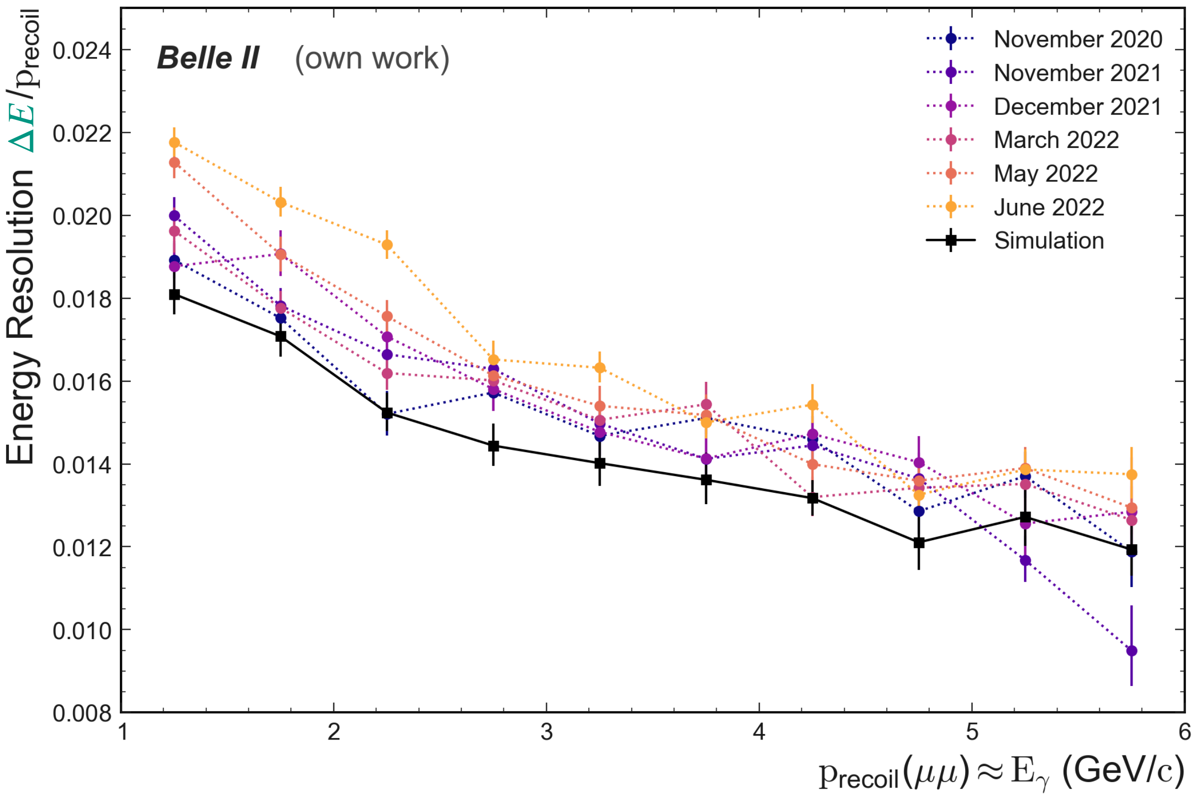


Photon Energy Resolution

- **Photon Covariance Matrix (PCM)**

$$\begin{pmatrix} \Delta E & \mathrm{cov}(E,\theta) & \mathrm{cov}(E,\phi) \\ * & \Delta\theta & \mathrm{cov}(\theta,\phi) \\ * & * & \Delta\phi \end{pmatrix}$$

  - Most relevant are the diagonals
  - Most impactful is the energy resolution for high photon energies
  - Resolution is dependent on detector conditions
- **Current state**:
  - Extract resolution from $e^+e^- \to \mu^+\mu^-\gamma$
  - Automated energy resolution calibration in the ECL barrel & $E_\gamma > 1$ GeV

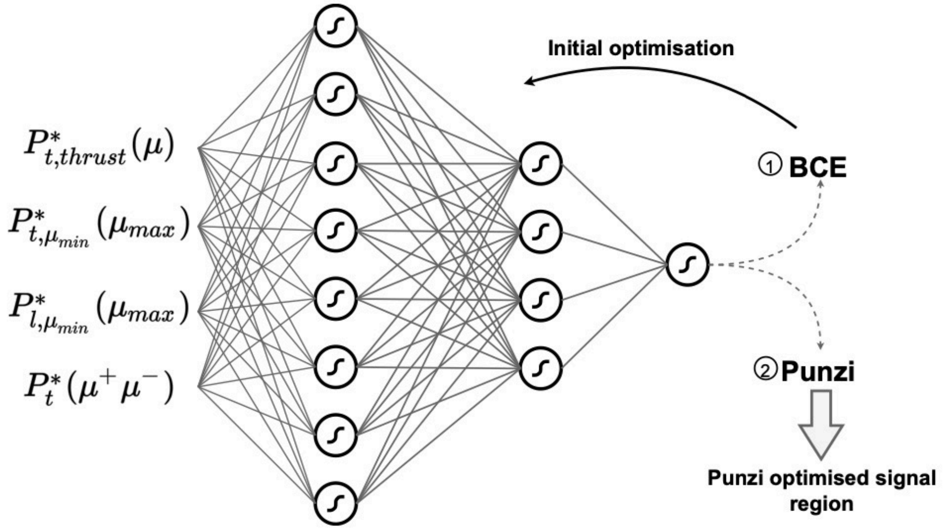


Belle II Group,
ETP, Karlsruhe Institute of Technology

Selection: Punzi Net

Simulation › Preparation › Selection › Study › Extraction

- **Punzi Net** (F. Abudinen et al., Eur. Phys. J. C 82 (2022) 121):
  - Uses minimal **detectable cross-section** as loss function (ϵ, B differentiable)
  - Feedforward network trained on input variables for different signal mass samples
  - **Generalises well** to mass hypotheses for which it was not trained
- Used at Belle II for the invisible Z' in $e^+e^- \to \mu^+\mu^- Z'$ analysis (Belle II, Phys. Rev. Lett. 130, 231801 (2023))
- Tested at KIT in BA thesis for $B^\pm \to K^\pm a, a \to \gamma\gamma$ sensitivity study



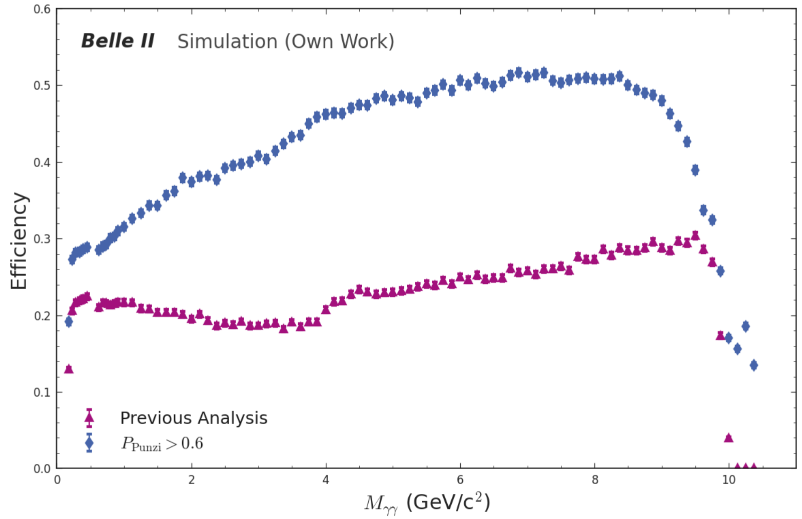
$$\sigma_{\min}(t) = \frac{\frac{b^2}{2}a\sqrt{B(t)} + \frac{b}{2}\sqrt{b^2 + 4a\sqrt{B(t)} + 4B(t)}}{\epsilon(t)L}$$

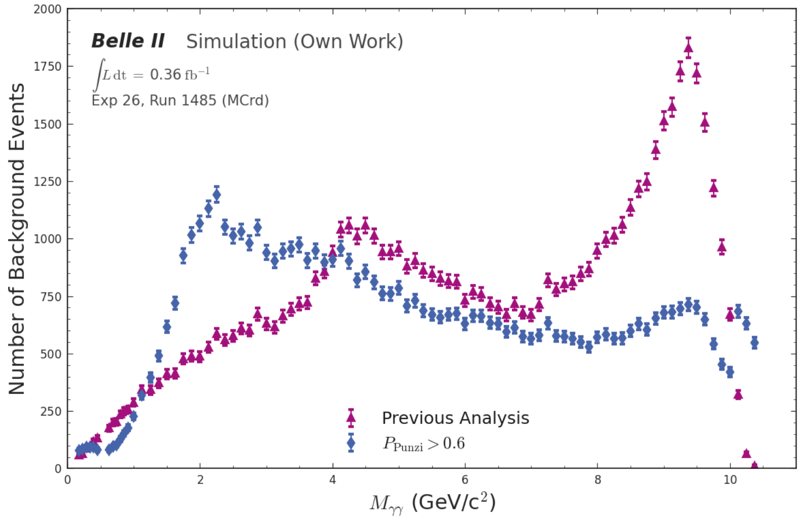
a, b : number of sigmas corresponding to Gaussian Test (5, 1.28)@(5σ, 90%CL)
ϵ: signal efficiency
B: number of background
L: luminosity

Selection: Result

- **Net input**:
  - *Optimisation Range*: $\pm 10\sigma$ (to be optimised)
  - *Features*: 9 selected features based on E_γ, θ_γ, and event shape variables
  - *Signal*:
    - $m_a \in [0.175, 10.375 : 0.05]$ GeV/c^2
      - Excluded 22 masses for generalisation tests
- **Hyperparameter** combination
  - *Architecture*: 10 layers, max depth 30, LeakyReLU
  - *Punzi Training*: 500 epochs, 20 000 batch size

Working point:
$P_{\text{Punzi}} > 0.6$
(to be optimised)





Signal Shape

- **Signal PDF:**
  - $N_{\mathrm{Sig}}\left(f\,\mathrm{DSCB} + (1\text{-}f)\ \mathrm{Poly}\right)$

Signal Shape

Simulation > Preparation > Selection > Study > Extraction

Signal PDF:

- $N_{\text{Sig}}\left(f\,\text{DSCB} + (1\text{-}f)\ \text{Poly}\right)$

Combined Signal Yield

- Multiply by f to get peaking yield

Signal Shape

Simulation › Preparation › Selection › Study › **Extraction**

Signal PDF:

- $N_{\mathrm{Sig}}\left(f\,\mathrm{DSCB} + (1\text{-}f)\ \mathrm{Poly}\right)$

Combined Signal Yield

- Multiply by f to get peaking yield

TM fraction

- Stabilises Combinatorial PDF parameter fit

Signal Shape

Signal PDF:

- $N_{\text{Sig}}\left(f\,\text{DSCB} + (1\text{-}f)\ \text{Poly}\right)$

Combined Signal Yield

- Multiply by f to get peaking yield

TM Signal PDF

- Double Sided Crystal Ball
- Both n parameters fixed
- Possible change: Generalized CB

TM fraction

- Stabilises Combinatorial PDF parameter fit

Signal Shape

Signal PDF:

- $N_{\text{Sig}}\left(f\,\text{DSCB} + (1\text{-}f)\ \text{Poly}\right)$

Combined Signal Yield

- Multiply by f to get peaking yield

TM Signal PDF

- Double Sided Crystal Ball
- Both n parameters fixed
- Possible change: Generalized CB

TM fraction

- Stabilises Combinatorial PDF parameter fit

Combinatorics PDF

- More or less constant contribution
- Chebyshev polynomial of first-order
- Lower order is not sufficient for some masses
- Higher orders add too many parameters

Signal Shape

Signal PDF:

- $N_{\text{Sig}}\left(f\,\text{DSCB} + (1\text{-}f)\ \text{Poly}\right)$

Combined Signal Yield

- Multiply by f to get peaking yield

TM Signal PDF

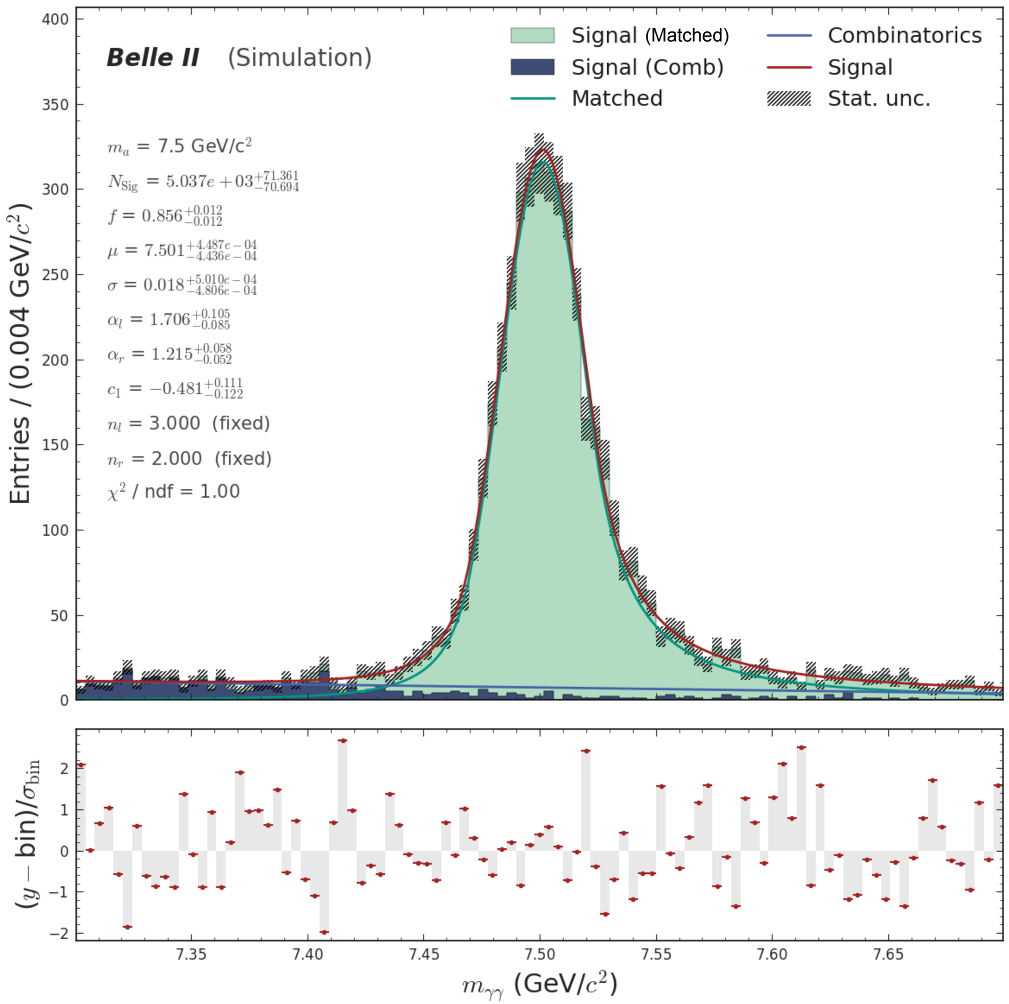
- Double Sided Crystal Ball
- Both n parameters fixed
- Possible change: Generalized CB

TM fraction

- Stabilises Combinatorial PDF parameter fit

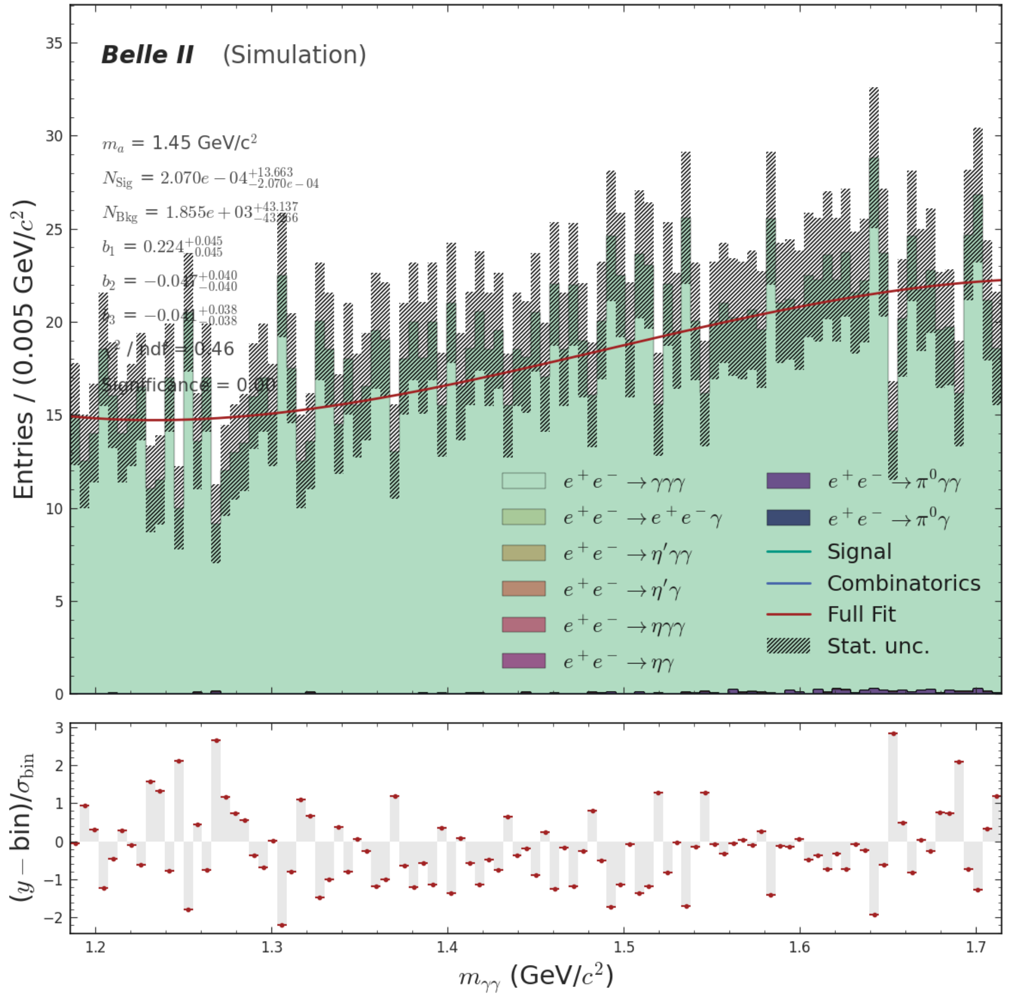
Combinatorics PDF

- More or less constant contribution
- Chebyshev polynomial of first-order
- Lower order is not sufficient for some masses
- Higher orders add too many parameters



Hypothesis Fit

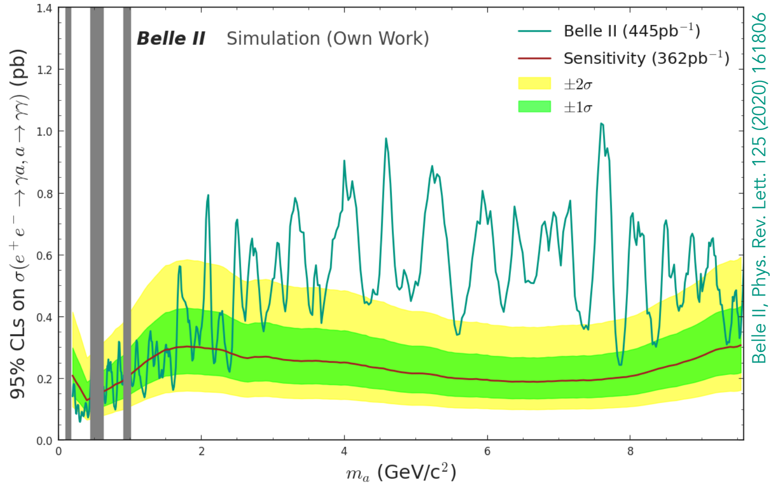
- **Total PDF**:
  - $N_{\text{Sig}}\left(f\text{Peak} + (1-f)\text{Comb}\right) + N_{\text{Bkg}}\text{Polynom}$
  - *Signal* PDF:
    - All shape parameters & f fixed to interpolation
  - *Background* PDF:
    - 3rd order polynomial
    - Polynomial parameter floating
      - Later: Fix them in side bands
- **Search Range**:
  - $m_a \in [0.175, 9.6 : 0.05]$ GeV
  - Excluded: π^0, η, η' mass regions

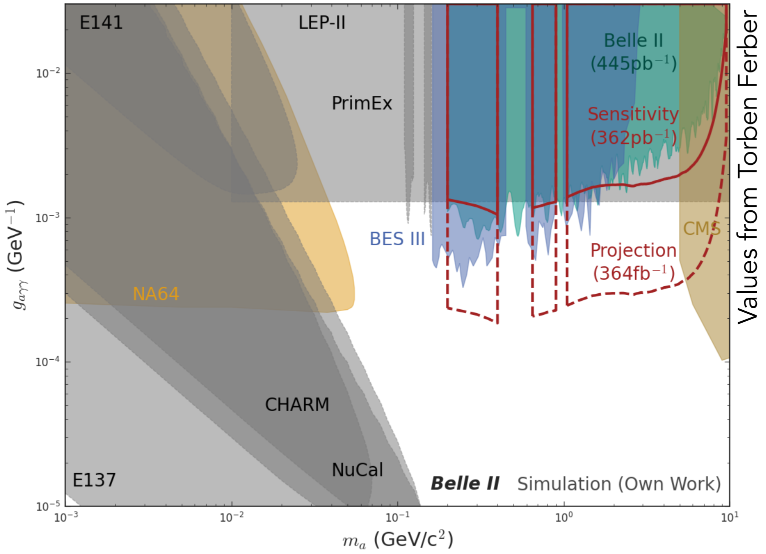


Search Sensitivity

Simulation | Preparation | Selection | Study | **Extraction**

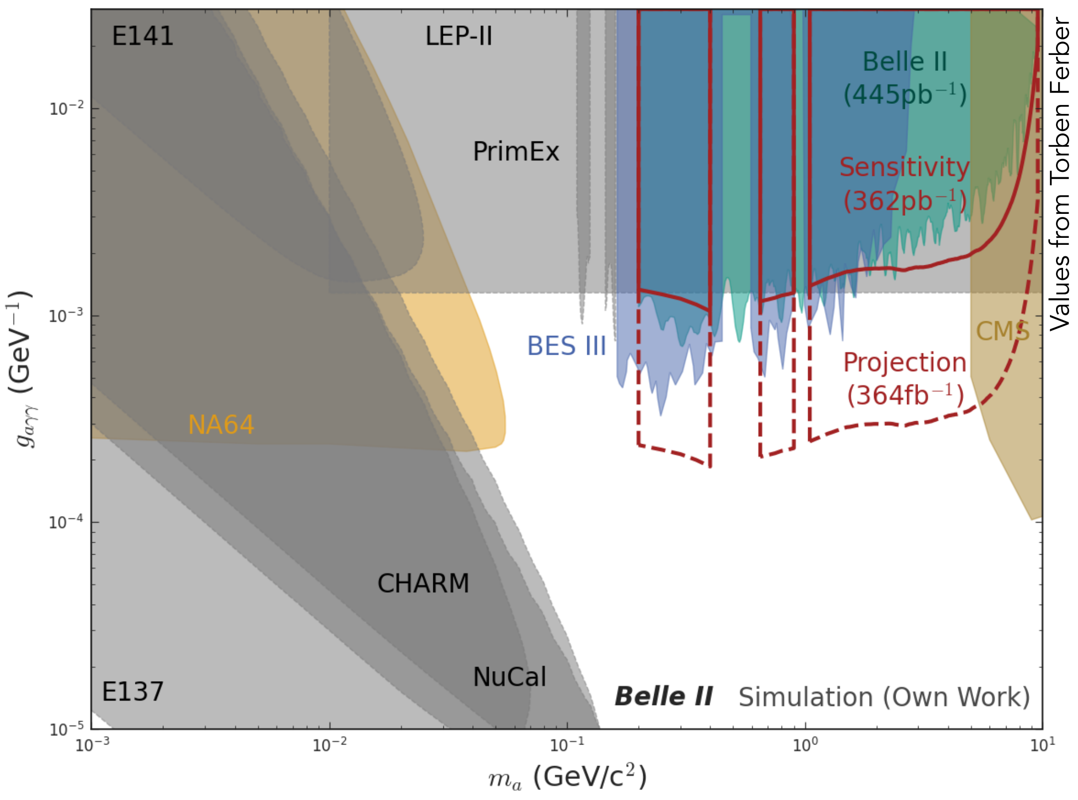
- **Upper Limit**
  - Asymptotic formulae @ 95% CLs
- Use the upper limit on the cross-section for **model-dependent coupling exclusion**
- **The plots show…**
  - Only the sensitivity for this analysis (MC only)
  - Not considered: trigger efficiency, systematic uncertainty
    - Expectation: Trigger Efficiency > 95% and negligible systematics





A full road ahead of us!

- **Analysis priority list**
  - Study of trigger efficiency
  - Extension to the full and automated photon resolution calibration
  - Run optimisation on selection and fitting range
- **We have a lot more data and understanding of Belle II since 2018**
  - **Goal**: find $a \to \gamma\gamma$ decay with increased sensitivity



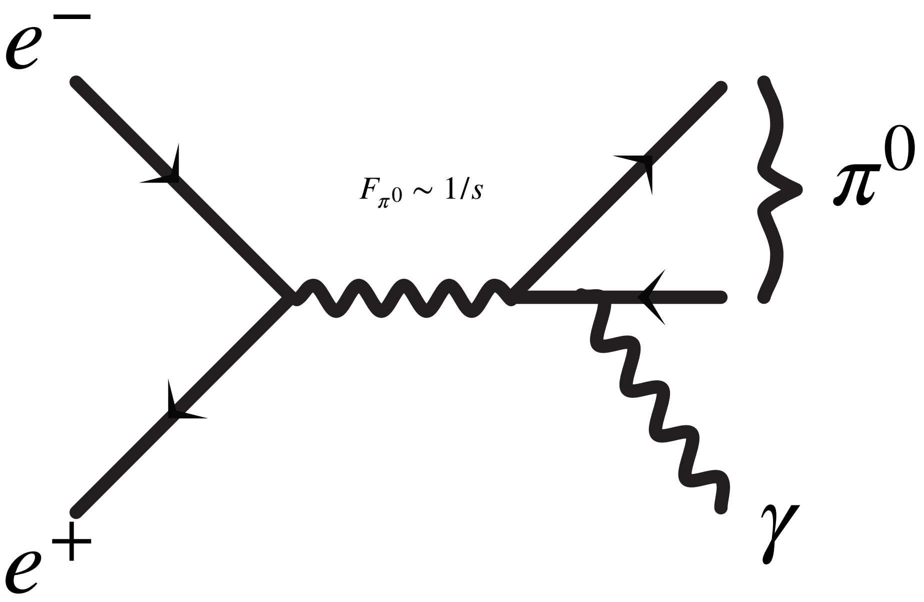
Backup

Irreducible Background Example: $e^+e^- \to \pi^0\gamma(\gamma)$

Institute of Experimental Particle Physics

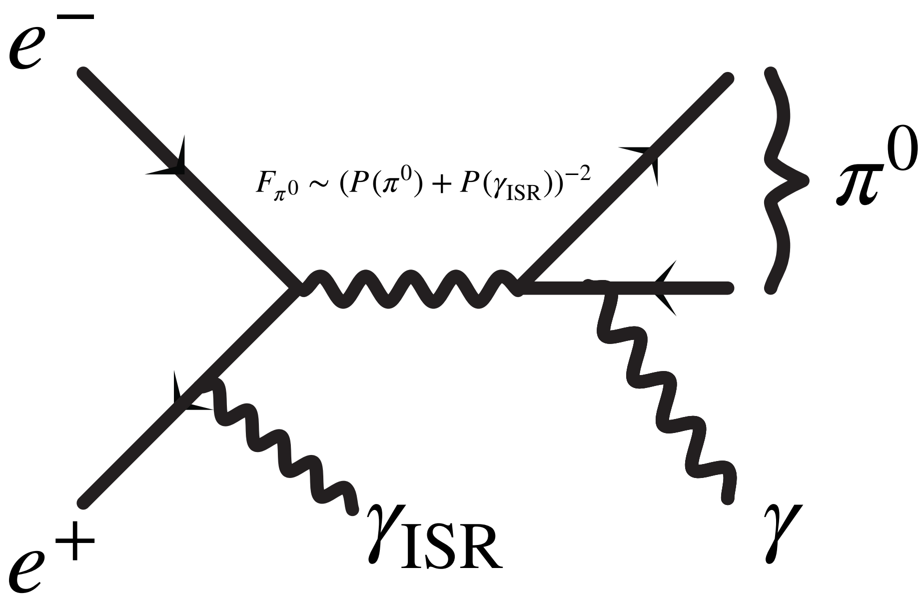
$e^+e^- \to \pi^0\gamma$

- Irreducible background at $M_{\gamma\gamma} \approx m_{\pi^0}$
- $\sigma^{\dagger}_{\text{eff}}(ee \to \pi^0\gamma) \approx 7.26 \cdot 10^{-6}\,\text{nb}$



$e^+e^- \to \pi^0\gamma\gamma$

- Missed a photon but $M_{\gamma\gamma\gamma}$ in selection range
- $\sigma^{\dagger}_{\text{eff}}(ee \to \pi^0\gamma\gamma) \approx 1.68 \cdot 10^{-3}\,\text{nb}$ ($\varepsilon \approx 1.3\,\%$)

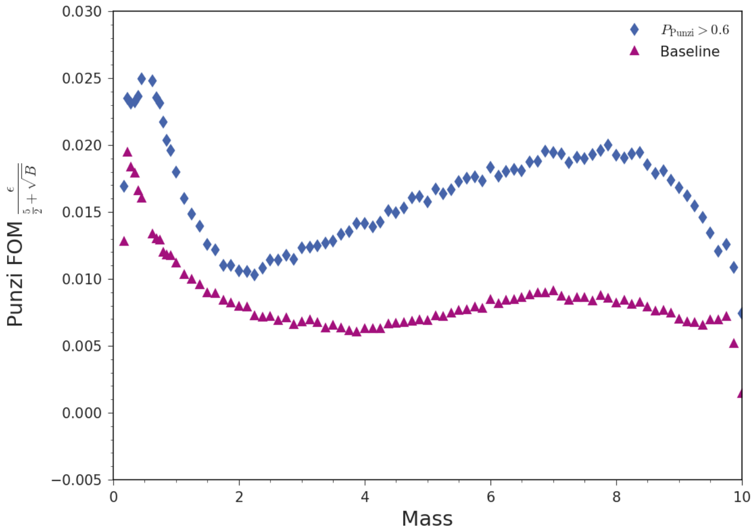


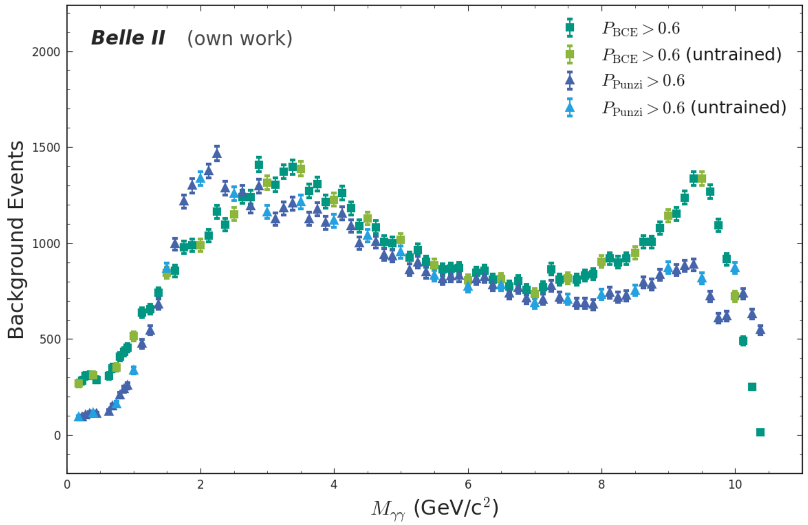
Hyperparameter "Optimisation"

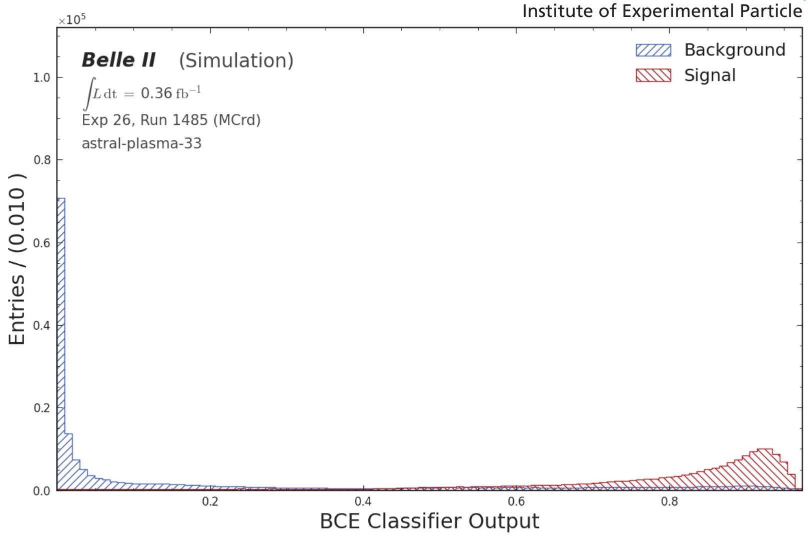
Institute of Experimental Particle Physics

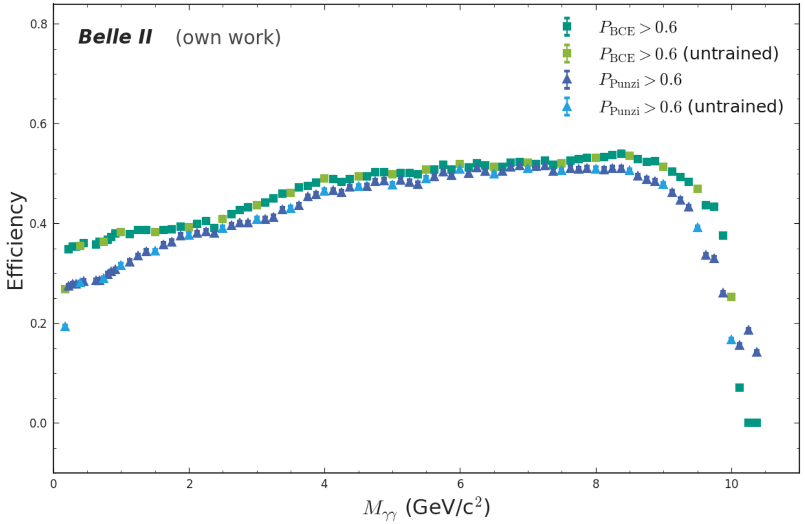
- **Grid search** of hyperparameters:
  - Layer depth: 4-256
  - Number of layers: 3-10
  - Activation functions: ELU, ReLU, LeakyReLU, Sigmoid, Tanh
  - Punzi batch size: $2^{11} - 2^{16}$
  - Punzi epochs: 200-5000
  - Fixed parameters:
    - BCE epochs: 200, BCE batch size: 2048
    - LR scheduler
- **Net input**:
  - "Most important" features
  - *Signal*:
    - $m_a \in [0.175, 10.375 : 0.05]$ GeV/c^2
      - Excluded 22 masses for generalisation
      - Only show 50% of signal masses
  - *Background*:
    - All background channels
    - Weight to luminosity

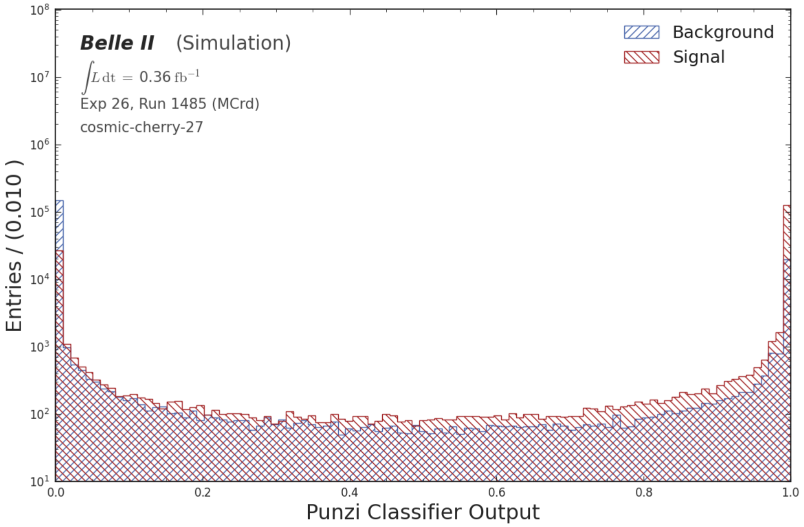
Training Plots

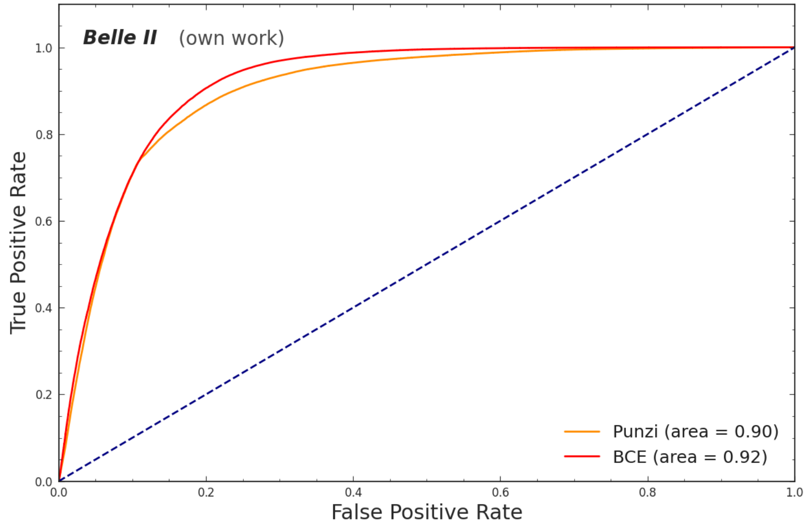








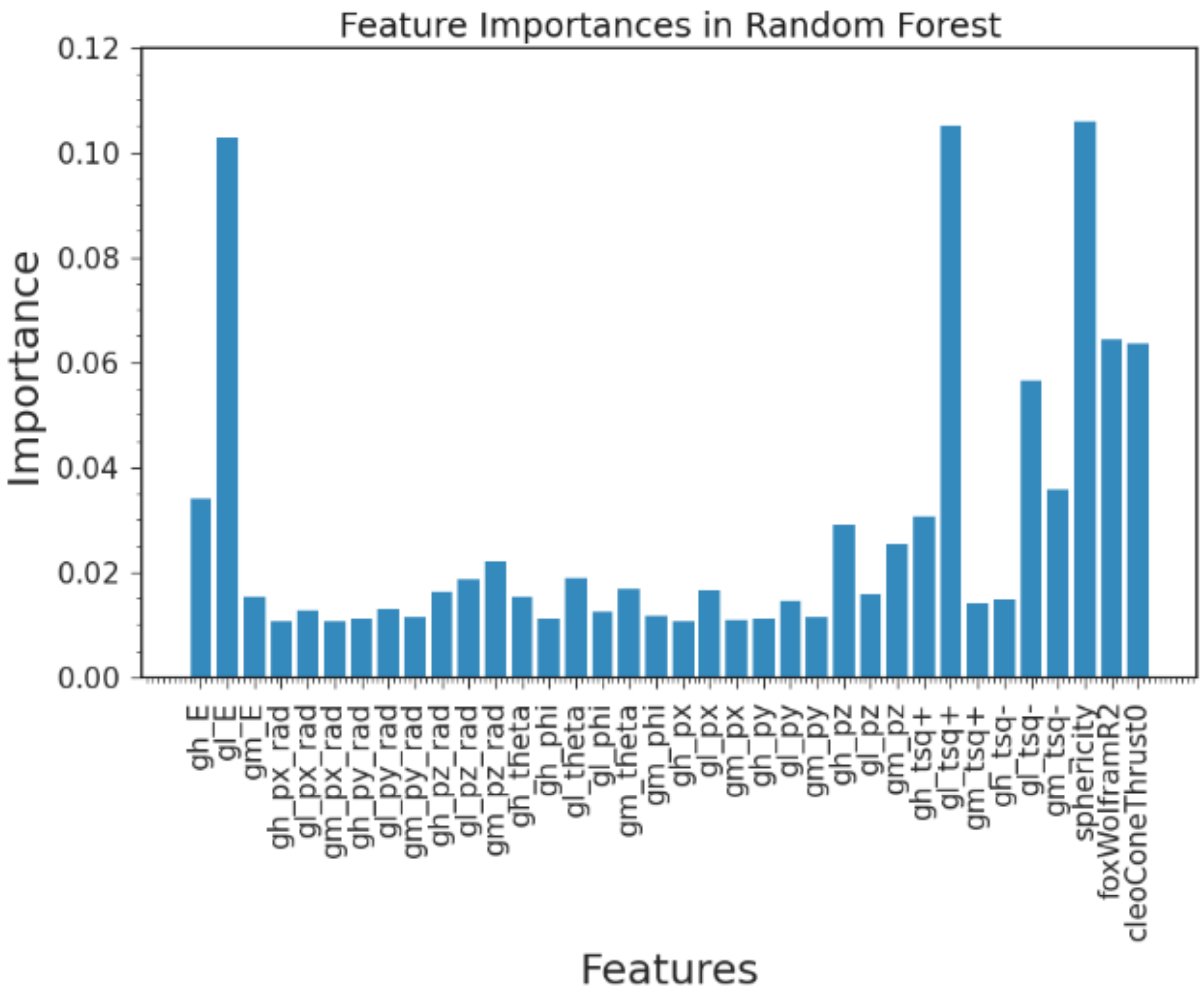




Feature Selection

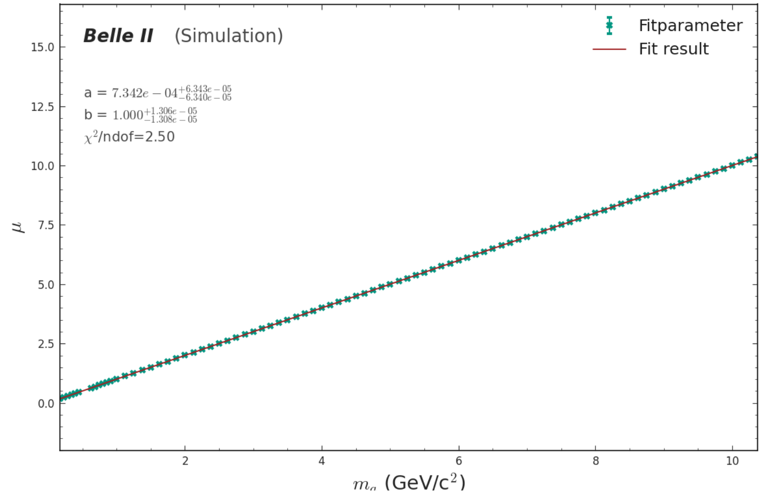
- **Random forest classifier** for feature importance estimation
- Features **used in selection**:
  - $E_{\gamma h}$, $E_{\gamma l}$
  - $u_{\gamma h}$, $u_{\gamma l}$, $t_{\gamma l}$, $t_{\gamma m}$
  - Sphericity, R_2, 0th Cleo cone with respect to the thrust axis

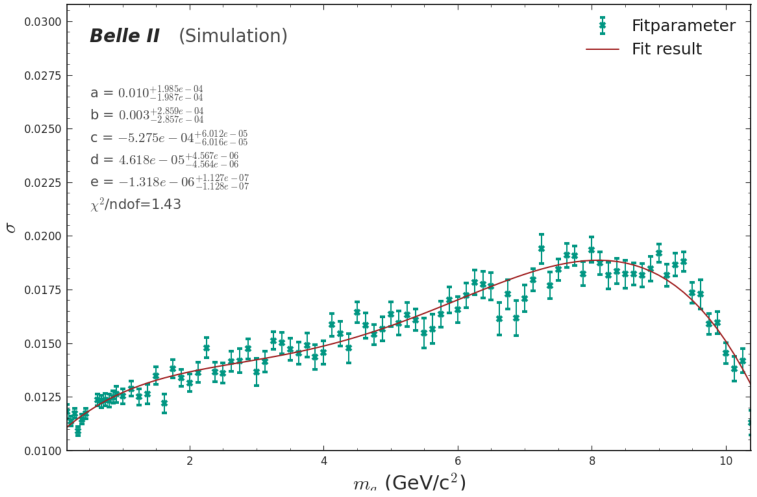
$$2u/t = E^2(1 \pm \cos\theta)$$

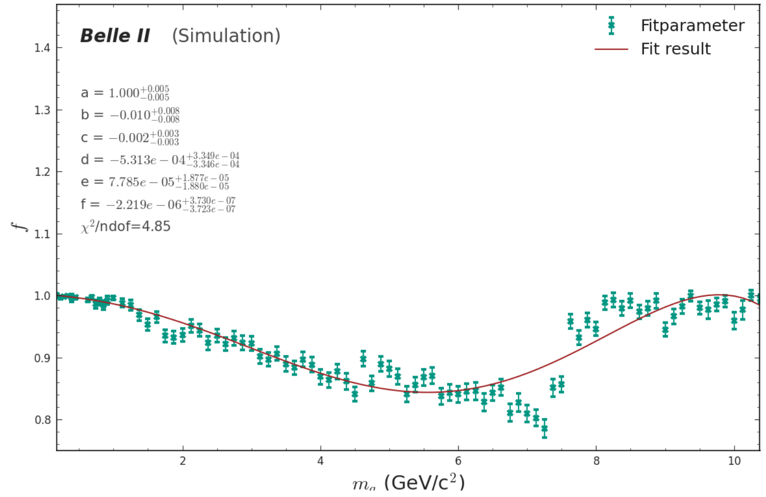


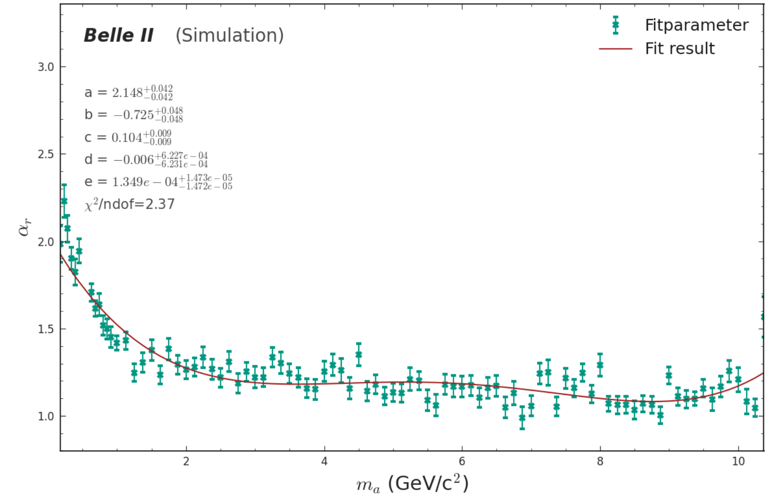
Signal Shape Interpolation

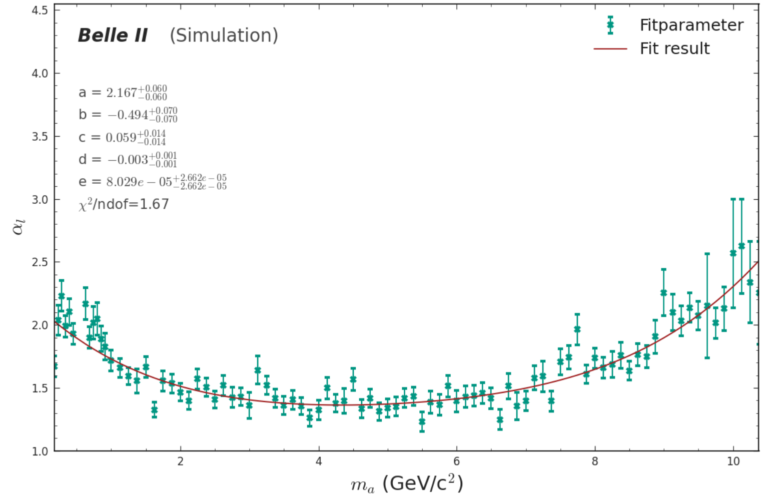
Institute of Experimental Particle Physics

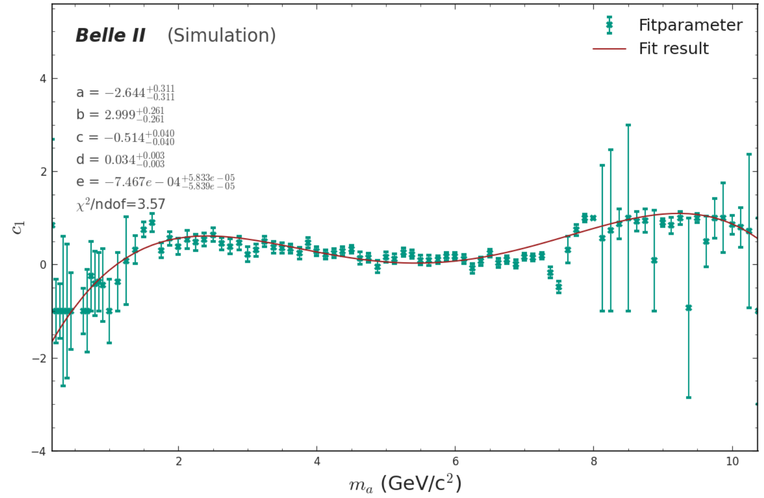






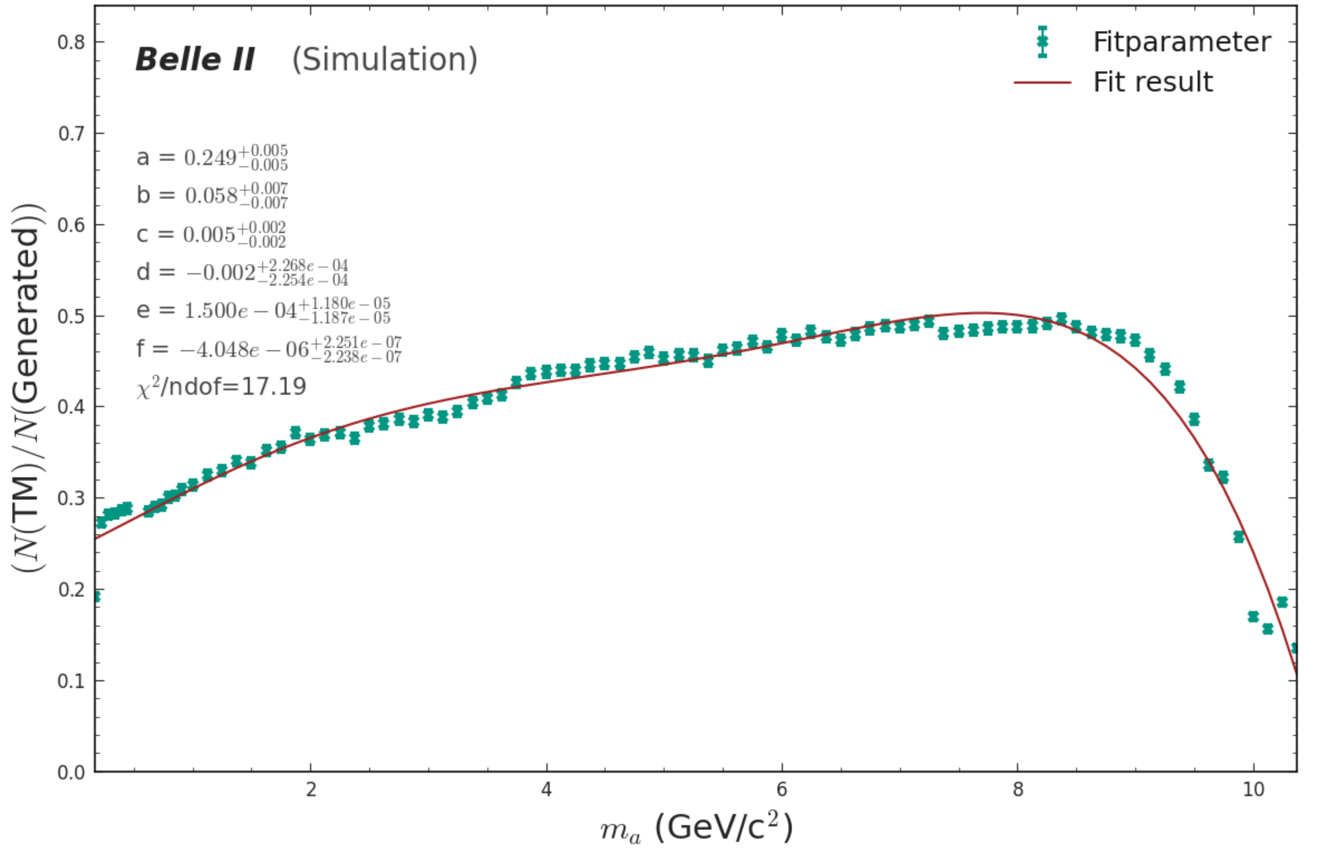






Signal Efficiency

- **Signal/Fit Range**:
  - For now: $[-20\sigma_{\text{DSCB}}, +20\sigma_{\text{DSCB}}]$
  - Signal shape is more or less symmetrical
- **Signal Efficiency**: $\dfrac{\text{Number of TM Events}}{\text{Number of Generated Events}}$



- It isn't easy to fit polynomial PDF to this shape
  - Possible solution:
    - Signal MC for each scan point
- Trigger efficiency not taken into account