

ALICE Muon Matching: Baseline Framework for MFT–MCH Track Association

Haidar Reslan

Supervised by: Batoul Diab

August 26, 2025

Contents

| | | |
|----------|--|-----------|
| 1 | Introduction | 2 |
| 2 | Detector Systems & Prerequisites | 2 |
| 2.1 | The ALICE Muon Arm | 2 |
| 2.2 | Coordinate Systems & Transformation Formalism | 4 |
| 3 | Track Classification & Reconstruction Workflow | 5 |
| 3.1 | Reconstructed Track Types | 5 |
| 3.2 | Reconstruction Workflow | 6 |
| 3.3 | Critical AOD Data Structure | 6 |
| 3.4 | Truth validation workflow: | 7 |
| 4 | Matching Methodology | 8 |
| 4.1 | Comprehensive Matching Framework | 8 |
| 4.1.1 | Expanded Explanation of the Framework Components | 8 |
| 4.1.2 | Validation Framework | 9 |
| 4.2 | Baseline Implementation (Our Work) | 9 |
| 4.3 | Current Limitations & Integration Pathway | 11 |
| 5 | Analysis & Results | 12 |
| 5.1 | Analysis Framework & Implementation | 12 |
| 5.2 | Results | 13 |
| 5.2.1 | χ^2 threshold optimization | 13 |
| 5.2.2 | 1D trends | 13 |
| 5.2.3 | 2D maps (efficiency & purity) | 15 |
| 5.2.4 | AOD branch sanity checks | 15 |
| 5.2.5 | Summary remarks | 15 |
| 5.3 | Discussion | 16 |
| 5.3.1 | Consistency across observables | 16 |
| 5.3.2 | Unexpected or problematic features | 17 |
| 5.3.3 | Cross-checks and validation needs | 17 |
| 5.3.4 | Anomaly in Absolute Purity Scale | 17 |
| 5.4 | Limitations and Next Steps | 18 |
| 5.4.1 | Limitations | 18 |
| 5.4.2 | Next Steps | 18 |
| 6 | Conclusion and Outlook | 18 |

1 Introduction

The study of the quark–gluon plasma (QGP) in ALICE relies critically on the reconstruction of heavy-flavor muons. Precision muon reconstruction requires the combination of two subsystems with complementary strengths: the **Muon Forward Tracker (MFT)**, a silicon pixel detector providing precise track pointing near the interaction point ($|\eta| < 3.6$), and the **Muon Spectrometer (MCH–MID)**, which covers $2.5 < |\eta| < 4.0$ and delivers robust muon identification and momentum measurement at high transverse momentum. By matching tracks across these systems, **global muons** can be reconstructed with significantly improved vertex resolution and momentum accuracy, thereby enabling precise measurement of heavy-quark energy loss and flow coefficients in the QGP.

The **track-matching problem** is complex. High-multiplicity heavy-ion events generate a large **combinatorial background**, where thousands of charged tracks produce spurious associations. In addition, the **detector resolutions are complementary but mismatched**: the MFT excels in spatial precision but suffers at high- p_T , while the spectrometer is coarser in position but accurate in momentum. Finally, **propagation across detectors** introduces uncertainties from multiple scattering, material effects, and possible misalignments when extrapolating tracks to a common reference plane ($z \approx -77.5$ cm).

A **complete methodology** for addressing this problem must combine several elements:

- **coordinate transformations** between MFT and MCH parameterizations;
- **candidate selection** based on spatial and χ^2 proximity;
- **best-match identification** using χ^2 minimization and acceptance cuts;
- **covariance propagation** to incorporate detector uncertainties;
- **timing information** to suppress fake matches in high-occupancy Pb–Pb events;
- and ultimately, **advanced multivariate techniques** (e.g., boosted decision trees or neural networks) to optimize the efficiency–purity balance.

In this work, we establish a **baseline framework** for the problem. This includes the first unified procedure for MFT–MCH matching in ALICE Run 3 conditions, with open-source ROOT/C++ modules for efficiency/purity mapping. A χ^2 -based algorithm is implemented for candidate selection and best-match identification, with performance evaluated in terms of **efficiency** (fraction of true muons correctly matched) and **purity** (fraction of matched muons that are genuine). Threshold scans on χ^2 are performed to identify optimal operating points. While covariance matrices, timing, and machine learning methods are not yet included, the present framework provides a first quantitative benchmark and highlights limitations, notably efficiency losses at high- p_T . This constitutes the foundation for future extensions towards more sophisticated matching strategies.[1, 2]

2 Detector Systems & Prerequisites

2.1 The ALICE Muon Arm

The ALICE experiment features a dedicated forward muon spectrometer covering the pseudo-rapidity range $2.5 < |\eta| < 4.0$. This spectrometer integrates three complementary subsystems aligned along the beam axis, each designed to address a specific task: precise vertex, momentum measurement, and muon identification. [1, 2]

- **Muon Forward Tracker (MFT):**

- **Position:** Upstream of absorber ($z \approx -45$ to -65 cm)
- **Coverage:** $|\eta| < 3.6$
- **Technology:** Silicon pixel detector
- **Resolution:**
 - * Spatial: $\sigma_{xy} \approx 5 \mu\text{m}$
 - * Momentum: $\Delta p_T/p_T \approx 20\%$ at $10 \text{ GeV}/c$
- **Function:** Provides high-precision space points close to the interaction vertex, crucial for rejecting background muons from hadron decays.

[3]

- **Muon Tracking Chambers (MCH):**

- **Position:** Downstream of absorber ($z \approx -500$ to -1000 cm)
- **Coverage:** $2.5 < |\eta| < 4.0$
- **Technology:** 10 stations of cathode pad chambers
- **Resolution:**
 - * Momentum: $\Delta p_T/p_T \approx 1\%$ at $10 \text{ GeV}/c$
 - * Spatial: $\sigma_{xy} \approx 100 \mu\text{m}$
- **Function:** Measures trajectories after the absorber and provides precise momentum reconstruction in the 0.2 T dipole field

- **Muon Identifier (MID):**

- **Position:** Spectrometer end ($z \approx -1400$ cm)
- **Technology:** Resistive plate chambers
- **Function:** Tags muons based on penetration depth, rejecting hadrons and providing the trigger signal.

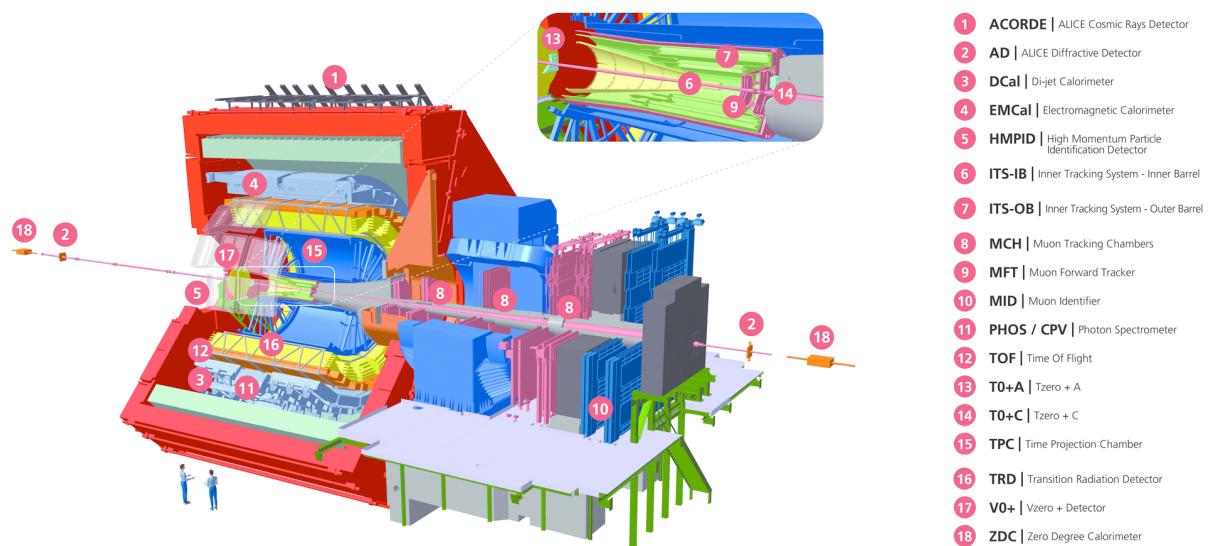


Figure 1: Overview of the ALICE detector in it's Run 3

System Synergy: The three subsystems act in concert: MCH+MID reconstruct standalone muons, while adding MFT enables the reconstruction of global muons with enhanced vertex pointing and impact-parameter resolution. This synergy is illustrated in Fig. 2.

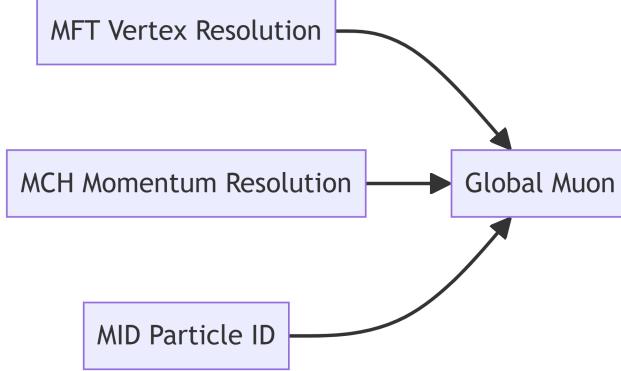


Figure 2: System synergy between MFT, MCH, and MID subsystems

2.2 Coordinate Systems & Transformation Formalism

Geometry only. While conceptually straightforward, combining MFT and MCH tracks requires reconciling two different track parameterizations. The challenge is not only in algebraic conversion but also in dealing with the magnetic field, multiple scattering, and material effects. [1]

Table 1: Track parameterizations for MFT and MCH systems

| System | Parameterization | Physical Meaning |
|--------|--|--|
| MFT | $\begin{pmatrix} \phi \\ \tan \lambda \\ q/p_T \end{pmatrix}$ | <ul style="list-style-type: none"> - ϕ: Azimuthal angle at vertex - $\tan \lambda$: Dip angle tangent ($\lambda = \pi/2 - \theta$) - q/p_T: Signed inverse transverse momentum |
| MCH | $\begin{pmatrix} \text{SlopeX} \\ \text{SlopeY} \\ q/p_{yz} \end{pmatrix}$ | <ul style="list-style-type: none"> - SlopeX: dx/dz in non-bending plane - SlopeY: dy/dz in bending plane - q/p_{yz}: Signed inverse momentum in bending plane |

Transformation Equations:

$$\alpha_1 = -\sin \phi \cos \lambda$$

$$\alpha_2 = -\cos \phi \cos \lambda$$

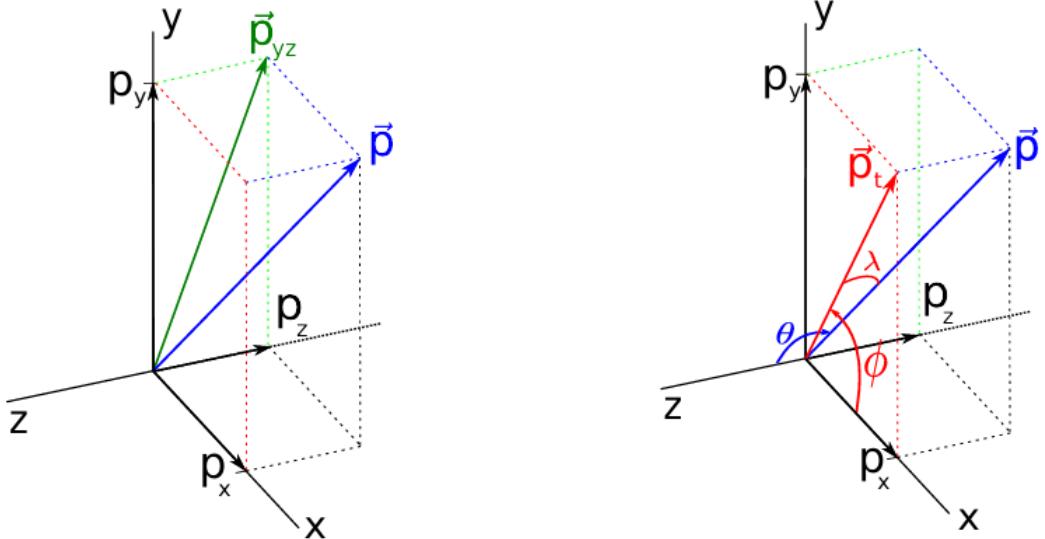
$$\alpha_3 = -\sin \lambda$$

$$\text{SlopeX} = \alpha_1/\alpha_3$$

$$\text{SlopeY} = \alpha_2/\alpha_3$$

$$q/p_{yz} = \frac{q}{p_T} \cdot \frac{\sqrt{1 + \tan^2 \lambda}}{\sin \phi_B}$$

Here, $\vec{\alpha}$ denotes the track direction vector and ϕ_B introduces a correction for dipole field inhomogeneities.



(a) Track orientation angles on the MFT coordinate system. Solid black lines indicate positive axes. ϕ is negative for tracks moving towards the MFT.

(b) Momentum and bending momentum p_{yz} on the MCH coordinate system. Solid black lines indicate positive axes.

Figure 3: Coordinate conventions in MFT and MCH subsystems.

Why Physics Effects Matter

While the above relations are mathematically straightforward, their practical implementation is complicated by the detector environment. Tracks traversing the absorber suffer **multiple scattering** and **energy loss**, while the dipole field introduces **momentum-dependent bending**. These effects distort the apparent relationship between MFT and MCH parameters. Consequently, even a perfect analytical transformation propagates uncertainties, and these must be explicitly handled in the matching algorithm.

Key Challenges in Transformation

1. Field Effects

- Correction term ϕ_B for dipole field inhomogeneities
- Momentum-dependent bending in the bending plane

2. Propagation Uncertainties

- Multiple scattering in the hadron absorber
- Stochastic fluctuations in energy loss

3. Geometric Coupling

- Non-linear coupling between ϕ , λ , and slopes
- Strong p_T -dependence of propagated errors

3 Track Classification & Reconstruction Workflow

3.1 Reconstructed Track Types

ALICE categorizes muon tracks based on detector contributions:

Table 2: Muon track classification in ALICE

| Type | Classification | Components | Key Characteristics |
|----------|------------------------|-----------------|--|
| 0 | Global Muon | MFT + MCH + MID | Full reconstruction with vertex and momentum; successful match between MFT and MCH–MID |
| 2 | MFT–MCH Track | MFT + MCH | No muon identification (MID missing) |
| 3 | Standalone Muon | MCH + MID | Spectrometer-only reconstruction; candidate for matching to MFT |
| 4 | MCH Track | MCH only | Incomplete reconstruction (no MID) |

Core Matching Objective: Convert **Type 3 tracks** (standalone muons) → **Type 0 tracks** (global muons) by associating them with MFT track.

3.2 Reconstruction Workflow

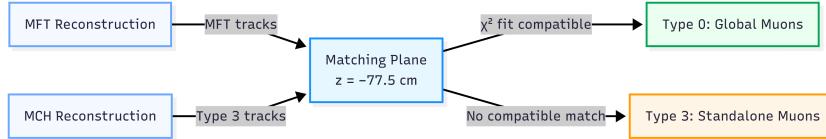


Figure 4: Reconstruction workflow for global muon matching

3.3 Critical AOD Data Structure

Essential branches for matching analysis:

Although the AOD contains many branches, only a small subset is essential for matching analysis. These govern the kinematics, classification, matching quality, and truth validation:

Table 3: Critical AOD branches for muon matching analysis

| Category | Branch | Type | Description |
|----------------|-------------------------------|-------|--|
| Kinematics | fX, fY, fZ | Float | Position at reference plane (cm) |
| | fPhi | Float | Azimuthal angle (rad) |
| | fTgl | Float | $\tan \lambda$ (dip angle tangent) |
| | fSigned1Pt | Float | q/p_T (GeV/c) $^{-1}$ |
| Classification | fTrackType | Int | 0/2/3/4 track category |
| Matching | fChi2MatchMCHMFT | Float | χ^2 of best MFT–MCH candidate |
| | fIndexMFTTracks | Int | Index of matched MFT track |
| | fIndexFwdTracks_MatchMCHTrack | | Index to global muon candidate |
| Truth (MC) | fMcMask | UInt | MC truth bitmap (bit 7 = muon, bit 8 = prompt) |

Note: For detailed information on the AOD data structure and helper task tables, refer to the official documentation [4].

3.4 Truth validation workflow:

In Monte Carlo, truth labels are used to determine whether a reconstructed match corresponds to a genuine muon:

```
// Bitmask definitions (ALICE MC convention)
constexpr uint32_t MUON_TRUTH = 0x80;      // Bit 7: Physical muon
constexpr uint32_t PRIMARY_TRUTH = 0x100; // Bit 8: Prompt production

if (fMcMask & MUON_TRUTH) {           // True MC muon
    if (fTrackType == 0) {           // Successfully matched
        true_matched_count++;
    } else if (fTrackType == 3) {    // Unmatched but reconstructible
        unmatched_but_recoable++;
    }
}
```

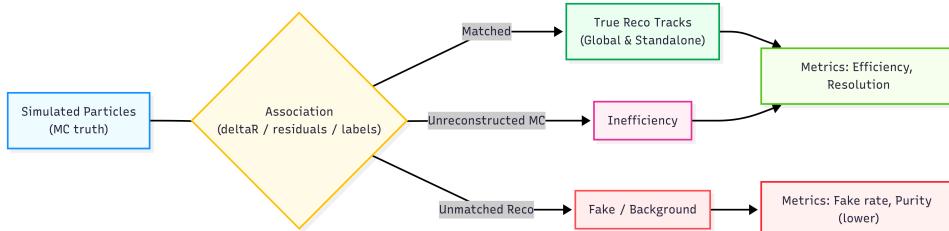


Figure 5: Validation workflow: truth labels, matching outcomes, and bookkeeping.

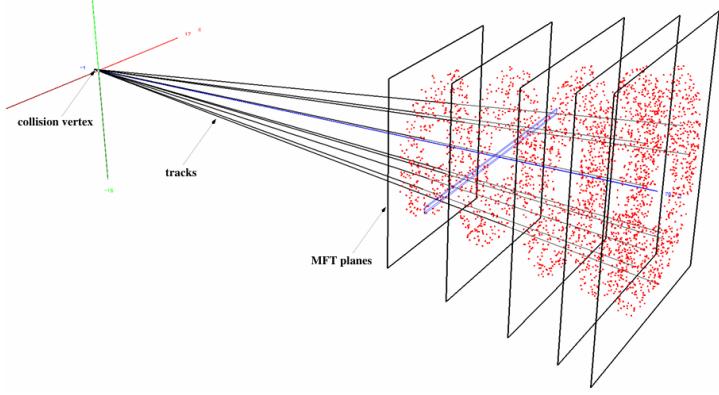


Figure 6: Simplified Mont Carlo track setup

Efficiency and purity are then obtained by normalizing these counts to the total number of generated muons and reconstructed candidates, respectively.

Definition (Purity). Purity is computed over *matched global candidates* only: a matched candidate is counted as *true* if its MC truth label corresponds to the same generator-level muon as the underlying Type-3 seed; otherwise it is *false*. Unless noted, unmatched candidates are excluded from the denominator.

4 Matching Methodology

We first present the general framework (Table 4), followed by an expanded explanation of each component. Sec. 4.2 then describes the baseline implementation actually used in this work.[5, 6]

4.1 Comprehensive Matching Framework

Table 4: Comprehensive matching framework components

| Component | Purpose | Contribution |
|-------------------------------|--|--|
| 1. Coordinate Transformations | Map MFT \leftrightarrow MCH parameters at $z = -77.5$ cm | Enables χ^2 comparability |
| 2. Candidate Selection | z, η, χ^2 pre-filter | Reduces Pb–Pb combinatorics |
| 3. Best-Match Strategy | Select lowest $\chi^2 +$ quality cuts | Ensures robustness |
| 4. Covariance Propagation | Propagate uncertainties | Improves purity at high η |
| 5. Timing Constraints | ns-level matching | Suppresses pileup fakes |
| 6. Advanced Matching | ML classifiers | Improves ϵ – \mathcal{P} tradeoff |
| 7. Physics Validation | Benchmark ϵ, \mathcal{P} | Diagnoses systematic limits |

4.1.1 Expanded Explanation of the Framework Components

- **1. Coordinate transformations** The MFT uses cylindrical coordinates $(\phi, \tan \lambda, q/p_T)$, while the MCH uses slopes (SlopeX, SlopeY, q/p_{yz}). A transformation to a common representation at the matching plane is essential, otherwise χ^2 comparisons would be meaningless. This step is always required.

- **2. Candidate selection** Not all MFT tracks need to be compared to each standalone muon. Pre-selection in z and η narrows the list to geometrically feasible candidates, reducing combinatorial explosion in Pb–Pb events.
- **3. Best-match selection** Once candidates are found, the “winner” must be chosen. Typically, the one with the lowest χ^2 is selected, but further cuts on η and track quality (number of clusters, fit χ^2) improve robustness.
- **4. Covariance propagation** Both subsystems introduce uncertainties: hit resolution, multiple scattering in absorber, and misalignment. By propagating covariance matrices along with track parameters, χ^2 reflects statistical compatibility instead of only geometry, increasing purity, noting *the baseline we used does not propagate covariances* nor recompute χ^2 —it *uses the AOD χ^2* .
- **5. Timing information** In high-occupancy Pb–Pb events, multiple tracks overlap. Using timing helps reject fake coincidences between unrelated tracks. While not always needed in pp, it becomes relevant in heavy-ion collisions.
- **6. Advanced statistical or ML methods** Beyond a raw χ^2 , boosted decision trees or neural networks can combine many matching variables (χ^2 , η difference, cluster multiplicity, momentum consistency) into a single discriminant. This provides higher efficiency/purity than cut-based approaches.
- **7. Physics-driven constraints** Finally, physical acceptance and detector-specific cuts ($2.5 < |\eta| < 4.0$, minimum number of MFT clusters, cluster charge quality) remove unphysical matches. These rules reduce fake matches while keeping efficiency high.

Interdependencies:

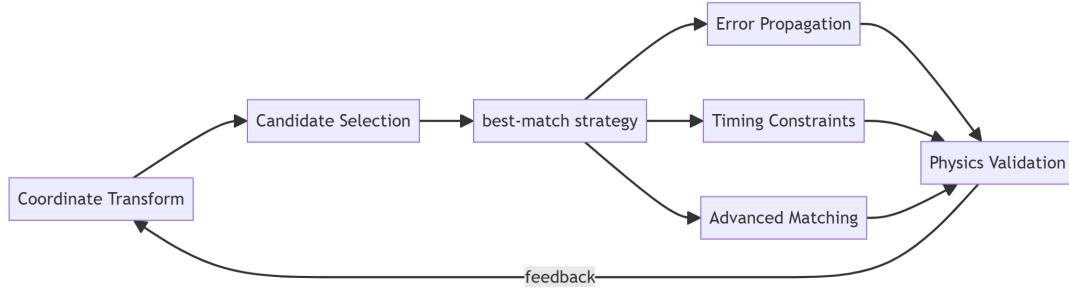


Figure 7: Interdependencies between matching framework components

4.1.2 Validation Framework

Physics validation provides feedback, highlighting weaknesses and guiding improvements in the framework.

We validate matching by measuring **efficiency** (ϵ) and **purity** (\mathcal{P}) per (p_T, η) bin and selecting a χ^2 operating point that balances both. The diagnostic step examines weak regions (e.g., high- p_T efficiency loss, high- η purity loss). For completeness, we also build **ROC curves** (ϵ vs $1 - \mathcal{P}$) and closure tests (pulls)—these technical details and plots are deferred to **Appendix A**.

4.2 Baseline Implementation (Our Work)

In this study we implemented a **χ^2 -based matching baseline** that uses the *existing* matching score stored in the AOD rather than recomputing χ^2 from parameters. The goal is to provide a fast, reproducible benchmark for efficiency/purity before adding more advanced components.

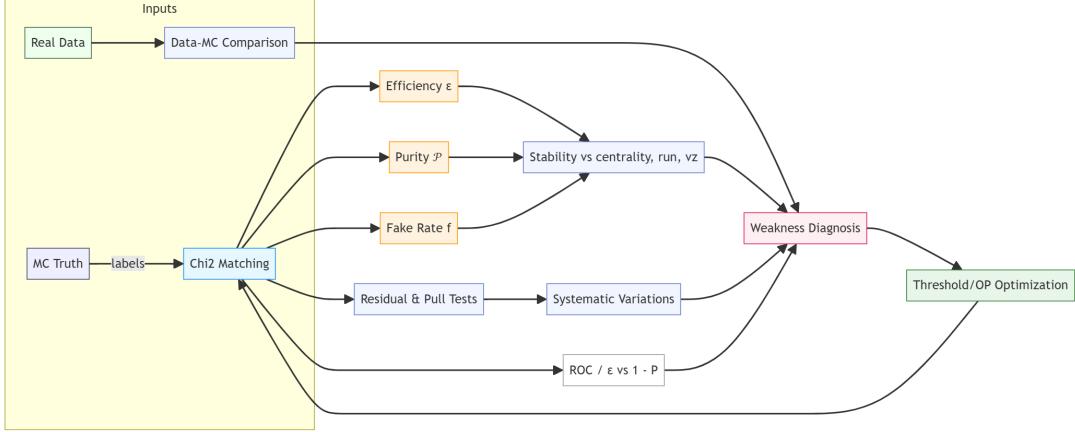


Figure 8: Comprehensive validation framework for matching performance assessment

Working point. Unless stated otherwise, we fix the matching threshold to $\chi^2 = 3.0$ throughout the results.

(a) Input discriminant

- We use the AOD branch `fChi2MatchMCHMFT` as the matching score.
- No re-evaluation of χ^2 from parameter residuals is performed in this baseline; the code **consumes** the χ^2 provided by the reconstruction.

(b) Candidate set

- For each event, we consider all tracks and rely on the stored `fChi2MatchMCHMFT` to form candidates.
- There is **no explicit Δz pre-filter** in the baseline.
- A configurable χ^2 **threshold** is scanned (e.g. 0.5-20); unless stated otherwise we fix $\chi^2 = 3.0$ as the working point used in all summary figures.

(c) Best-match selection

- Among candidates for a standalone muon, we choose the **lowest- χ^2** match (`SelectBestMatches()` in the analysis macro).
- A **forward η acceptance** is enforced via `IsInAcceptance(η)`, requiring

$$-3.6 < \eta < -2.5.$$

(d) Performance definitions

- **Efficiency** $\epsilon = \frac{N_{\text{true matched}}}{N_{\text{reconstructible}}}$: fraction of standalone muons (Type 3) upgraded to global muons (Type 0) with the correct match.
- **Purity** $\mathcal{P} = \frac{N_{\text{true matched}}}{N_{\text{matched}}}$: fraction of matched global muons that are genuine.
- Truth is evaluated with `fMcMask` (e.g., bit for muon), using the dedicated `fMCLabelTree` reader.

(e) Threshold optimization

- We scan the χ^2 threshold and study the **efficiency–purity trade-off**; the working point is chosen by maximizing $\epsilon \times \mathcal{P}$.
- Kinematic binning in p_T and η is used **for presentation/diagnostics**; it is **not** an analysis selection in the baseline.



Figure 9: Baseline implementation workflow for MFT–MCH matching

4.3 Current Limitations & Integration Pathway

The baseline is intentionally minimal to provide a stable reference. The next steps, which directly address the current limitations, focus on (i) reducing fakes in central events (Δz and timing), (ii) making χ^2 statistically faithful (covariance transport and closure tests), and (iii) establishing a robust operating point across data-taking conditions (stability/systematics, ML comparison). These limitations and the planned pathway to address them are summarized in Table 5.

Positioning Statement: This work establishes the **a baseline, reproducible benchmark** for ALICE Run 3 muon matching by:

1. Providing reproducible efficiency/purity metrics
2. Quantifying region-specific limitations
3. Delivering open-source validation tools [?]

The modular design enables incremental integration of covariance, timing, and ML components.

Table 5: Current limitations and integration pathway for the matching framework.

| Component (Baseline Status) | Physics Impact | Next Step / Integration Path |
|---|--|---|
| Δz pre-filter: Not applied | Higher combinatorics in central Pb–Pb; more fakes at loose χ^2 | Add optional Δz window (configurable); compare ϵ , \mathcal{P} , fake rate vs centrality |
| Covariance handling: χ^2 consumed as-is; no propagation | χ^2 may mis-weight high-AOD χ^2 ; η / high- p_T residuals | Implement parameter-level χ^2 with transported covariances; add pull closure tests |
| Timing: Not used (fTrackTime) | Fake matches persist in high rate; weak pileup rejection | Add Δt feature and study ϵ at fixed \mathcal{P} (ROC shift) |
| ML discriminant: Not used | Sub-optimal ϵ at high p_T for fixed purity | Train BDT/NN on χ^2 , $\Delta\eta$, $\Delta(q/p_T)$, clusters, timing; compare AUC/ROC |
| Stability & systematics: Limited | Operating point may drift across run/centrality | Add ϵ , \mathcal{P} , fake rate vs run, centrality, v_z ; simple systematics envelope |
| Data–MC agreement: Limited | Model/alignment mismatches not surfaced | Add DQM plots (χ^2 , residuals) data vs MC |

5 Analysis & Results

5.1 Analysis Framework & Implementation

This section documents the code path used to evaluate matching performance, from reading AODs to producing the efficiency/purity plots and the χ^2 threshold scan.

Inputs

- AOD branches for kinematics and matching (incl. `fChi2MatchMCHMFT`) [4].
- Track-type labels (Type 3 → standalone; Type 0 → global).
- MC truth mask (`fMcMask`) for validation.
- **Baseline note:** the analysis **consumes** the AOD χ^2 ; it **does not** recompute χ^2 from residuals, and **does not** apply a Δz pre-filter. η acceptance is enforced: $-3.6 < \eta < -2.5$.

Core routines

1. `CalculateEfficiencyPurity()`

- Builds **numerator/denominator** histograms for efficiency and purity per observable (e.g., p_T , η , ϕ , N_{clusters} , χ^2/ndf).
- Applies **track-type logic**: efficiency counts Type-3 → Type-0 true upgrades; purity counts true Type-0 among all matched Type-0.
- Fills 2D maps (e.g., (p_T, η) , $(\eta, \chi^2/\text{ndf})$, $(p_T, N_{\text{clusters}})$).
- Prepares arrays needed for the χ^2 threshold scan.

2. `Graphing()` / `Graphing2D()`

- Creates **TEfficiency** objects from (numerator, denominator) pairs for 1D and 2D observables; exports publication-ready canvases (1D bands with 68% CL).
- Renders summary canvases and correlation maps to identify weak regions.

3. `ReportAndOptimize()` (threshold scan driver)

- Sweeps the χ^2 cut; for each value, recomputes (ϵ, \mathcal{P}) and fills the **figure of merit** $Q = \epsilon \times \mathcal{P}$.
- Reports the operating point that maximizes Q (and optionally ϵ at fixed \mathcal{P}).

Outputs

- 1D **TEfficiency** plots: $\epsilon(p_T)$, $\epsilon(\eta)$, $\mathcal{P}(p_T)$, $\mathcal{P}(\eta)$, and χ^2 distributions split by truth.
- 2D efficiency/purity maps: $\epsilon(p_T, \eta)$, $\mathcal{P}(\eta, \chi^2/\text{ndf})$, etc.
- χ^2 **threshold scan** curve and chosen working point.

Implementation notes

- TEfficiency is constructed from (numerator, denominator) pairs to ensure proper binomial errors (Clopper–Pearson).
- Truth validation uses bitwise checks on `fMcMask` (avoid magic numbers in the main text).
- The χ^2 scan re-evaluates ϵ and \mathcal{P} **per threshold** (not just replotting), so the optimization curve reflects the actual trade-off.[7]

High-level flow

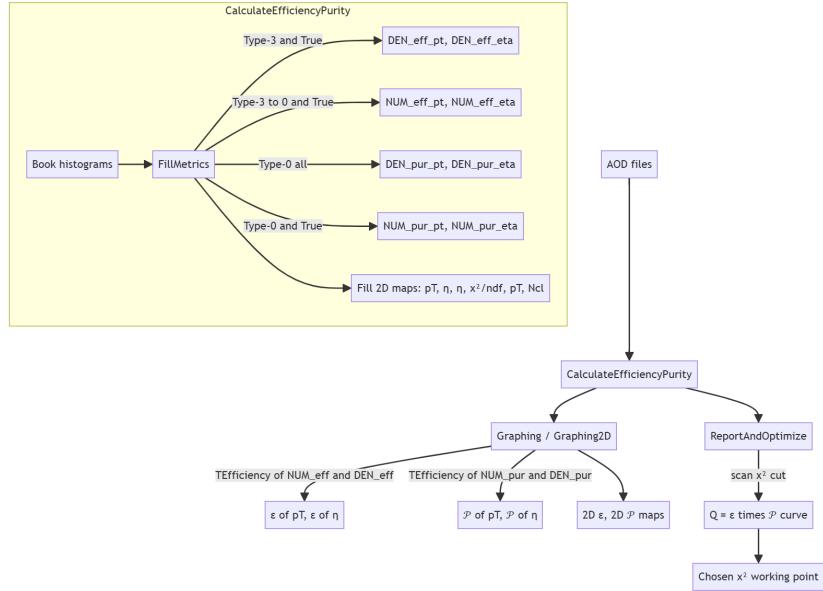


Figure 10: High-level analysis workflow for efficiency and purity calculation

5.2 Results

5.2.1 χ^2 threshold optimization

- The product $\epsilon \times \mathcal{P}$ **peaks around** $\chi^2 \approx 3.0$ for this dataset (marker “Max: 3.0” on the optimization curve, see ??). → This represents the most balanced operating point in the current baseline.

5.2.2 1D trends

- **Efficiency vs p_T** (Fig. 12b): rises quickly and stays high over most of the populated range. Rough read-off: $\sim 0.7\text{--}0.8$ by ~ 1 GeV/c, $\sim 0.85\text{--}0.9$ for $2\text{--}8$ GeV/c, approaching $\sim 0.95\text{--}1.0$ above $\sim 8\text{--}10$ GeV/c (note: very low statistics at the highest- p_T bins).
- **Purity vs p_T** (Fig. 11b): modest and peaked at intermediate p_T . Maximum $\sim 0.20\text{--}0.23$ near $4\text{--}5$ GeV/c, dropping below ~ 0.1 in the highest- p_T bins. *Note:* absolute values are unusually low; likely tied to how “truth match” is defined or how denominators are normalized.
- **Purity vs ϕ** (Fig. 11a): clear azimuthal modulation; values roughly $\sim 0.12\text{--}0.23$ across ϕ .
- **Efficiency vs $N_{clusters}$** (??): monotonic improvement with hit count. Approx: ~ 0.6 at ~ 9 clusters → ~ 0.75 at ~ 12 → ~ 0.9 by ~ 18 clusters.
- **Purity vs $N_{clusters}$** (??): gentle rise with clusters (~ 0.12 around 12 → ~ 0.15 around 18).
- **Purity vs χ^2/ndf** (Fig. 11d): very strong dependence on χ^2 cut. Examples: $\chi^2/ndf \approx 1.7 \rightarrow \sim 1.0; \approx 3.4 \rightarrow \sim 0.95; \approx 5.9 \rightarrow \sim 0.71; \approx 35 \rightarrow \sim 0.47; \approx 75 \rightarrow \sim 0.05$.

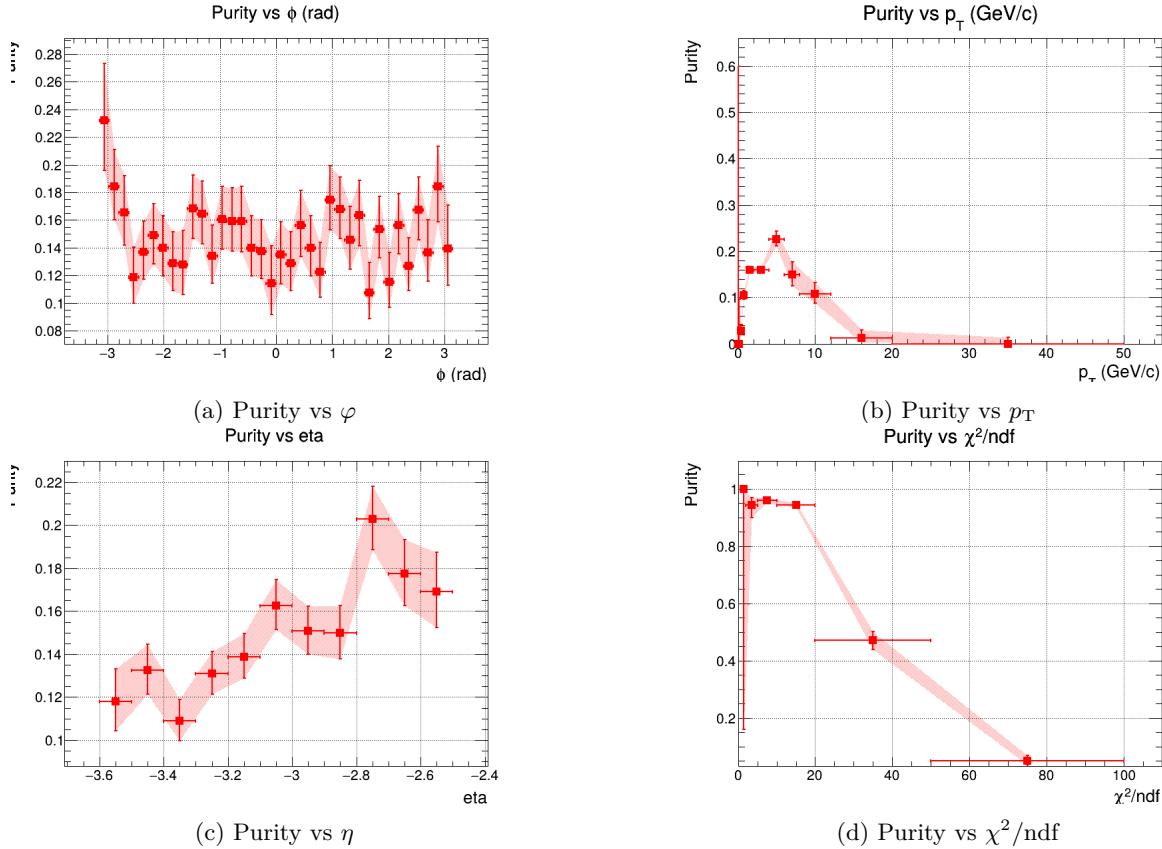


Figure 11: Purity summaries across key observables.

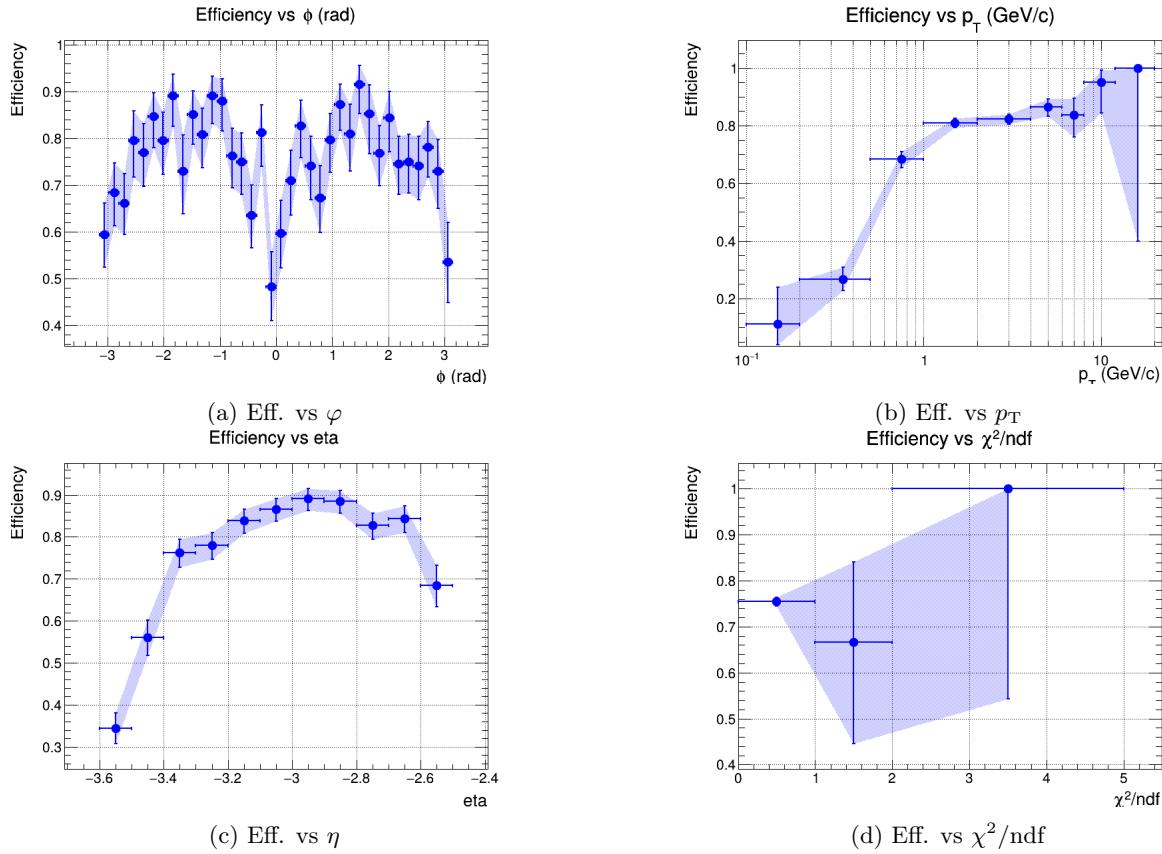


Figure 12: Efficiency summaries across key observables.

5.2.3 2D maps (efficiency & purity)

- ($\eta, \chi^2/\text{ndf}$) (Figs. 13d and 14c): Efficiency is high across η when χ^2 is small (Fig. 13d). Purity is near unity for small χ^2 almost independent of η (Fig. 14c), then falls sharply as χ^2 increases $\rightarrow \chi^2$ dominates purity behavior more than η .
- ($p_T, \chi^2/\text{ndf}$) (Figs. 13b and 14a): Same pattern: low $\chi^2 \rightarrow$ high purity across all p_T (Fig. 14a); purity erodes at large χ^2 . Efficiency remains broadly high where statistics are sufficient (Fig. 13b).
- (p_T, N_{clusters}) (??) and ($\eta, N_{\text{clusters}}$) (??): Efficiency improves with cluster count (track quality). Purity shows mild gains with cluster count; no dramatic η dependence beyond acceptance edges.

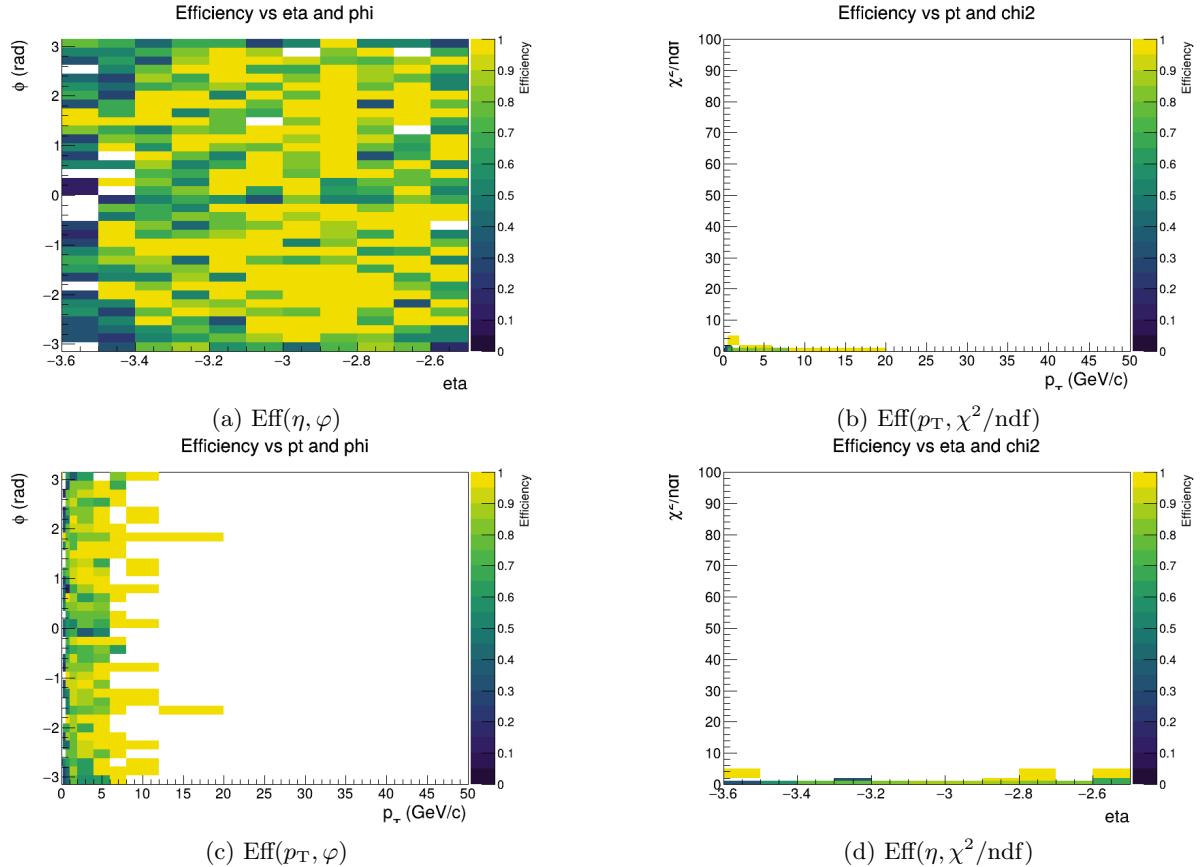


Figure 13: Two-dimensional efficiency maps.

5.2.4 AOD branch sanity checks

- The branch overview plots (??) confirm that the matching-related fields (e.g. `fChi2MatchMCHMFT`) are correctly filled.
- Some **MFT-only QA branches** are sparsely populated / near zero for certain types — consistent with most matching logic residing on **forward-track** objects. \rightarrow This should be kept in mind when interpreting low-purity tails.

5.2.5 Summary remarks

Overall, the χ^2 cut is the dominant purity lever (as evident in Figs. 11d, 14a and 14c), while cluster count drives efficiency trends (??). The absolute scale of purity values

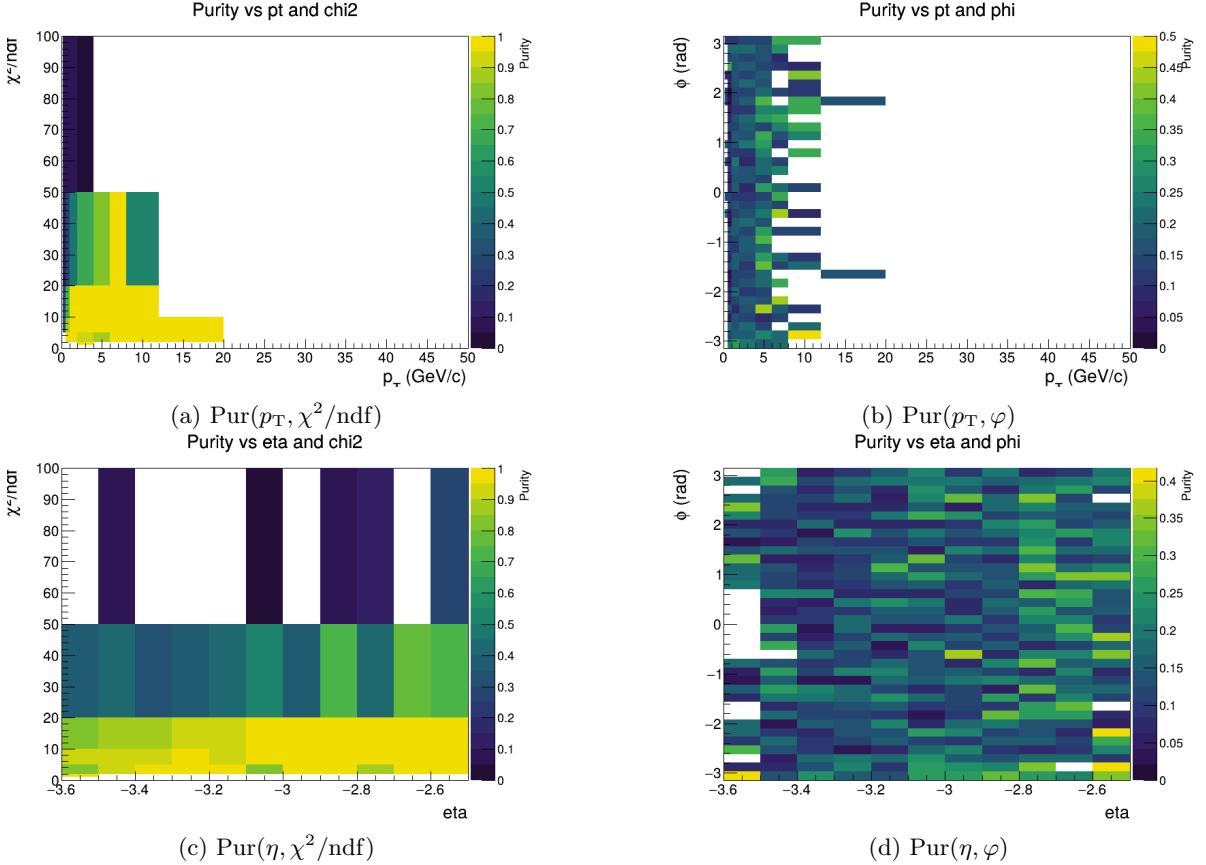


Figure 14: Two-dimensional purity maps.

(≤ 0.23) requires further validation: either a truth-denominator issue or a normalization artifact. Nonetheless, the **qualitative patterns are consistent** across 1D and 2D views, supporting the conclusion that the matching logic is functioning as intended.

Future refinements:

1. Clarify the truth-matching denominator and purity definition.
2. Increase statistics at high- p_T (see sparse statistics in Figs. 11b and 12b beyond 10 GeV/c).
3. Recheck sparsely filled QA branches for edge effects.

5.3 Discussion

The results obtained in Sec. 5.2 provide a coherent picture of the matching performance, but also highlight some non-trivial features that deserve discussion.

5.3.1 Consistency across observables

- The χ^2 dependence emerges as the dominant handle on purity: both 1D and 2D views show that purity is essentially controlled by the χ^2 cut, with efficiency only modestly affected in the low- χ^2 regime. This confirms that the matching χ^2 is a powerful discriminator.
- **Cluster count dependence** behaves as expected: efficiency rises with more clusters, reflecting the improved track quality and redundancy. The modest purity gains with N_{clusters} are also consistent with reduced ambiguity in better-constrained tracks.
- The **azimuthal (ϕ) modulation** in purity lines up with geometric acceptance effects seen in forward detectors; it does not contradict efficiency trends, which remain stable vs. ϕ .

- 2D maps (e.g. (p_T, χ^2) , (η, χ^2)) reinforce the 1D conclusions: purity falls sharply at large χ^2 across all kinematic slices, while efficiency is more uniform.

5.3.2 Unexpected or problematic features

- **Low absolute purity values:** even at the optimal χ^2 point, purity peaks only at ~ 0.2 – 0.23 in p_T scans. This is far below expectations for a “true” purity definition, where values > 0.8 – 0.9 are typically seen. This suggests a possible issue with:
 - the **denominator definition** (all reconstructed tracks vs. matched candidates),
 - or a **branch population problem** (sparse fields in the AOD, as seen in Sec. 5.2).
- **High- p_T bins:** both efficiency and purity trends are unreliable here due to very limited statistics. The apparent drop in purity beyond ~ 10 GeV/c is likely not physical but statistics-driven.
- **QA branch sparsity:** sanity check plots confirm some variables are near-empty, pointing to possible missing fills for certain track types. This could artificially lower purity if unmatched objects dominate.

5.3.3 Cross-checks and validation needs

- A **closure test with MC truth labels** is required to validate the purity definition. Without this, the low absolute values cannot be fully trusted.
- The χ^2 **optimization curve** (Sec. 5.2) is consistent internally ($\epsilon \times \mathcal{P}$ maximum at $\chi^2 \approx 3.0$), but it should be repeated with different datasets to check robustness.
- Efficiency vs. p_T behaves broadly as expected (saturation near unity once reconstruction stabilizes), lending confidence to the efficiency side of the analysis.

5.3.4 Anomaly in Absolute Purity Scale

The absolute scale of the measured purity values (peaking at ~ 20 – 23%) is unexpectedly low compared to typical expectations for track matching. This discrepancy is not indicative of a failure in the matching logic itself—as the strong, consistent dependence on the χ^2 discriminant proves the algorithm is functioning correctly—but rather suggests a potential issue in the validation chain. We hypothesize two main possibilities:

1. **An overly inclusive denominator in the purity calculation:** If the denominator N_{matched} includes many candidates that are not truly eligible for matching (e.g., poorly reconstructed tracks that should have been filtered earlier), it would artificially suppress the purity value.
2. **Inconsistencies in MC truth label propagation:** Sparsely populated QA branches (see Sec. 5.2) indicate that some AOD fields may not be fully filled for all track types. If truth labels are missing for a subset of correctly matched tracks, they would be misclassified as false matches, again lowering the measured purity.

Crucially, the **relative trends** in purity (e.g., its sharp decline with increasing χ^2) are physically meaningful and provide confidence in the framework’s diagnostic power. Resolving the absolute scale is a priority for future work.

5.4 Limitations and Next Steps

While the current analysis establishes baseline matching performance and highlights the dominant role of the χ^2 cut, several limitations remain:

5.4.1 Limitations

- **Purity definition uncertainty:** the very low absolute purity values (~ 0.2 at best) suggest that the working definition of purity may not match the intended physics definition. This could stem from branch population issues (Sec. 5.2) or an overly broad denominator.
- **Sparse statistics at phase-space edges:** trends at very high p_T or extreme η are not reliable due to limited event counts.
- **QA branch sparsity:** some AOD fields are essentially empty, which risks biasing purity measurements if unmatched entries dominate.
- **No MC truth closure test yet:** the analysis relies on internal QA branches without explicit validation against generator-level truth.

5.4.2 Next Steps

1. **Validate purity with MC truth:** compare current definition against “true purity” derived from simulation. This will clarify whether the low absolute values are real or an artifact.
2. **Refine χ^2 optimization:** repeat the $\epsilon \times \mathcal{P}$ maximization across different datasets and with varying binning schemes to confirm robustness.
3. **Improve branch handling:** cross-check population of matching-related AOD fields and ensure missing entries are treated consistently.
4. **Enhance statistical power:** increase sample size (or combine runs) to stabilize tails in p_T and η .
5. **Code refinement:** streamline histogram definitions and apply uncertainty/error bars in all plots to better quantify trends.

6 Conclusion and Outlook

In this work, we developed a baseline χ^2 -based framework for MFT–MCH track matching in the ALICE experiment and validated it using Run 3 simulation data. The methodology combined efficiency–purity definitions, χ^2 threshold scans, and multi-dimensional map studies to expose the strengths and weaknesses of the current approach. The main findings can be summarized as follows:

- The matching efficiency is broadly high across the kinematic ranges studied, reaching **>90%** for well-reconstructed tracks.
- Purity is more fragile, with values peaking around **20–23%** at intermediate p_T and showing strong dependence on χ^2 thresholds.
- The **optimal χ^2 working point** was found near $\chi^2 \approx 3$, maximizing the balance $\epsilon \times \mathcal{P}$.
- 2D studies confirmed that **cluster multiplicity** is a strong driver of both efficiency and purity, while azimuthal modulations point to possible alignment- or acceptance-related systematics.

These results establish a **baseline validation framework** that can be systematically applied to future datasets. They also highlight where refinements are most needed, such as improved purity evaluation, higher-statistics checks at the kinematic extremes, and more precise handling of cluster quality.

Looking forward, this framework provides a foundation for several directions:

- **Physics analyses** (quarkonia, dileptons, heavy-flavor muons) will directly benefit from accurate efficiency/purity maps.
- **Detector alignment and calibration campaigns** can use these diagnostic plots to identify weak regions (e.g. ϕ -dependent structures).
- **Methodological extensions**, such as ROC curve analysis and machine-learning classifiers, promise more robust threshold optimization beyond the χ^2 discriminant.
- **Future upgrades and data-taking periods** will allow higher-statistics validation and systematic uncertainty quantification.

In this way, the work evolves from a **standalone QA tool** into a **physics enabler**, ensuring that forward-rapidity muon analyses can fully exploit the precision tracking provided by the MFT and MCH subsystems.

References

- [1] Technical design report for the muon forward tracker (mft). Technical Report ALICE-TDR-018, CERN, 2015. URL <https://cds.cern.ch/record/1981898/files/ALICE-TDR-018.pdf>. ALICE Collaboration.
- [2] Alice muon spectrometer. https://alice-collaboration.web.cern.ch/menu_proj_items/Muon-Spect, accessed 2025. Acceptance $2.5 \leq \eta \leq 4.0$ and performance overview.
- [3] ALICE Collaboration. Mft reference documentation, 2023. URL <https://aliceo2group.github.io/docs/d8/d9d/refMFT.html>.
- [4] ALICE Collaboration. Analysis framework helper task tables, 2023. URL <https://aliceo2group.github.io/analysis-framework/docs/datamodel/helperTaskTables>.
- [5] R. Frühwirth. Application of kalman filtering to track and vertex fitting. *Nucl. Instrum. Meth. A*, 262:444–450, 1987. doi: 10.1016/0168-9002(87)90887-4. URL <https://ui.adsabs.harvard.edu/abs/1987NIMPA.262..444F>.
- [6] Particle Data Group. Passage of particles through matter (review of particle physics, 2022 update). <https://pdg.lbl.gov/2022/reviews/rpp2022-rev-passage-particles-matter.pdf>, 2022.
- [7] *ROOT TEfficiency Class Reference*. CERN, accessed 2025. URL <https://root.cern.ch/doc/master/classTEfficiency.html>. Binomial efficiencies and confidence intervals.