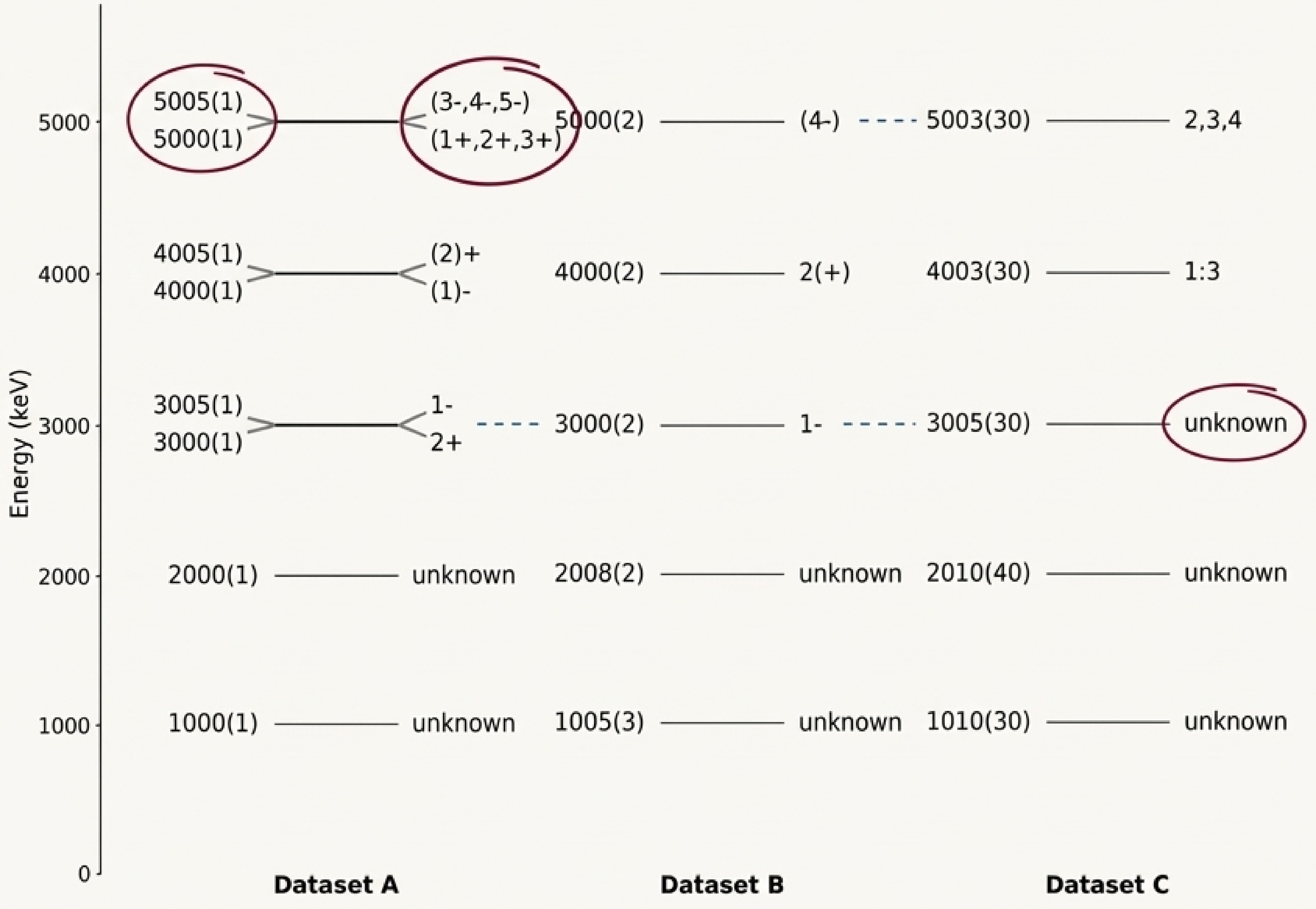
A Physics-Informed System for Nuclear Level Reconciliation

Combining Machine Learning and Rule-Based Logic to Synthesize Experimental Data

Developed by the FRIB Nuclear Data Group

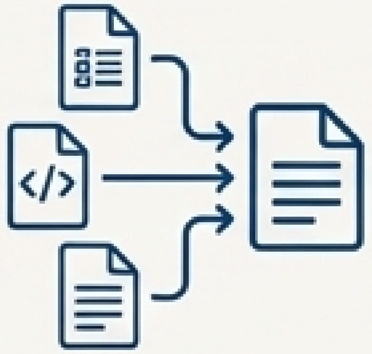
The Challenge: Reconciling Ambiguous Data from Disparate Experiments



- Nuclear physics experiments often measure the same phenomena, but produce data with slight variations and different uncertainties.
- The raw output is a set of level schemes that are similar but not identical, containing ambiguous or conflicting assignments for energy (E), spin (J), and parity (π).
- **The fundamental question:** Which observed levels across different datasets (A, B, C) are representations of the same underlying physical state? Manually resolving these conflicts is time-consuming and prone to subjective bias.

Our Solution: An Agentic, Four-Stage Reconciliation Pipeline

Ingest & Standardize



Raw JSON data from multiple experiments is loaded and parsed into a uniform structure.

Feature Engineering

Physical properties (E, J, π) are translated into a four-dimensional numerical vector that a machine can understand. *This encodes domain expertise*.

ML-Powered Inference

A physics-informed XGBoost model predicts the match probability for every possible pair of levels across datasets.

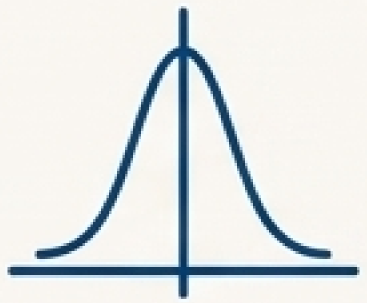
Rule-Based Clustering

A deterministic graph algorithm uses the probabilities to group levels into consistent, physically sound clusters. *This imposes logical certainty*.

This is a hybrid approach where physics expertise guides a machine learning model, and the final conclusions are drawn from pure, verifiable logic.

Stage 1: Translating Physics into a Machine-Readable Language

To compare two nuclear levels, we construct a four-dimensional feature vector. Each feature is designed to be **monotonically increasing**: a higher value provides stronger evidence for a match.

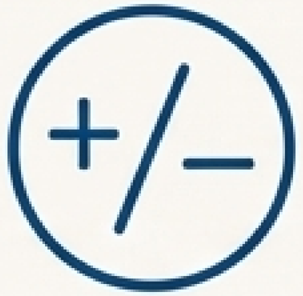


Energy Similarity

Measures the overlap of energy values, accounting for their measurement uncertainties.

Spin (J) Similarity

Scores the compatibility of spin assignments based on nuclear selection rules.



Parity (π) Similarity

Scores the compatibility of parity assignments, enforcing conservation laws.

Specificity

A penalty function that reduces confidence in matches involving ambiguous levels with multiple J^{π} possibilities.

Feature Deep Dive: Quantifying Energy Overlap and Ambiguity

Energy Similarity

A Gaussian kernel measures the energy difference (ΔE) in units of combined uncertainty (z-score).

$$\text{Similarity} = \exp\left(-0.1 \cdot Z^2\right) \quad \text{where} \quad Z = \frac{|E_1 - E_2|}{\sqrt{\sigma_1^2 + \sigma_2^2}}$$

The score decays smoothly from 1.0 (perfect overlap) towards 0.0. The `Sigma_Scale` of 0.1 is a lenient setting that tolerates separations up to ~3σ, reflecting real-world experimental variance.

Specificity (Ambiguity Penalty)

Penalizes matches involving levels with multiple possible J^π assignments. The penalty scales with the combined number of possibilities (multiplicity).

$$\text{Specificity} = \frac{1}{\sqrt{\text{multiplicity}}}$$

This formula provides a balanced penalty, naturally representing how uncertainty grows in quantum measurements. It was chosen over alternatives that were either too gentle (logarithmic) or too aggressive (reciprocal).

For each practical effect:

`multiplicity=1` → `Specificity=1.00` (No penalty)
`multiplicity=4` → `Specificity=0.50` (50% penalty)
`multiplicity=9` → `Specificity=0.33` (67% penalty)

Feature Deep Dive: Encoding Nuclear Selection Rules

For Spin and Parity, we employ an optimistic matching strategy: if any valid combination of J^{π} possibilities between two levels is compatible, we consider it evidence for a match. The final score is the maximum similarity found across all possible pairings.

Table 1: **Spin (J) Similarity Scoring**

Scenario	Example	Score	Rationale
Firm Match	2 vs 2	1.0	Strongest evidence
Tentative Match	2 vs (2)	0.9	High confidence, slight penalty
Weak Mismatch	2 vs (3)	0.2	Unlikely, but possible
Firm Mismatch	$\Delta J \geq 2$	0.0	Physics Veto. Incompatible.

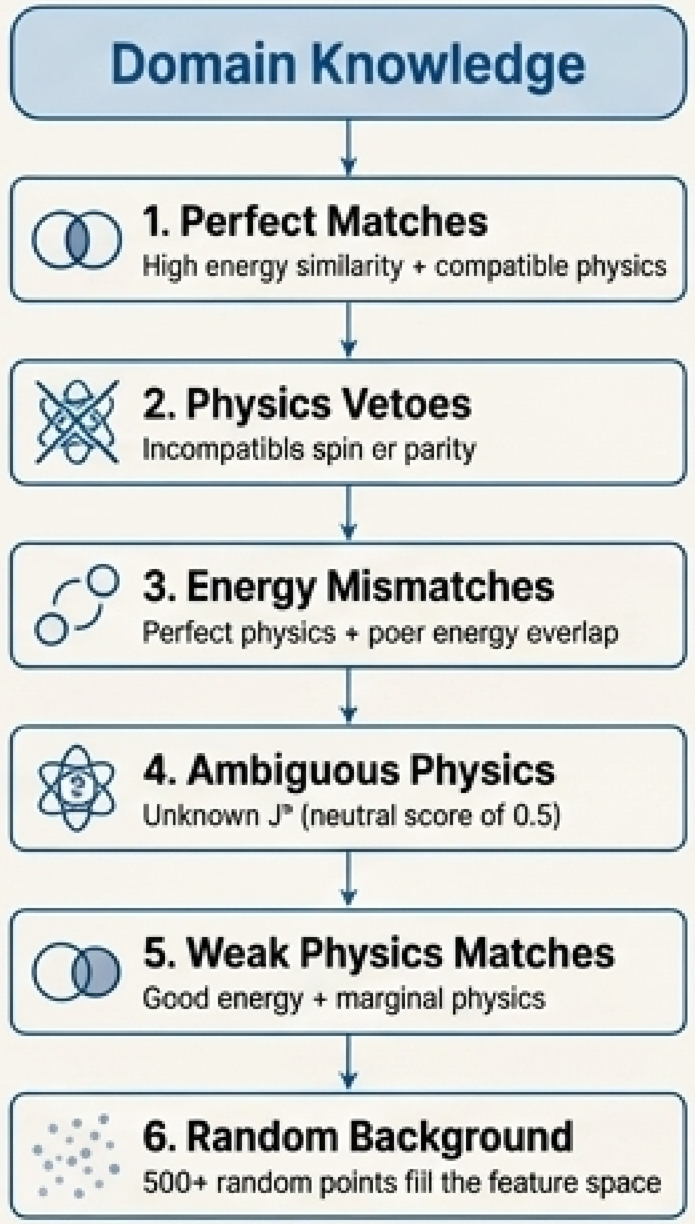
Table 2: **Parity (π) Similarity Scoring**

Scenario	Example	Score	Rationale
Firm Match	+ vs +	1.0	Strongest evidence
Tentative Match	+ vs (+)	0.9	High confidence, slight penalty
Weak Mismatch	+ vs (-)	0.1	Very unlikely, but possible
Firm Mismatch	+ vs -	0.0	Physics Veto. Incompatible.

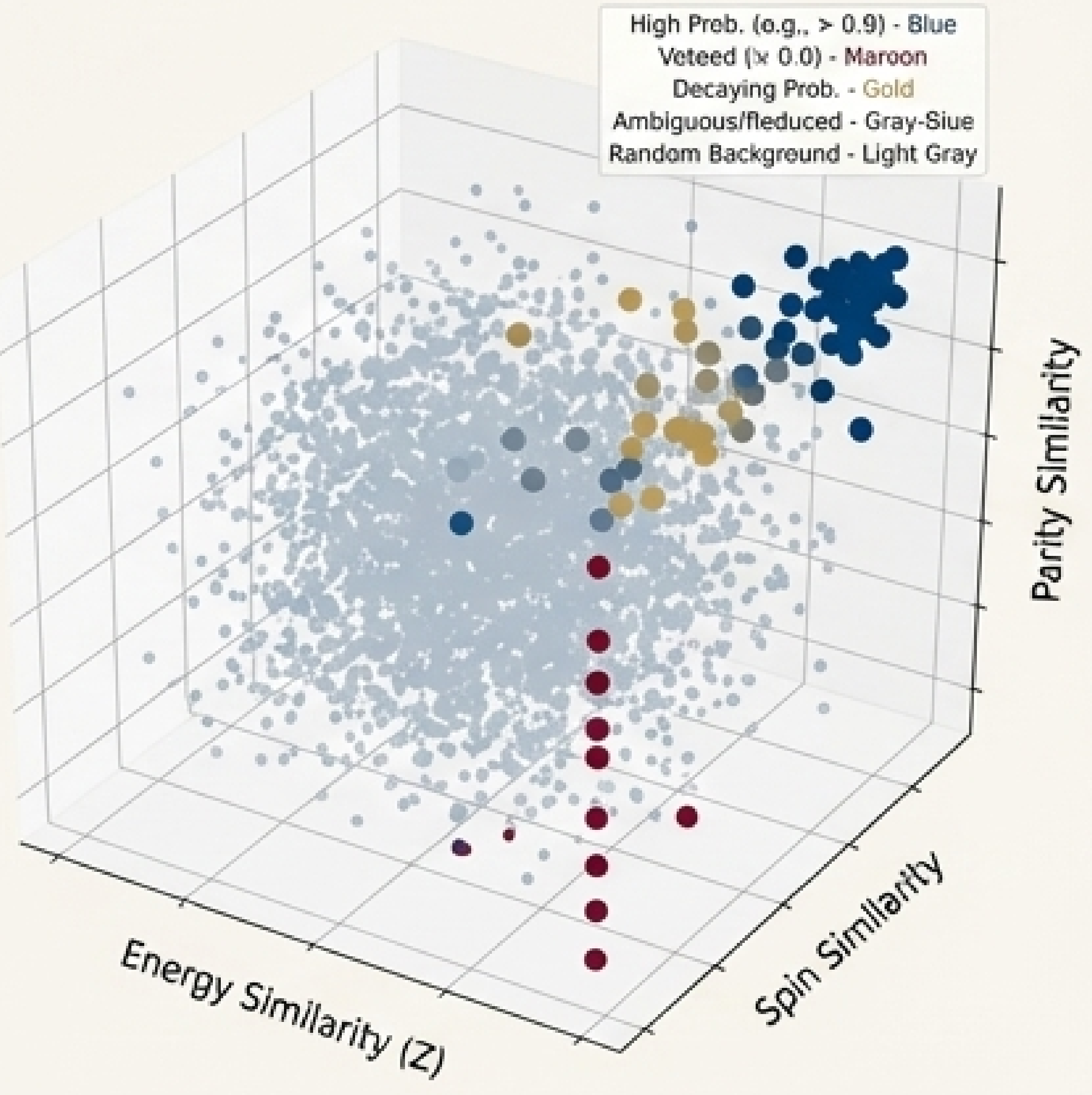
Tentativeness is not a separate feature; it is encoded directly into the similarity scores (e.g., 1.0 for firm vs. 0.9 for tentative). This makes the feature space more efficient and robust.

Stage 2: Training a Model to Think Like a Physicist

The XGBoost model is not trained on experimental data. Instead, we generate over 580 synthetic training points that explicitly encode our domain knowledge of nuclear physics. This ensures the model learns the desired physical constraints and behaviors.

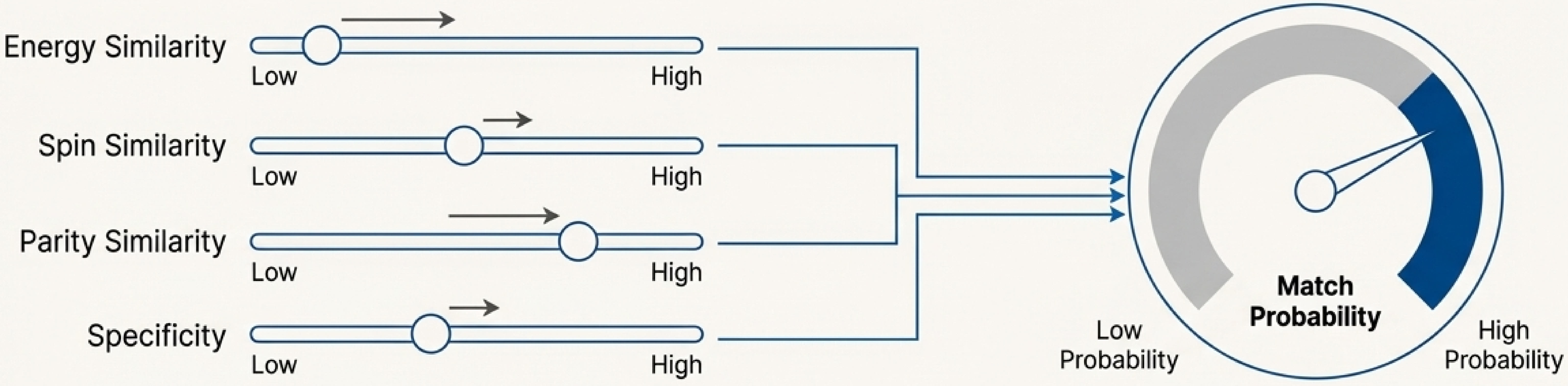


- 1. **Perfect Matches:** High energy similarity + compatible physics → **Probability > 0.9**
- 2. **Physics Vetoes:** Incompatible spin or parity → **Probability = 0.0**, regardless of energy.
- 3. **Energy Mismatches:** Perfect physics + poor energy overlap → **Probability decays to 0.0**
- 4. **Ambiguous Physics:** Unknown J^{π} (neutral score of 0.5) → **Probability depends solely on energy**
- 5. **Weak Physics Matches:** Good energy + marginal physics (e.g., ΔJ=1 tentative) → **Reduced probability**
- 6. **Random Background:** 500+ random points fill the feature space to prevent overfitting and ensure robust, rule-based labeling.



NotebookLM

Quality Assurance: Enforcing Physical Laws with Monotonic Constraints



We enforce a critical safeguard on the XGBoost model using **monotonic constraints.** This is a guarantee that the model's output will always follow physical intuition.

```
monotone_constraints = (1, 1, 1, 1)
```

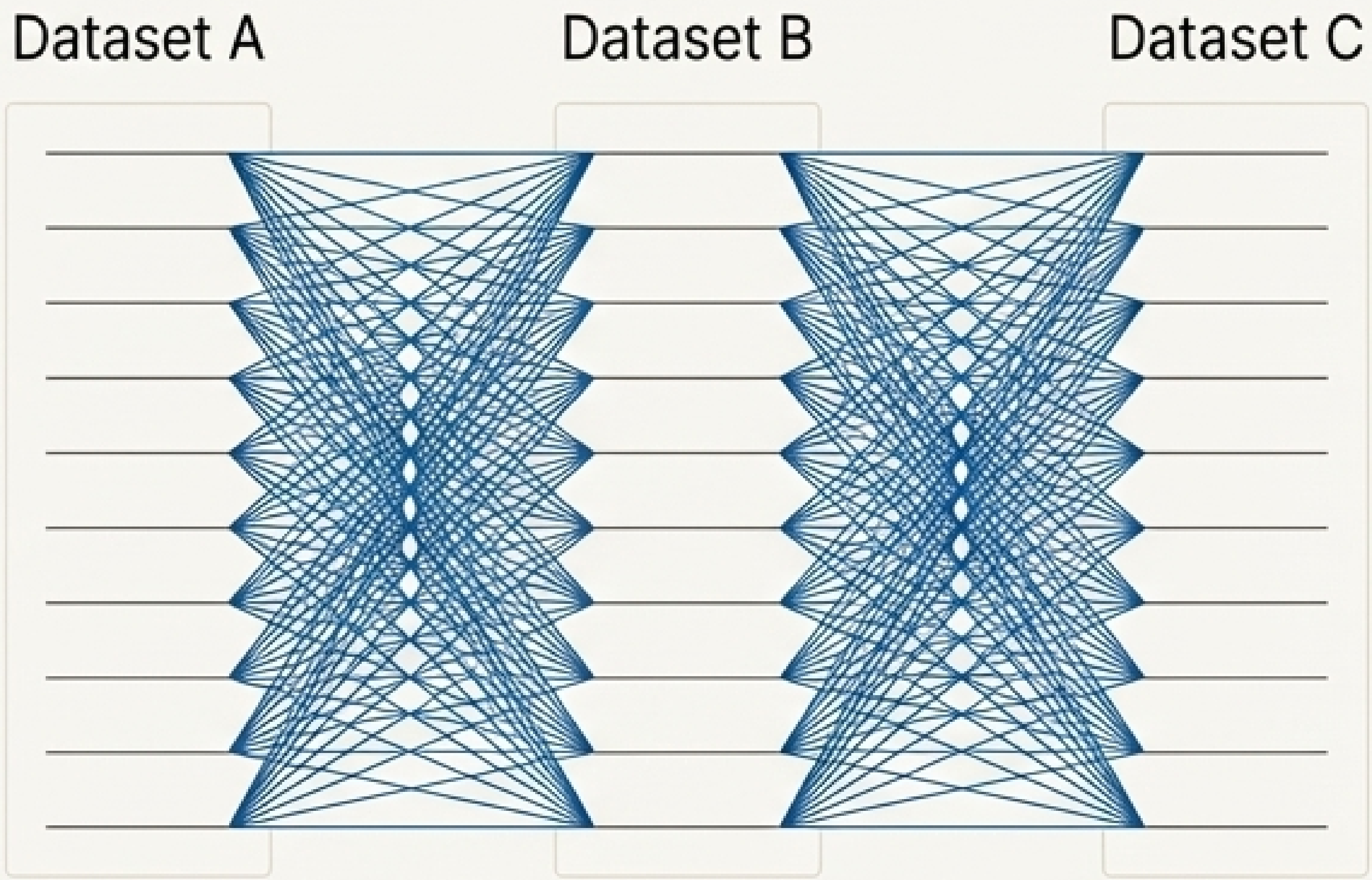
What this means:

- ↑ An increase in **Energy Similarity** can *only* increase the match probability.
- ↑ An increase in **Spin Similarity** can *only* increase the match probability.
- ↑ An increase in **Parity Similarity** can *only* increase the match probability.
- ↑ An increase in **Specificity** (less ambiguity) can *only* increase the match probability.

The model is mathematically forbidden from learning spurious, non-physical correlations. For example, it can never decide that a worse energy match should result in a higher probability score. This enforces strict intellectual honesty within the model itself.

Stage 3: Systematic Pairwise Inference at Scale

- **Exhaustive Comparison**: The system iterates through every possible cross-dataset pair of levels (A vs B, A vs C, B vs C). Same-dataset pairs are skipped.
- **Feature Extraction**: For each pair, the four-dimensional feature vector is constructed.
- **Prediction**: The trained, constrained XGBoost model predicts a match probability between 0.0 and 1.0.
- **Output**: All pairs with a probability greater than a low threshold (1%) are written to `level_pairs_inference.txt`, creating a comprehensive ranked list of potential matches for the next stage.

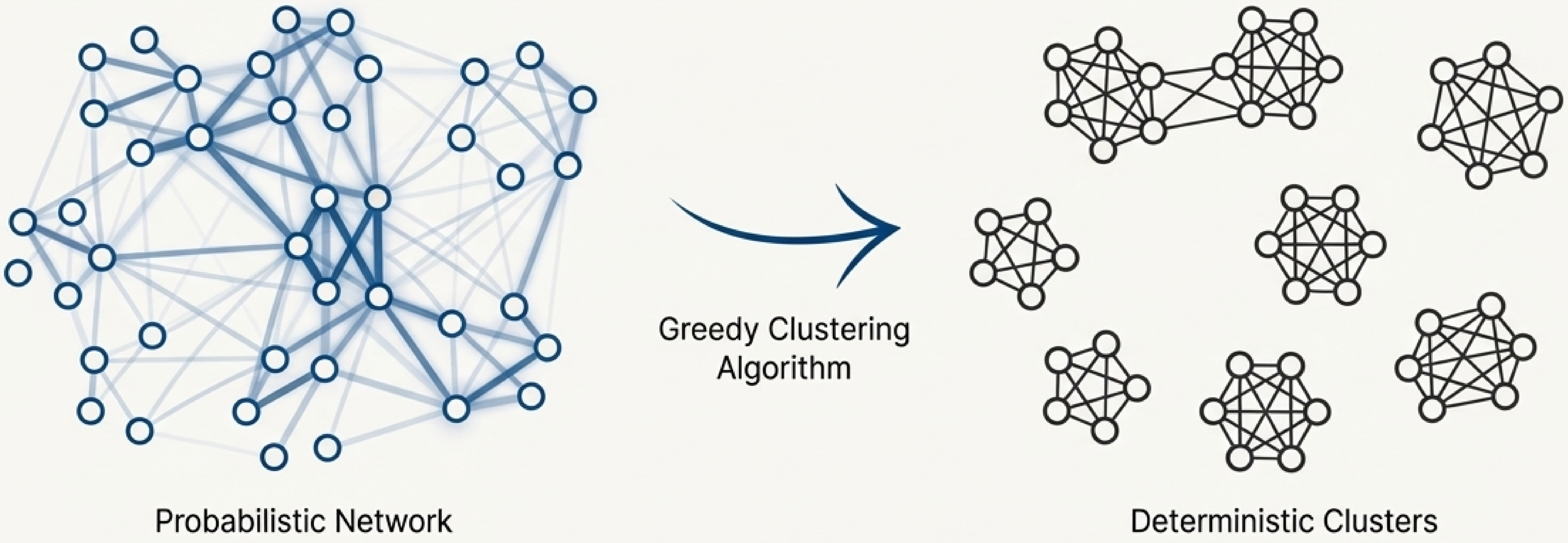


A_3000 <-> B_3000 | Probability: 98.5%
Features: Energy_Sim=0.99, Spin_Sim=1.00, Parity_Sim=1.00, Specificity=1.00

A_5005 <-> B_5000 | Probability: 85.1%
Features: Energy_Sim=0.91, Spin_Sim=0.90, Parity_Sim=0.90, Specificity=0.71

Stage 4: Rule-Based Synthesis—From Probabilities to Definitive Clusters

The final step is not based on machine learning. It is a deterministic graph clustering algorithm that uses the ML-generated probabilities as input to perform logical grouping. This ensures the final "Adopted Levels" are the product of verifiable rules, not a black box.

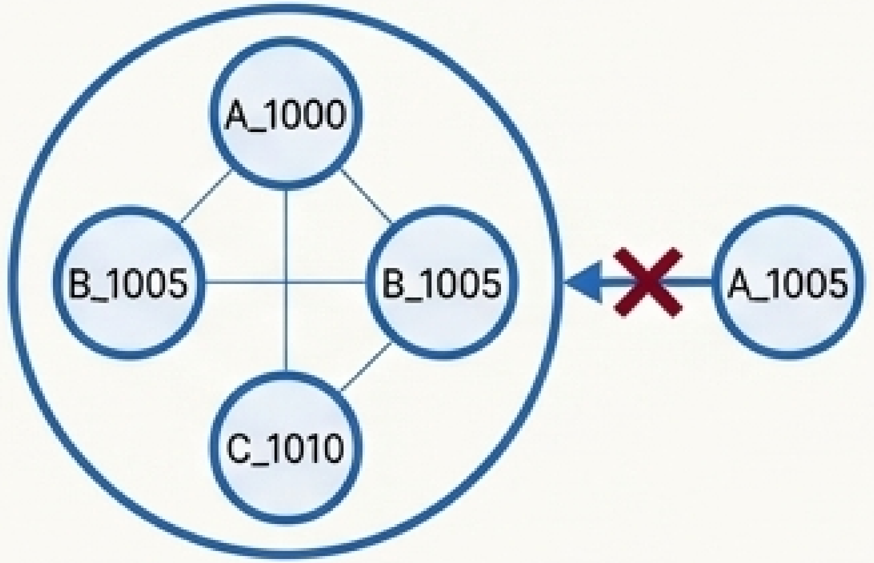


Method: Greedy cluster merging.
Order: It processes the level pairs in descending order of match probability, starting with the most confident matches first.
Threshold: Only pairs with a probability above a strict threshold (clustering_merge_threshold = 0.15) are considered for merging.

The Logic of Clustering: Certainty Through Unbreakable Rules

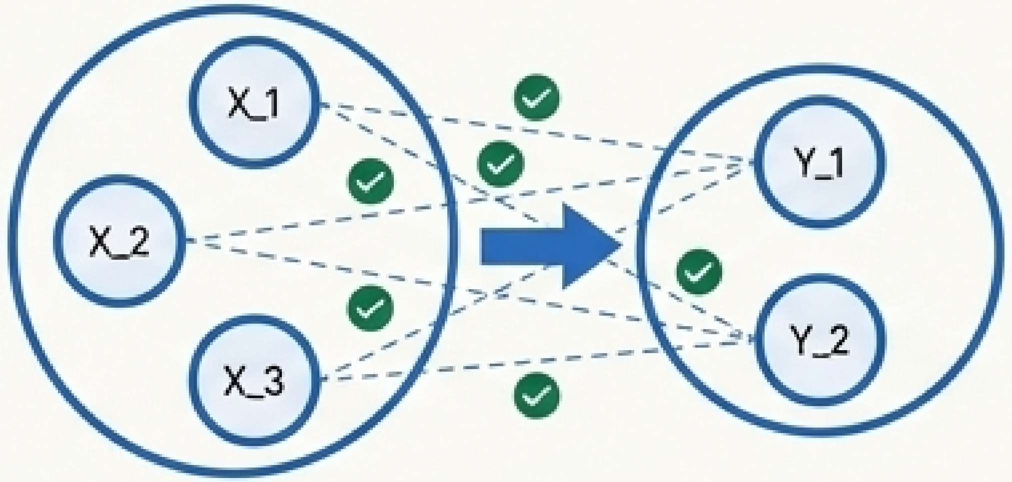
The greedy merging algorithm is governed by three strict constraints to ensure every cluster is physically and logically sound.

1. Dataset Uniqueness



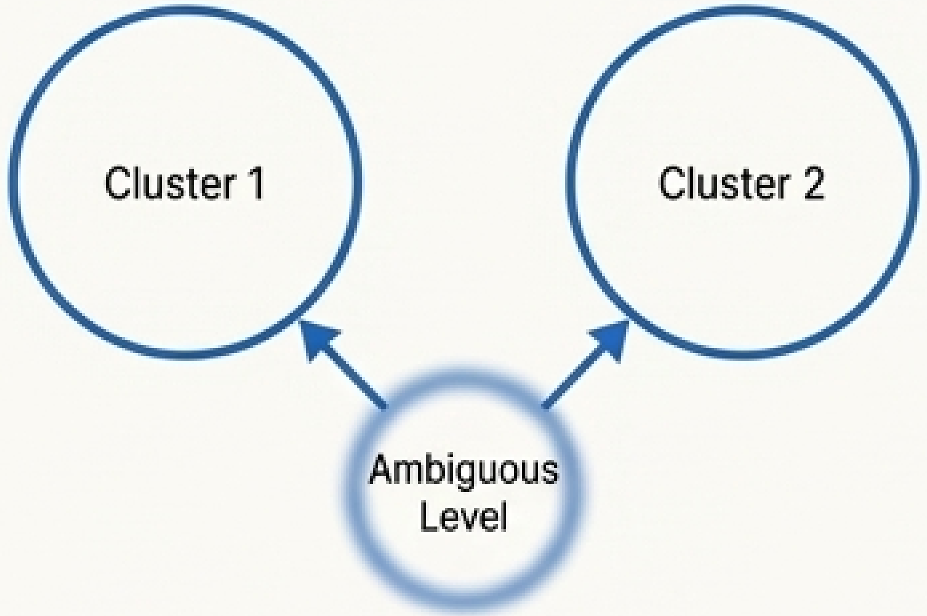
A cluster can contain at most one level from any given dataset (e.g., a cluster cannot contain both A_1000 and A_1005). This is the fundamental constraint.

2. Mutual Consistency



Before merging two clusters, the algorithm verifies that *every member* of the first cluster is compatible (above the 15% threshold) with *every member* of the second. This creates a fully connected, clique-like structure.

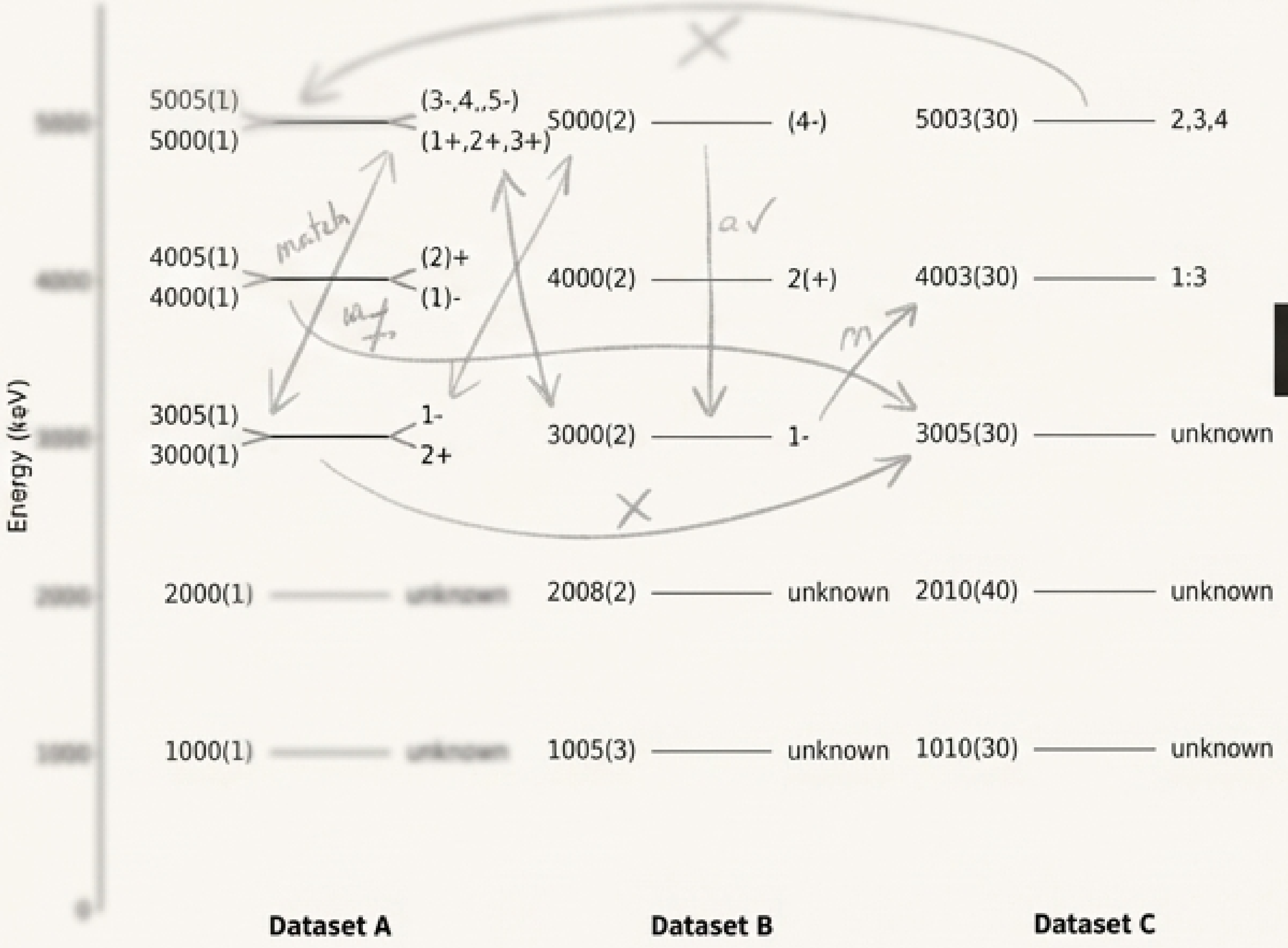
3. Overlap Support for Ambiguity



If a merge is blocked by dataset uniqueness, the system doesn't just fail. It checks if an ambiguous level can be logically added to multiple clusters, reflecting genuine physical ambiguity. This allows the system to model complex scenarios without violating core principles.

The Result: From Ambiguity to Actionable Insight

Initial State: Disparate & Ambiguous



Final Result: Reconciled & Anchored Clusters

```
Cluster 5:
  Anchor: B_3000 | E=3000.0±2.0 keV | Jπ=1-
  Members:
    [A] A_3000: E=3000.0±1.0 keV, Jπ=1- (Match Prob: 99.2%)
    [B] B_3000: E=3000.0±2.0 keV, Jπ=1- (Anchor)
    [C] C_3005: E=3005.0±30.0 keV, Jπ=unknown (Match Prob: 88.4%)
```

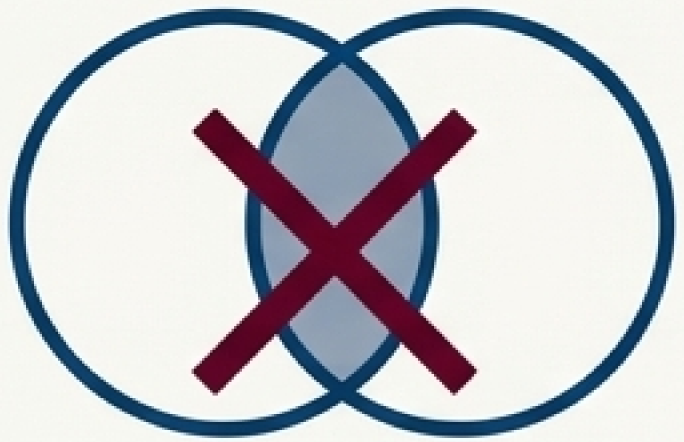
Key Logic:

Anchor Selection: The level with the *smallest energy uncertainty* is automatically chosen as the anchor, representing the most precise measurement.

Reporting: All other members are reported with their properties and their calculated match probability relative to the anchor.

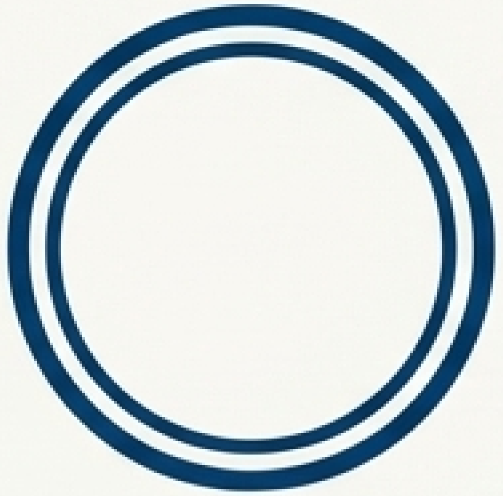
Acknowledging Limitations & Design Choices

A robust system is defined as much by what it does as by what it is designed *not* to do. Our approach includes several deliberate choices to maintain focus and avoid redundancy.



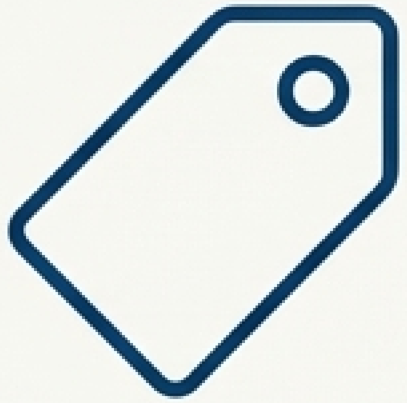
Removed Redundancy

Early versions included separate '**Spin Certainty**' and '**Parity Certainty**' features. These were removed because tentativeness is more effectively encoded directly into the similarity scores (1.0 vs 0.9). This simplifies the model and prevents feature redundancy.



Current Scope

The model's decisions are based exclusively on **energy**, **spin**, **parity**, and their ambiguities. It does not yet incorporate other crucial physical observables.



Interpreting Probabilities

The model's output is a **confidence score**, not a formal statistical probability. It is a powerful tool for ranking and guiding expert evaluation.

The Path Forward: Incorporating Deeper Physics

The current architecture is designed to be extensible. Future work will focus on incorporating additional physical observables to further increase the system's resolving power and accuracy, requiring extensions to the underlying data schema.

Future Feature	Physics Value	Implementation Strategy
Gamma-Ray Branching Ratios	**Critical:** Can distinguish close-lying states with identical J^{π}.	Compare decay intensity patterns using cosine similarity of branching ratio vectors.
Half-Life / Lifetime	**High**: Orders of magnitude differences provide a definitive veto.	Use Gaussian similarity on log-scale lifetimes to compare values.
Band Assignment	**Medium**: Useful for identifying members of the same rotational band structure.	Add a feature to check for and reward shared band assignments.