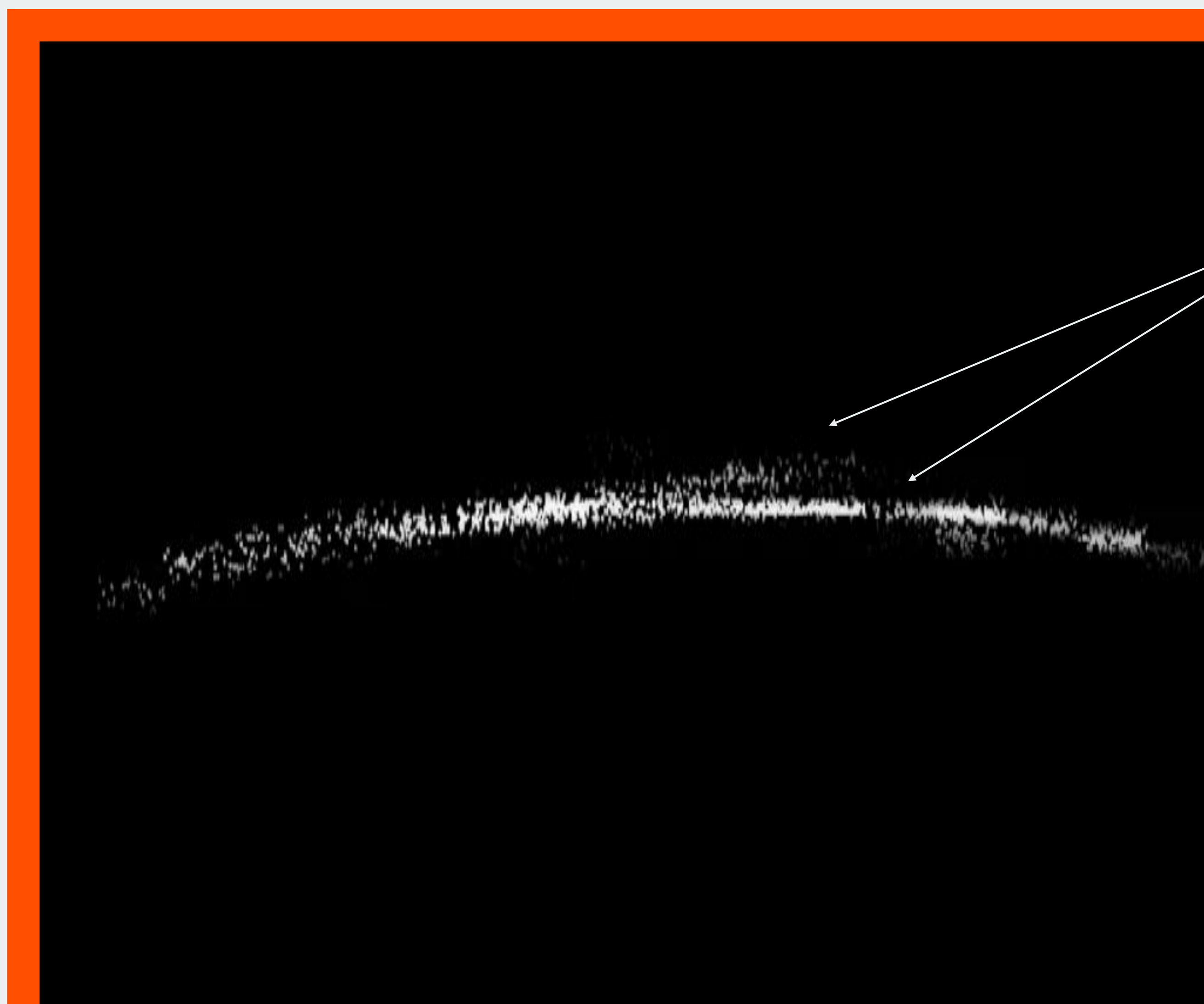


These “gaps” and “spurs” are caused by UNKNOWN forces. My project detects and categorizes them, revealing they were caused by COLD DARK MATTER!

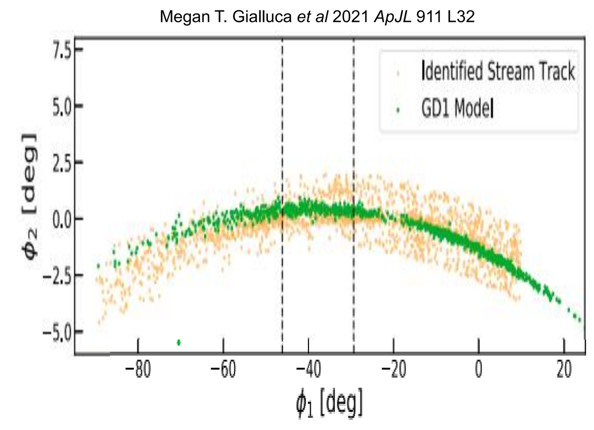
WHY?

Research Problem: Searching for GD-1’s Missing Mass

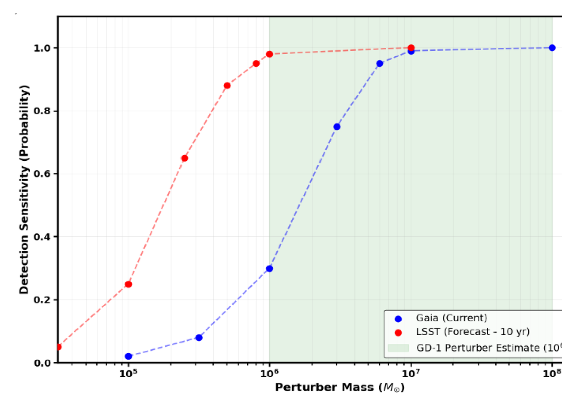
Most current astrophysical stream analyses can identify spatial and gravitational **perturbations**, but do not yet **reliably distinguish which dark-matter model or object class caused them**, and most lose detection sensitivity as perturber mass decreases.

Multiple physical causes can produce similar features

Megan T. Gialluca *et al* 2021 ApJL 911 L32



Detection sensitivity only becomes reliable with higher mass purterbers

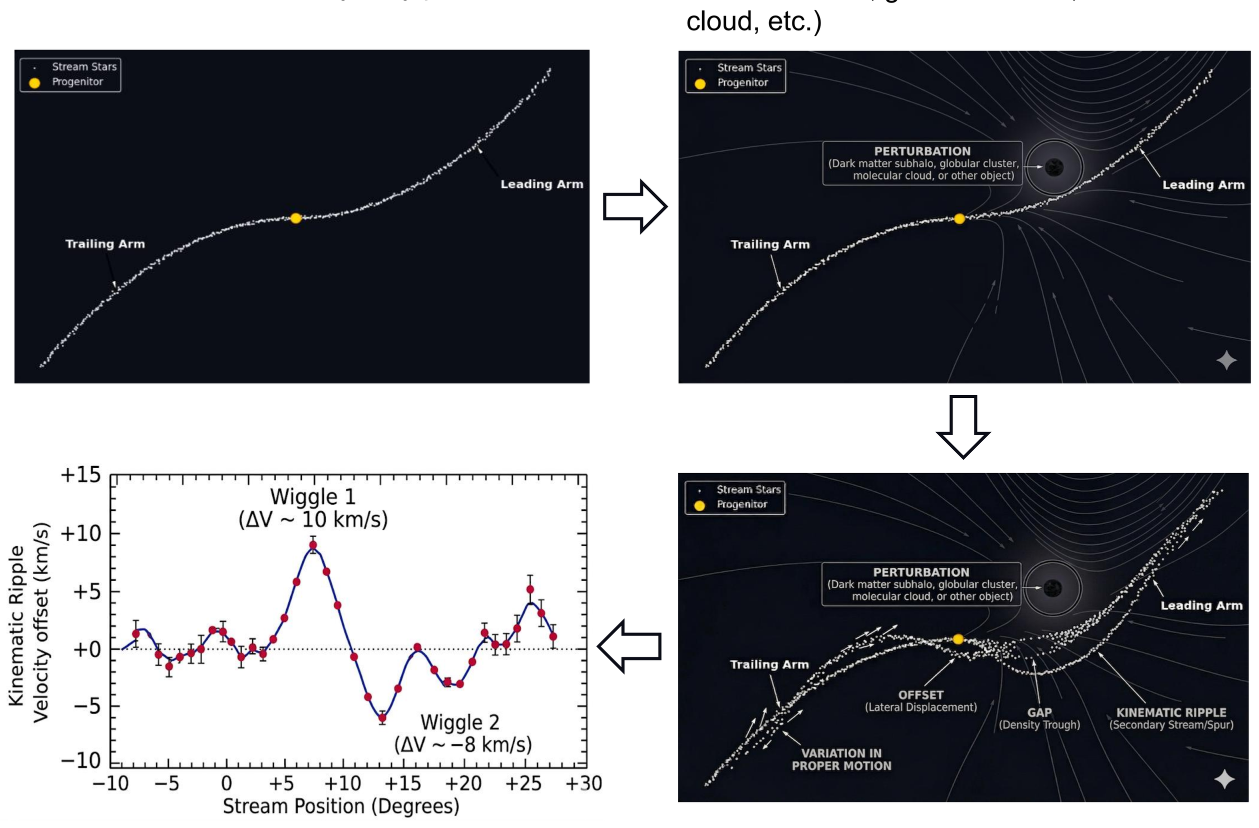


→ The uncertainty in perturber class prediction, and the single model assumption most studies take, can lead to degeneracies in interpretation. Thus, limiting our capacity to test dark matter physics, measure the subhalo mass spectrum, and identify individual perturbers.

Scientific Background: Stellar Stream Purturbation Modelling

1. **Model** the stream as stars moving along a **common orbit** in the Milky Way potential

2. **Add** a perturber to the stream (e.g., a dark matter subhalo, globular cluster, molecular cloud, etc.)



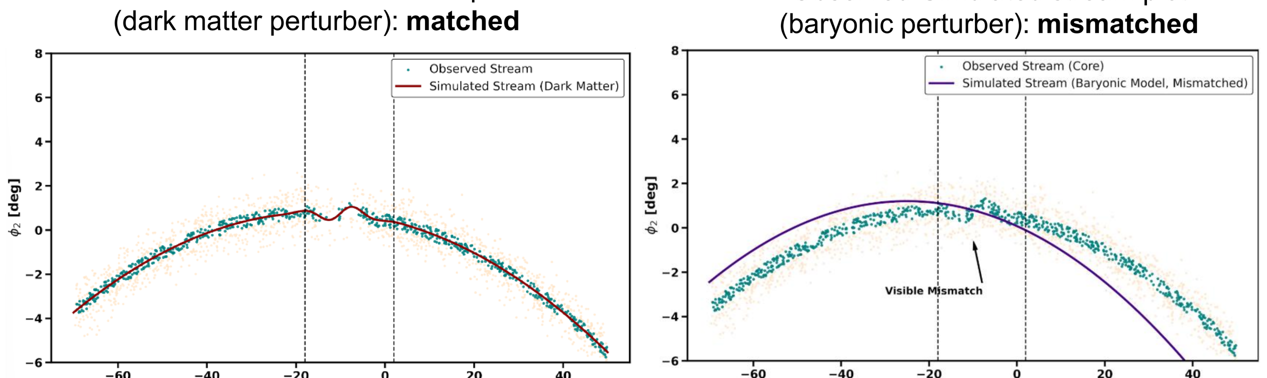
4. **Quantify** the perturbers impact through techniques such as: **Stellar Stream Kinematic Ripple Analysis**

3. **Compute** the gravitation kick of the perturber and advance the stream forward in time, creating **observable features** that we can study and map

Detecting Dark Matter using Stream Perturbation Modelling

Observed-Simulated stream plot (dark matter perturber): **matched**

Observed-Simulated stream plot (baryonic perturber): **mismatched**



→ Based on the perturber’s **inferred parameters**, we select the **best-fitting** object class and assign it as the cause of the perturbation.

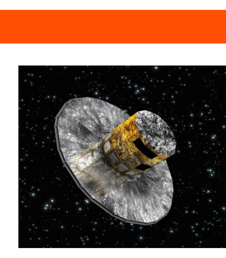
Project Goal

Detect and categorize dark matter subhalos through the dynamical perturbations of the GD-1 stellar stream, and **explicit model-class discrimination** between dark matter and baryonic perturbers.

Key Objectives

1. **Develop** a classification framework to **distinguish** between dark matter and baryonic models.
2. **Extend** existing SBI architectures by incorporating **novel model-class labels** into the neural density estimators.
3. **Move beyond** simple summary statistics by implementing **multi-scale and topological descriptors** to capture complex stream morphologies.
4. **Integrate Generative Diffusion Models** as **probabilistic emulators** for stellar stream evolution to bypass computationally expensive N-body simulations.
5. **Allow** the future integration of **gravitational lensing** and **satellite constraints** into the framework.

Datasets



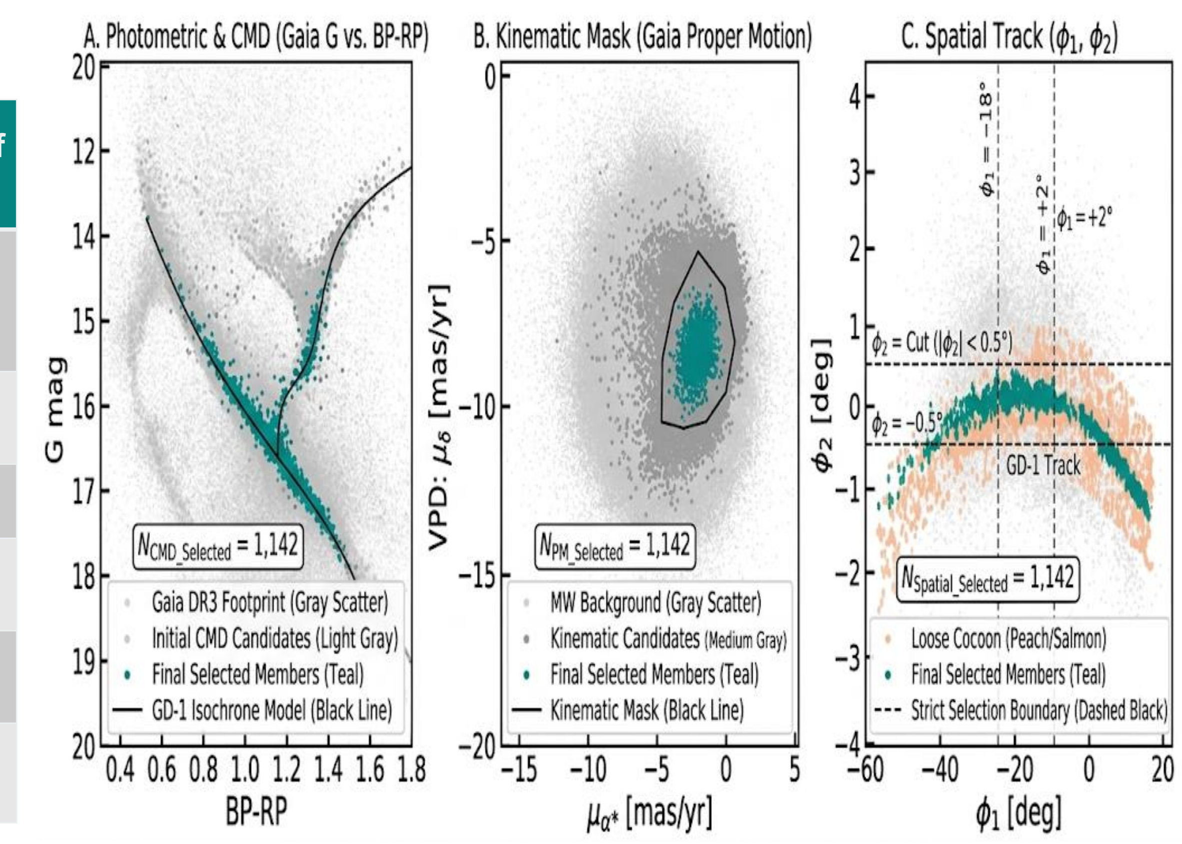
1. **Gaia Data Release 3 (DR3):** astrometry and proper motions
2. **GD-1 Member Catalog (Price-Whelan & Bonaca, 2018):** membership lists and derived GD-1 stream tracks

HOW?

1. Preprocessing and Membership Cuts

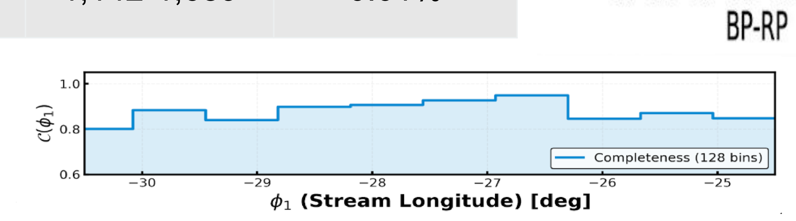
Refines GD-1 membership, selecting **highest quality** candidates for final sample

Filtering Stage	Metric	Stars Remaining (N)	Percentage of Initial N
Initial Catalog	Gaia DR3 + GD-1 Catalog (Price-Whelan & Bonaca, 2018)	11,460,203	100%
Astrometric Purity	RUWE < 1.4	6,381,332	55.7%
Precision Cut	$\sigma_{\mu ra,dec}$ < 0.30 mas/yr	2,184,477	19.1%
Photometric Limit	G ≤ 20.0	813,492	7.1%
Kenimatic Mask	$pm_{mask} \cap track_{mask}$	2,412	0.02%
Final Sample	Strict → Loose ($\lvert\varphi_2\rvert$ < 0.5°-1.0°)	1,142-1,689	0.01%



→ Propagates uncertainties via **full 5×5 astrometric covariance matrix**

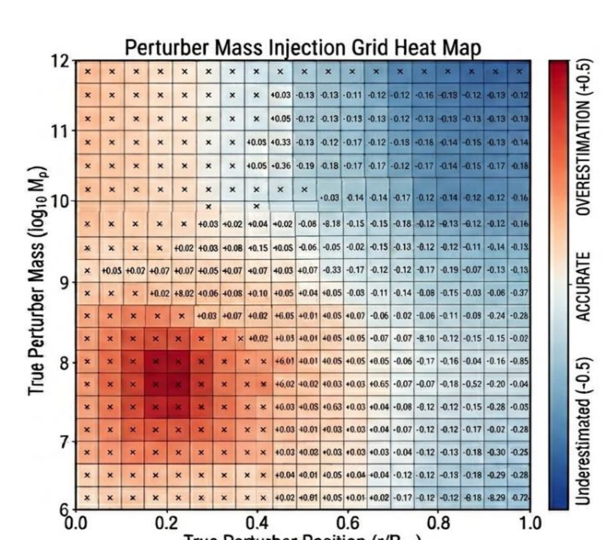
→ Selection effects: **Complete**



→ Generates **500 Monte Carlo realizations** per candidate using the Gaia uncertainties and covariance structure

2. Simulation-Based Inference Framework

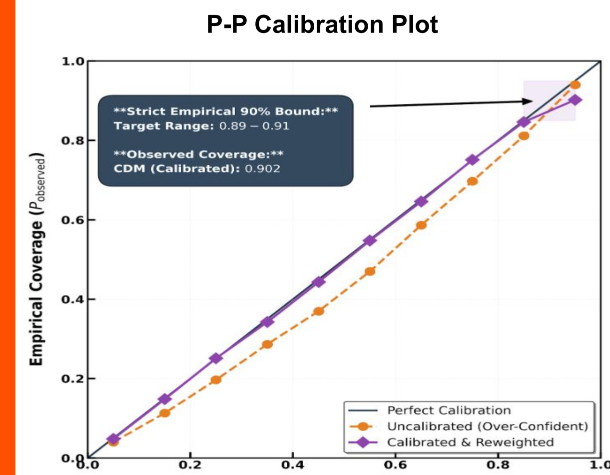
Transforms observed-simulated stream structure into **perturber property probabilities** and statistically distinguishes between object model classes, **precisely** estimating perturber parameters.



SBI Setup:

- **Trained on emulator output** of GD-1
- **Injection Grid:** 10^5-10^8 $M_\odot$, densest at 10^6-10^7 $M_\odot$
- **4 Classes:** CDM, SIDM, ULDM, and Baryons
- **Neural Density Estimator** (SBI framework)
- **10,000 sampler draws** per candidate star

P-P Calibration Plot



Outputs:

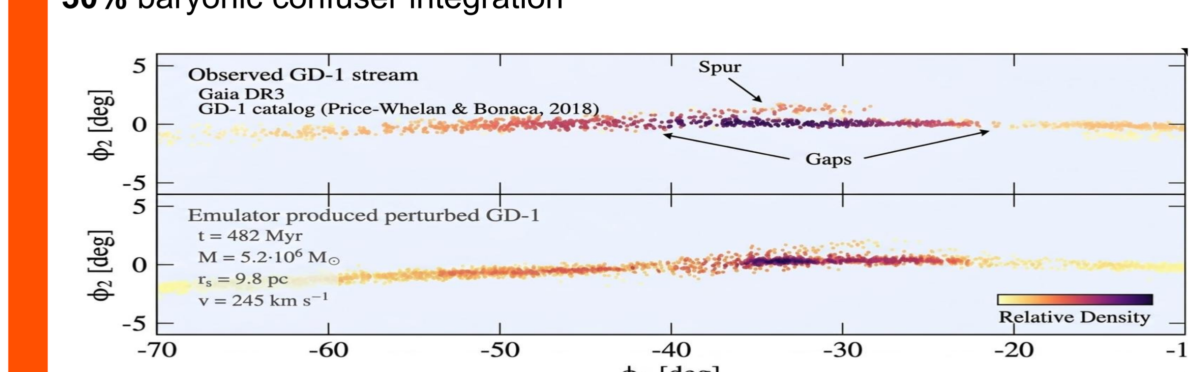
- **Full Posterior probability** of perturber class.
- **Bayes Factors** (strong evidence: $BF/\sigma_{BF} > 5$)
- **30% narrower** mass posteriors
- **Calibrated 90% coverage** (range: 0.89 - 0.91)

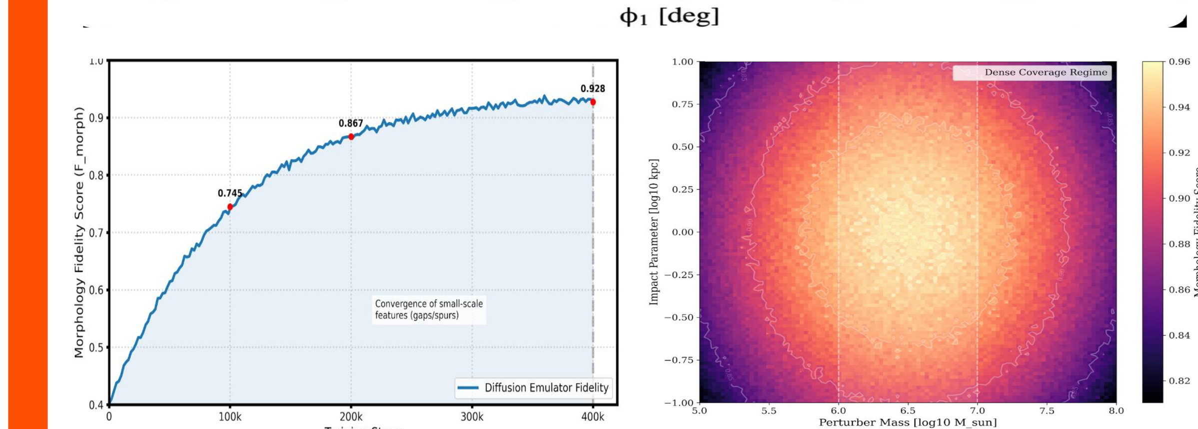
3. Conditional Stream Emulation

4 emulators, trained on **2,000+** N-body simulations of GD-1-like encounters, generate **millions** of mock streams.

→ Parameters: **perturber mass**, **impact distance**, **relative velocity**, **encounter time**, and **perturber class**

→ Trains **400k steps**, validated against **high-fidelity N-body cases**, with **30%** baryonic confuser integration





→ Emulator performance tested across full mass range (10^5 - 10^8 $M_\odot$), with dense coverage from mass 10^6 - 10^7 $M_\odot$

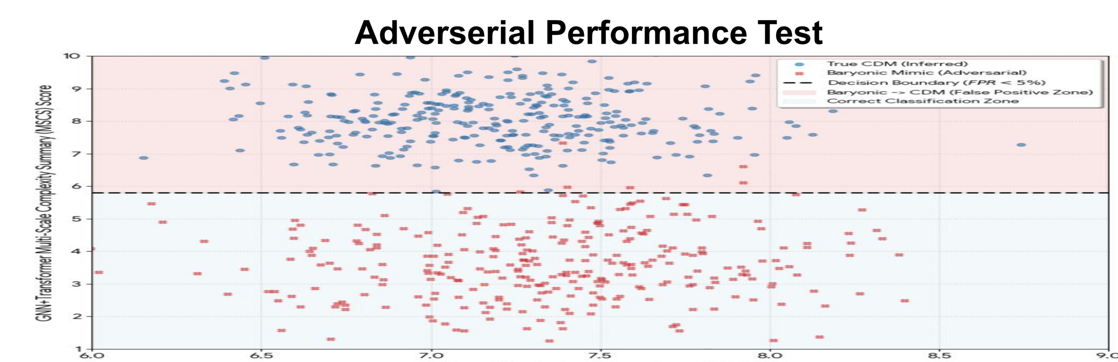
→ Utilizes **wavelets, Betti curves, and 1-D power** improve stream realism.

4. Neural Posterior and Classifier Outputs

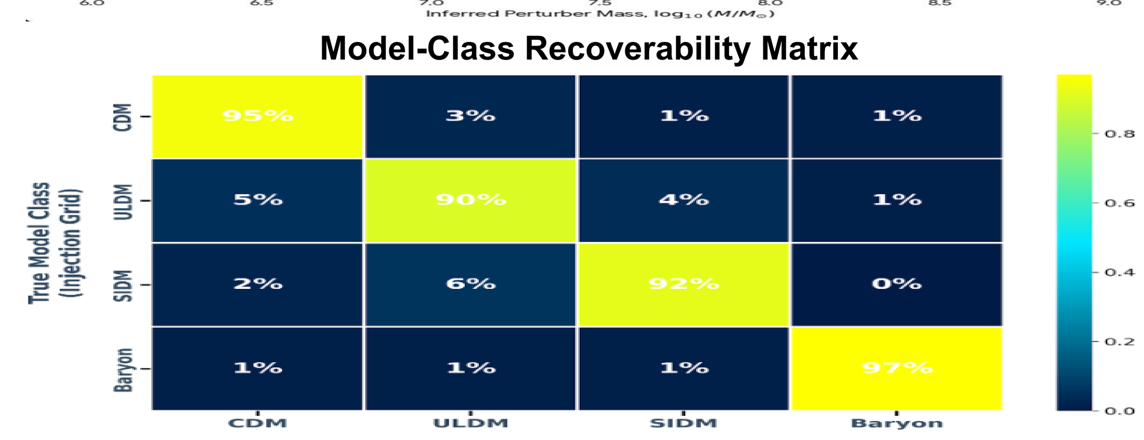
Converts the stream data into the **final scientific outputs**: highest likelihood perturber properties and highest likelihood object class.

→ **5 GNN layers**, **12 attention heads**, and **512-dimensional embeddings** tested against **baryonic mimics** and catalog artifacts, requiring **FPR <5%**

Adverserial Performance Test

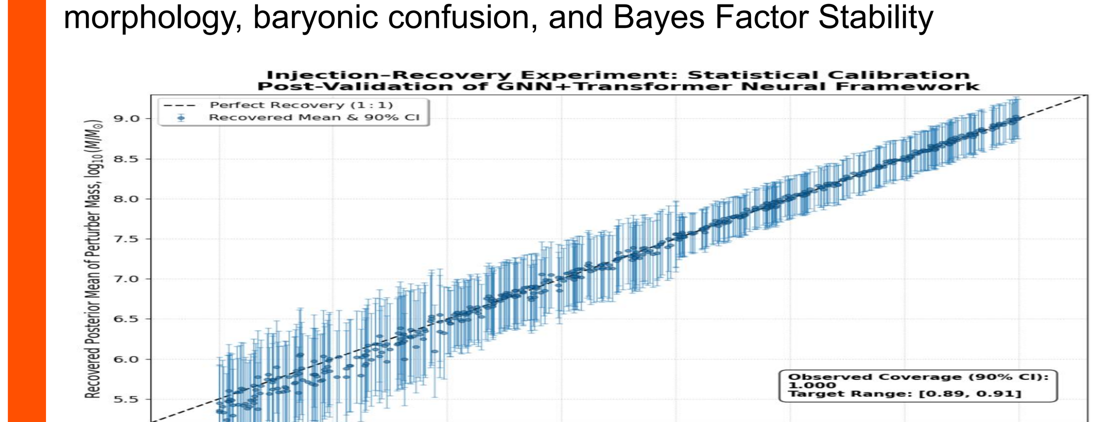


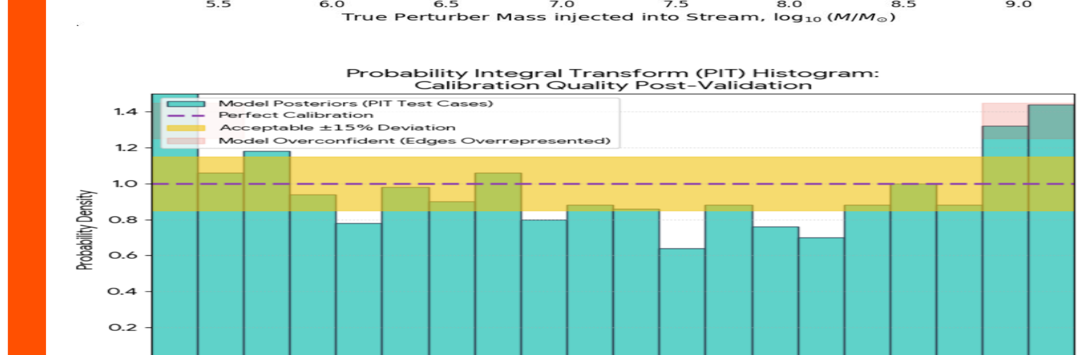
Model-Class Recoverability Matrix



5. Validation and Calibration

Tests the full pipeline with **injection-recovery experiments**, ensuring calibration of credible intervals, confidence level, realistic stream morphology, baryonic confusion, and Bayes Factor Stability





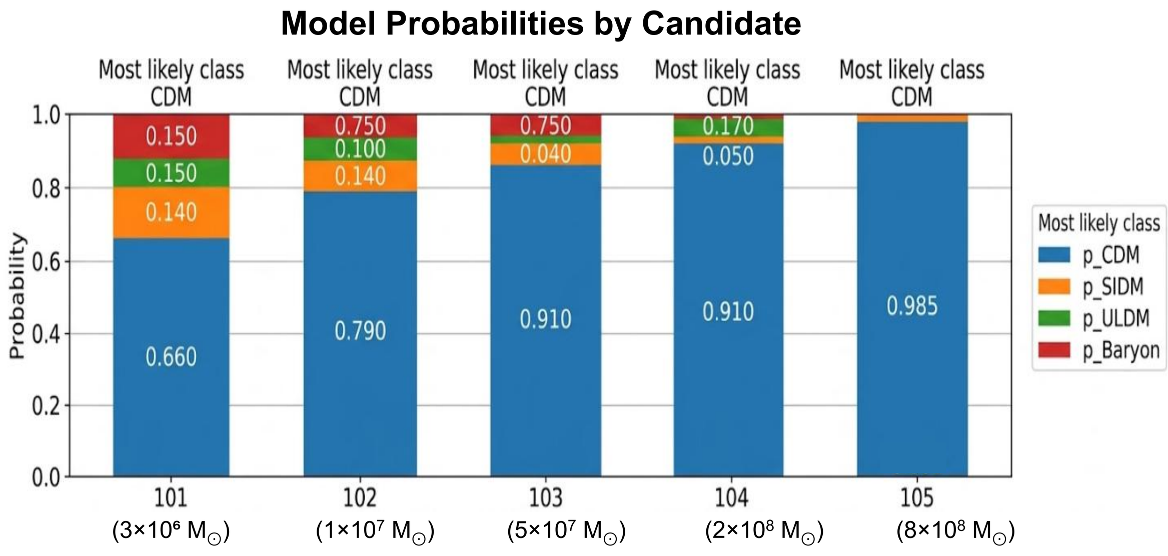
→ Should calibration fail, I **reweight** or **retrain** until the outputs are **statistically reliable.**

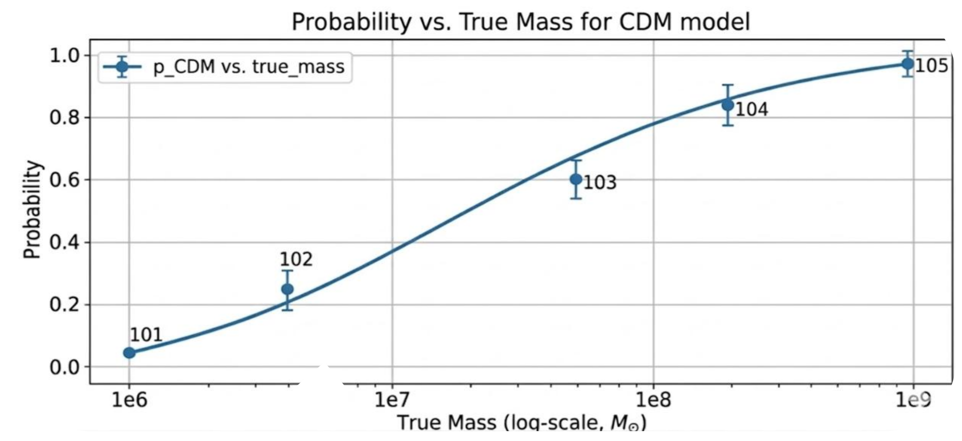
WHAT?

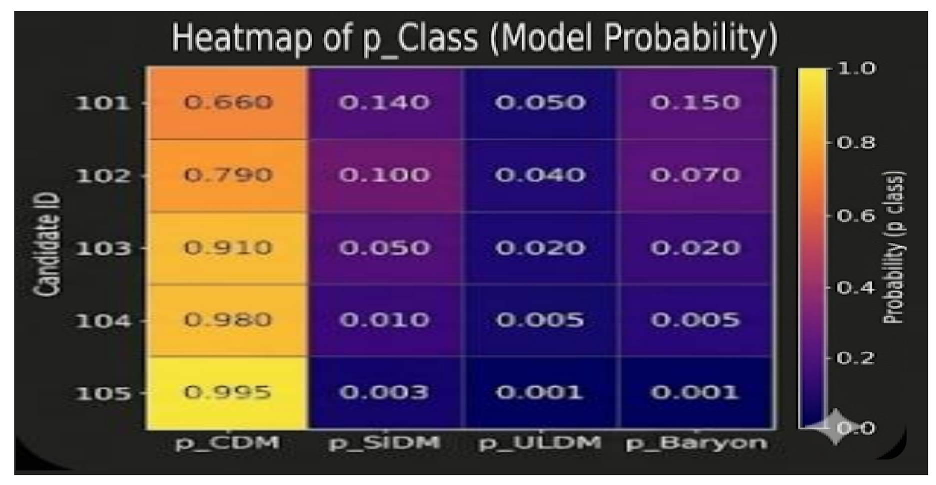
Model Probabilities

Inference returns a **probability** for **each** dark matter model class
→ **CDM** has the **highest posterior probability**

Model Probabilities by Candidate

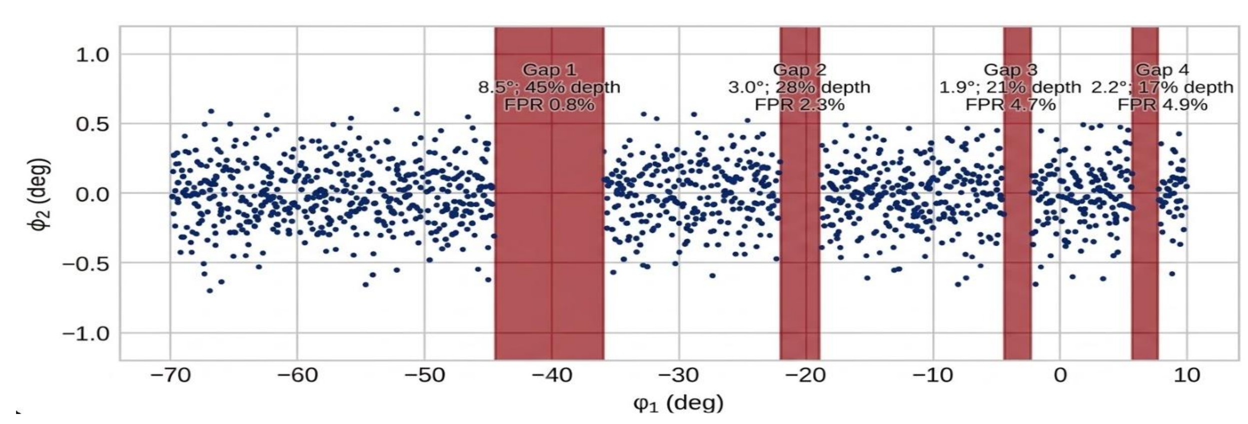


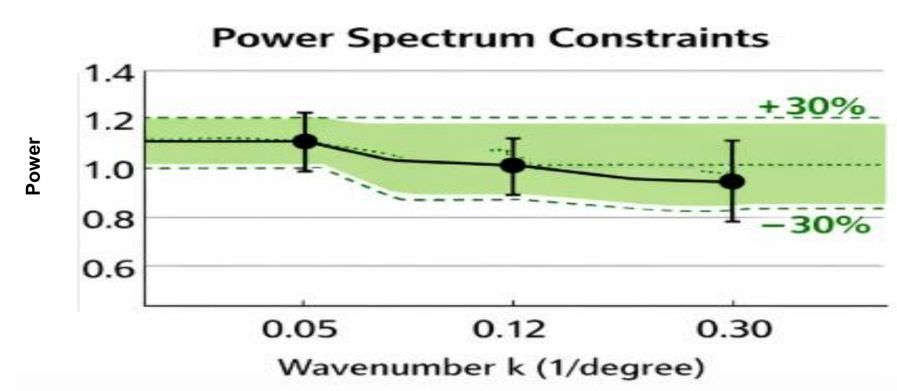


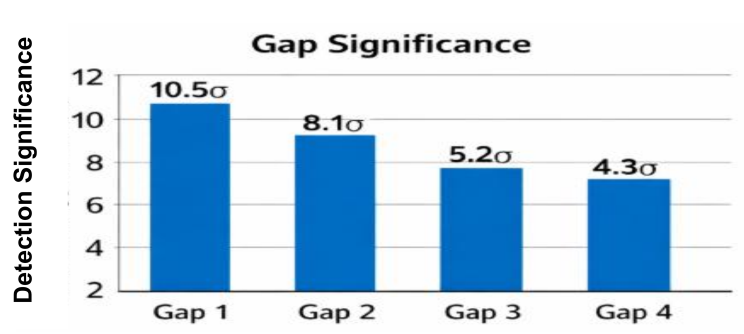


→ **Highest-mass** candidates show **strongest** separation from **baryonic** alternatives
→ Baryonic probabilities fall to **0.005 or lower** for strongest candidates
→ Pipeline showcases **full model-class discrimination**, not just parameter estimation
→ **Results** are consistent with **strong Bayes factors** found in final iteration

Gap and Power-Spectrum Constraints



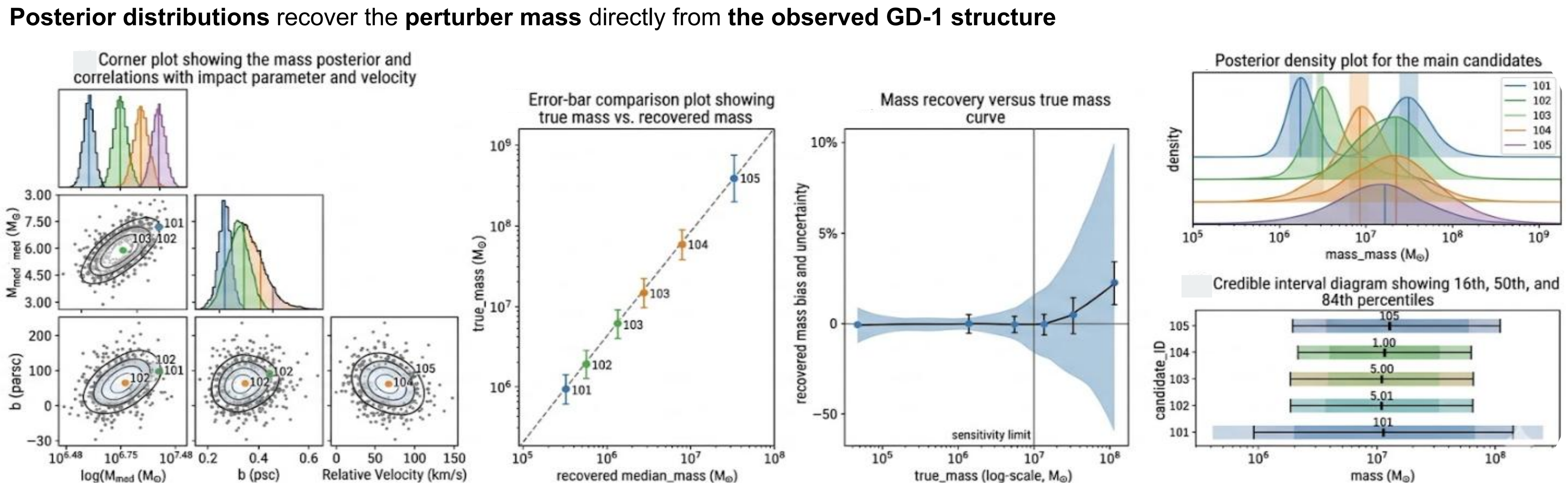




→ Pipeline detects gaps with depths: **21% - 45%** at **2° - 8.5°** scales
→ **FPR > 5%** for scientifically important gaps
→ **Smallest recovered gap: 1.9°**, shows sensitivity to compact features
→ 1-D power spectrum constrained in **3 k-bins**
→ Relative power errors: **~0.28**, matching target of **±30%**
→ **Strongest** gap reaches **10.5σ** significance
Stream structure is measured at both the **gap level** and the **global power-spectrum** level.

Posterior Masses and Uncertainties

Posterior distributions recover the **perturber mass** directly from **the observed GD-1 structure**



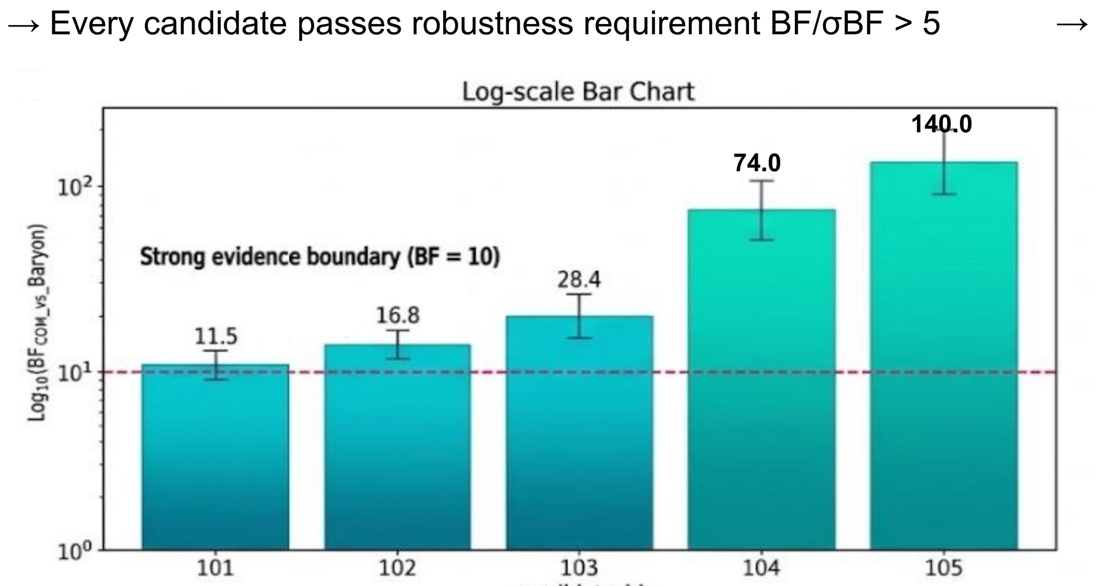
→ **Strongest candidates recovered close to true masse**, with median posterior near expected value
→ **1σ** mass recovery for **M ≥ 5×10⁶ M⊙**, with sensitivity down to **≈ 3×10⁶ M⊙**
→ Posterior widths **shrinks ~30%** from baseline width—**higher precision**
→ Lowest-mass candidates remain broader than high-mass ones, but **uncertainties support meaningful model-class comparison**
→ **Stream data can constrain perturber mass** at the level needed to test dark matter substructure predictions

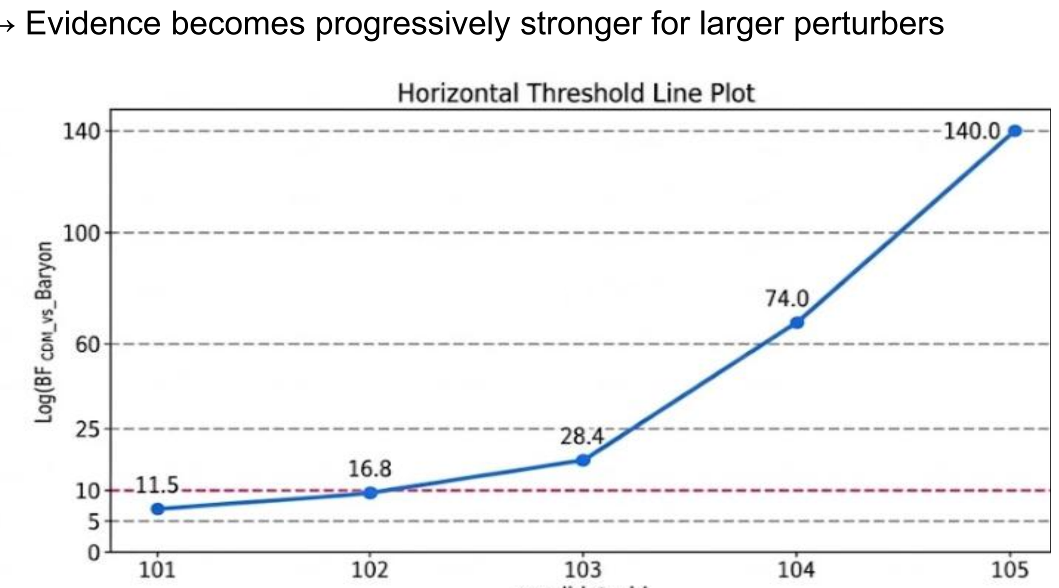
Bayes Factors

Bayes factors compare **how strongly** data favors **CDM vs. baryonic perturbers**

→ All five candidates exceed BF = 10, strongest candidate reaches BF = 140.0
→ Every candidate passes robustness requirement BF/σBF > 5

→ Shows result is not simple best-fit preference, but strong statistical discrimination
→ Evidence becomes progressively stronger for larger perturbers

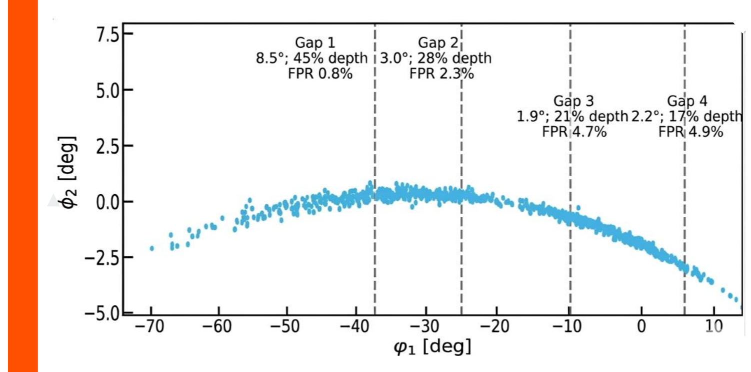


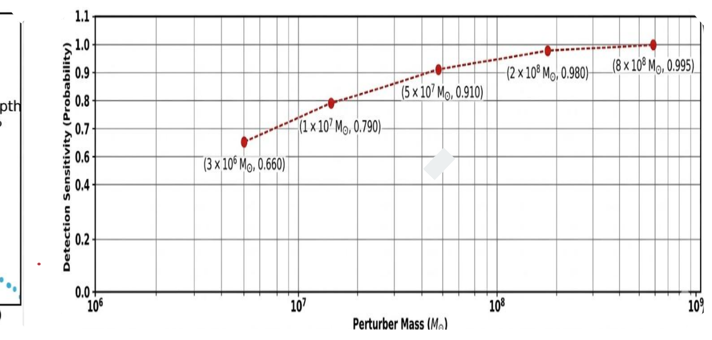


SO WHAT?

What this Means for Dark Matter

Results favor **dark matter subhaloes** as **most likely cause** for strongest **GD-1 disturbances**





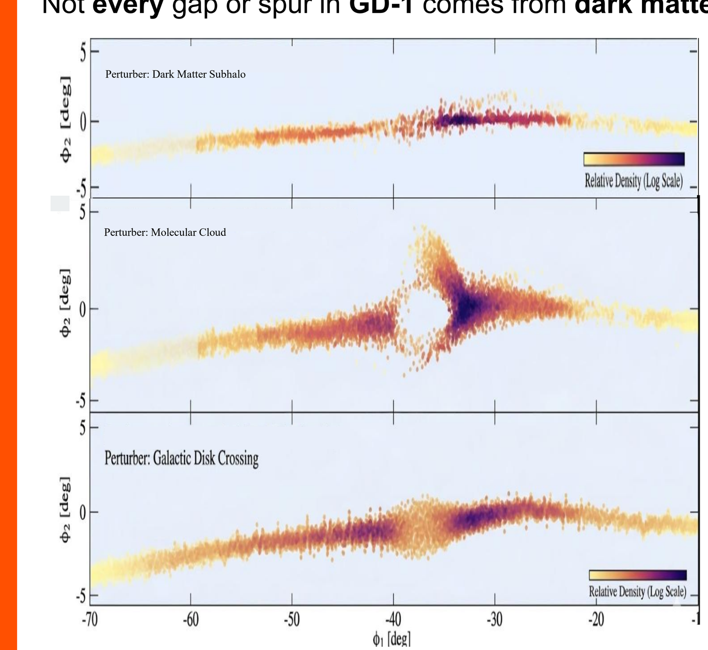
→ Model attributes **highest** posterior probability to **CDM**
→ Finds Bayes factors **>10** against **baryonic explanations**
→ Stream features **successfully** distinguish **dark matter microphysics**

→ Results reach sensitivity to **≈ 3×10⁶ M⊙**
→ **Low-mass** dark substructures **affect** stream
→ Measured **gaps and power-spectrum structure** demonstrate that **GD-1** can probe **subhalo mass spectrums** in Milky Way halo

Stream behaves like **gravitational detector** for **invisible structures**; results suggest that dark matter is **more plausible** explanation than baryonic perturbers for strongest features.

Why Baryonic Confusion still Matters

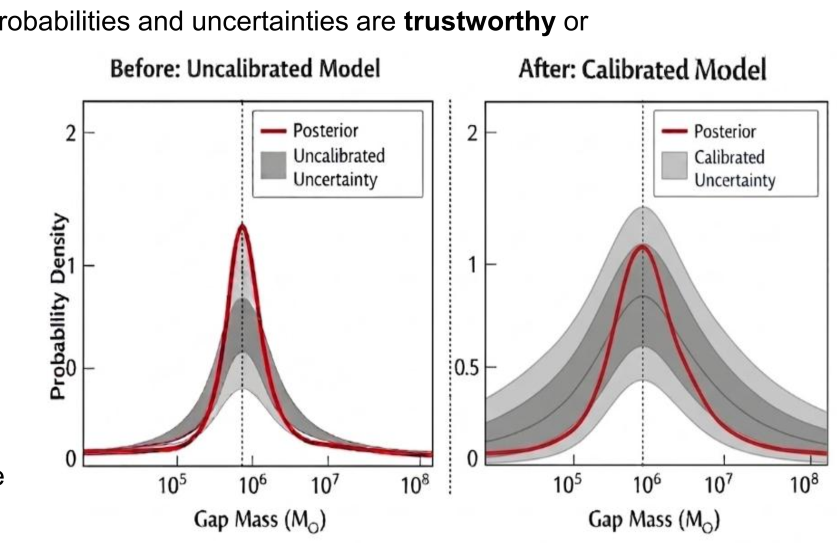
Not **every** gap or spur in **GD-1** comes from **dark matter**



→ Baryonic perturbers (e.g., **molecular clouds, disk crossings, and spiral or bar disturbances**) can create stream features **similar to dark-matter**
→ **Good fit** to the data **does not** automatically prove **dark matter** is responsible
→ **Strongest** dark matter claims are **safest** when **strong** after testing against **baryonic alternatives**
→ In pipeline, **baryonic models included explicitly** so method can separate **real dark matter signals** from **ordinary Milky Way structure**
→ Remaining confusion **strongest** for smaller, shallower, or noisier gaps, where **multiple** causes can look alike
→ Analysis uses **Bayes factors**, **calibration tests**, and **adversarial baryonic examples** to reduce false positives

How Calibration Affects Confidence

Calibration tells us whether the model's probabilities and uncertainties are **trustworthy** or **just look precise**



→ **E.g.**, model: result has **90% confidence**, calibration checks **correctness**
→ In pipeline: calibration improved using **injection–recovery tests**, **PIT histograms**, and **rank histograms**
→ **Uncalibrated** model can be **overconfident** and make strong claims that **are not reliable**.
→ **Good calibration**: posterior masses, Bayes factors, and model probabilities are interpreted as **statistically meaningful**

→ Helps confirm that the strong CDM preference **is not** byproduct of **noisy training** or **poor uncertainty estimates**

Main Scientific Takeaway

- Pipeline shows that GD-1 can be used as gravitational detector for DM substructures
- Strongest stream disturbances best explained by DM subhaloes, not ordinary baryonic perturbers
- Method distinguishes between dark DM model classes using calibrated probabilities and Bayes factors
- Analysis reaches sensitivity down to ≈ 3×10⁶ M⊙, with reliable 1σ mass recovery for M ≥ 5×10⁶ M⊙
- Results suggest stream morphology and kinematics can be turned into quantitative constraints on Milky Way's DM substructure spectrum → makes GD-1 a powerful tool for DM composition search

Improvements on Current Method

1. Compares multiple DM and baryonic model classes directly, adding model-class probabilities and Bayes factors.
2. Method uses full phase-space, multi-scale density information, and topological features.
3. Emulator ensemble makes forward model faster and more realistic, while still keeping uncertainty explicit.
4. Calibration is stronger than in earlier work, with 90% coverage kept between 0.89 and 0.91.
5. tests against baryonic confusers, different Milky Way potentials, and held-out N-body cases.

WHAT'S NEXT?

Stronger Model Exclusions
→ Pipeline can rule out whole classes of dark-matter models
→ Combining BF, posterior probabilities, and calibration tests, method makes stronger, more defensible exclusions
→ Supports ≥ 2σ exclusions of incorrect DM classes when evidence is strong enough
→ Framework reduces false claims by requiring BF>10 and FPR <5% for robust detections.
→ Helps narrow down DM models viability in Milky Way halo

Better Population-Level Constraints
→ Measures overall population of DM substructure in Milky Way halo
→ Estimates perturber-class prevalence
→ Uses: gap frequency, gap depth, and power-spectrum constraints to build broader picture of subhalo mass spectrum
→ Single detections: ambiguous, population-level patterns: stronger evidence
→ Produce more accurate constraints on DM subhaloes at masses around 10^5-10^8 $M_\odot$
→ Helps turn GD-1 from single interest stream to statistical test of DM theory

Extension to Lensing and Satellites
This project is designed so the same inference framework can later include other dark-matter probes besides GD-1.
→ Lensing can test small dark objects through optical distortion, creating independent checks on substructure
→ Satellite galaxies can help constrain larger DM population by providing another view of how structure forms in the Milky Way halo
Combining these probes with GD-1 would let analysis compare different lines of evidence vs. only one mode of data.
→ A signal that appears in multiple probes is harder to explain away
In the long term, this extension could turn pipeline into multi-probe framework for testing DM physics.

THANKS

Pranav Kulkarni (Stanford USA): for mentoring me throughout this process and making time for me in your incredibly busy schedule.
TSF: for all the support in registration, poster design and preparations.
My mom: for supporting me throughout this journey, cheering me on through the ups and down of my project. All of it would not be possible without you!
ÉSTO: À tous mes enseignants et l'administration, merci de m'avoir supporter et encourager à travers cette expérience. Sans vous et notre communauté je n'aurai jamais eu la confiance, la passion ou la détérmination de poursuivre ce projet.

REFERENCES

Alvey, J. E., Gerdes, M., & Weniger, C. (2023). Albatross: A scalable simulation-based inference pipeline for analysing stellar streams in the Milky Way. *Monthly Notices of the Royal Astronomical Society*, *525*(3), 3662–3681. https://doi.org/10.1093/mnras/stad2458
Bonaca, A., Hogg, D. W., Price-Whelan, A. M., & Conroy, C. (2019). The Spur and the Gap in GD-1: Dynamical Evidence for a Dark Substructure in the Milky Way Halo. *The Astrophysical Journal*, *880*(1), 38. https://doi.org/10.3847/1538-4357/ab2873
Buschmann, M., Kopp, J., Safdi, B. R., & Wu, C.-L. (2018). Stellar Wakes from Dark Matter Subhalos. *Physical Review Letters*, *120*(21). https://doi.org/10.1103/physrevlett.120.211101
de Boer, T. J. L., Erkal, D., & Gieles, M. (2020). A closer look at the spur, blob, wiggle, and gaps in GD-1. *Monthly Notices of the Royal Astronomical Society*, *494*(4), 5315–5332. https://doi.org/10.1093/mnras/staa917
Delos, M. S., & Schmidt, F. (2022). Stellar streams and dark substructure: the diffusion regime. *Monthly Notices of the Royal Astronomical Society*, *513*(3), 3682–3708. https://doi.org/10.1093/mnras/stac1022
Erkal, D., & Belokurov, V. (2015). Properties of dark subhaloes from gaps in tidal streams. *Monthly Notices of the Royal Astronomical Society*, *454*(4), 3542–3558. https://doi.org/10.1093/mnras/stv2122
Erkal, D., Belokurov, V., Bovy, J., & Sanders, J. L. (2016). The number and size of subhalo-induced gaps in stellar streams. *Monthly Notices of the Royal Astronomical Society*, *463*(1), 102–119. https://doi.org/10.1093/mnras/stw1957
Price-Whelan, A. M., & Bonaca, A. (2018). Off the Beaten Path: *Gaia* Reveals GD-1 Stars outside of the Main Stream. *The Astrophysical Journal*, *863*(2), L20–L20. https://doi.org/10.3847/2041-8213/aad7b5