AI-Driven Optimization of Detector Design for the FCC-ee

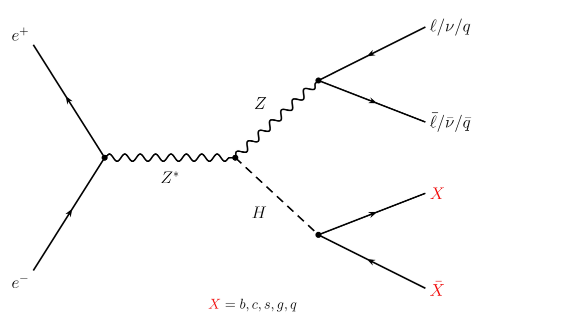
Ivan Brillo[2,3], Dimitris Ntounis[1,2], Umar Sohail Qureshi[1,2], Zhi Zheng[1,2],
Yi-Ren Wu[5], Hanyi Ying[4], Caterina Vernieri[1,2], Loukas Gouskos[4]

[1]Stanford University [2]SLAC National Accelerator Laboratory
[3]Sant'Anna School of Advanced Studies [4]Brown University [5]National Taiwan University

Challenges in FCC-ee detector optimization

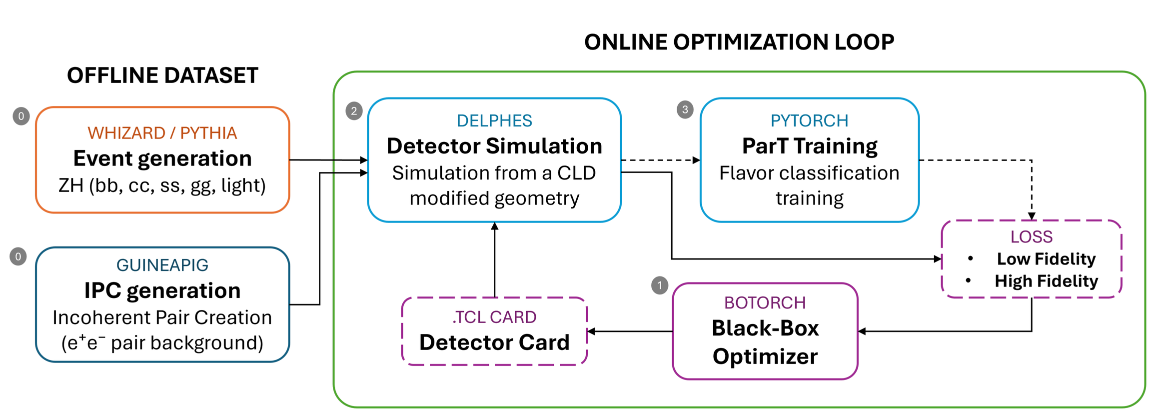
FCC-ee detectors: Enormous & coupled design space + hard constraints → grid search and manual scans infeasible

- **Stringent Physics goals:** FCC-ee targets precise measurements of EW, Higgs and top observables. Achieving the desired precision requires an optimized detector design.
- **High-dimensional, coupled parameter space:** Geometry variables for different subdetectors are large in number and correlated in non-trivial ways.
- **Per candidate geometry:** Fast Sim → Tagger retrain ($b\bar{b}$, $c\bar{c}$, $s\bar{s}$, gg, qq) → Analysis $\sim 1\,$h $\Rightarrow$ brute force infeasible.
- **Complex feasibility constraints:** Detector geometries must satisfy a set of geometric and engineering constraints simultaneously, restricting the feasible space.



Simulation and Optimization Pipeline

Bayesian optimization builds a probabilistic surrogate of the loss surface from past evaluations and proposes the next candidate that maximizes the improvement, achieving better sample efficiency than random or grid search for the same compute budget.



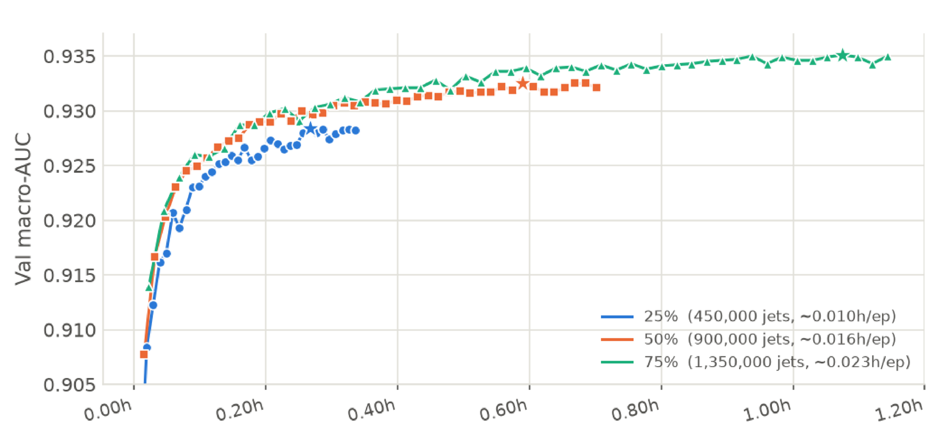
Full optimization loop. MF-BO proposes a new detector card, Delphes re-simulates, and the tagger retrains.

A. Offline dataset generation *(reused across all geometries)*

1. **Signal:** WHIZARD [1] + Pythia6, ZH events ($H \to b\bar{b}$, $c\bar{c}$, $s\bar{s}$, gg, qq) at $\sqrt{s} = 240\,$GeV. 150k events simulated per class.
2. **Beam-induced background:** GuineaPig [2] incoherent e^+e^- pairs.

B. Online optimization loop *(iterated per candidate geometry)*

1. **Propose:** Multi-Fidelity Bayesian Optimization (MF-BO) selects the new feasible geometry.
2. **Simulate:** Delphes [3] fast simulation on all 5 channels to create the PFO dataset.
3. **Train:** Particle Transformer (ParT) [4] retrained from scratch. Each jet is represented as a sequence of up to 64 constituents, each described by 36 features (kinematics, track covariance, impact params, and PID).
4. **Evaluate:** $\overline{\text{macroAUC}} = \frac{1}{3}\sum_{f\in\{b,c,s\}} \text{AUC}_f$ fed back to MF-BO.



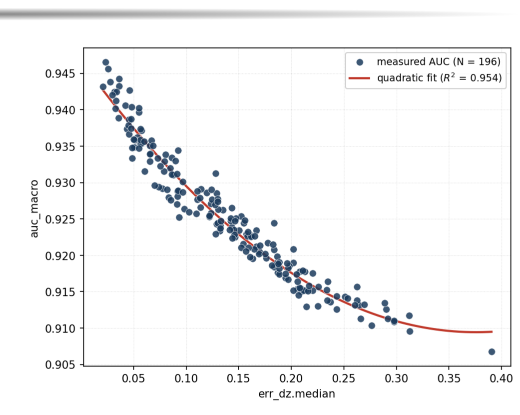
Validation macroAUC vs. cumulative training time for different dataset sizes.

Multi-Fidelity Trust Region Bayesian Optimization

Tagger retraining dominates loop cost, but $\hat{\sigma}_{d_z}$ is strongly anti-correlated with AUC.

We exploit this with **multi-fidelity BO** [5]. Two fidelity levels ($s \in \{0,1\}$):

1. **LF** ($s = 0$, $\sim$1 min): $\hat{\sigma}_{d_z}$ (median longitudinal IP error).
2. **HF** ($s = 1$, $\sim$1 h.): $\overline{\text{macroAUC}}$.



Anti-correlation between $\hat{\sigma}_{d_z}$ and $\overline{\text{AUC}}$.

MF-GP model over $(\mathbf{x}, s) \in \mathbb{R}^{11} \times \{0, 1\}$:

$$k\big((\mathbf{x}, s), (\mathbf{x}', s')\big) = \sigma^2\Big[k_0(\mathbf{x}, \mathbf{x}') + (1-s)(1-s')(1+ss')^p\, k_1(\mathbf{x}, \mathbf{x}')\Big]$$

k_0, k_1: independent ARD Matérn-5/2. k_1 absorbs LF-only variance.

In high dimensions, unsampled regions dominate the acquisition function, causing over-exploration. **TuRBO** [6] addresses this by restricting proposals to a local trust region.

Proposal Loop

1. **Thompson Sampling** inside TuRBO trust region. Draw MF-GP posterior samples at $s = 1$; take $\mathbf{x}^* = \arg\max$ over feasible geometries. TR is centred on the current best HF point, stretched per-dimension by ARD lengthscales of k_0.
2. **Update trust region.** After τ_{succ} consecutive successes ($f_{\text{new}} > f^* + \epsilon$), expand the region; after τ_{fail} consecutive failures, shrink it. If its size falls below $L_{\min}$, restart from a Sobol point.
3. **Select fidelity (MF-MES).** Pick s^* maximising information per unit cost:
$$s^* = \arg\max_{s\in\{0,1\}} \frac{I(y(\mathbf{x}^*, s);\, f^*)}{c(s)},$$
where $f^* = \max_{\mathbf{x}} f^{(1)}(\mathbf{x})$.

Detector Design Space

This study focuses on a **silicon tracker** (barrel layers + endcap disks of the vertex); future work will extend the comparison to TPC and drift-chamber concepts.
The VTX geometry is an 11-dimensional vector $\mathbf{x}$ using an equal-spacing model:

$$r_i = r_1 + (i-1)\,s_r, \quad i = 1,2,3 \quad \text{(barrel radii)},$$
$$z_j = z_1 + (j-1)\,s_z, \quad j = 1,2,3 \quad \text{(disk } |z| \text{ positions)}.$$

#	Symbol	Description	Domain
1	ρ_b	Barrel resolution	$[3,\ 30]\ \mu$m
2	ρ_d	Disk resolution	$[3,\ 30]\ \mu$m
3	w_b	Barrel thickness	$[0.05,\ 5]$ mm
4	w_d	Disk thickness	$[0.05,\ 5]$ mm
5	Z_b	Barrel half-length	$(0,\ Z_{K_1}]$
6	z_1	$\lvert z\rvert$ of first disk	$(0,\ Z_{K_1}]$
7	s_z	Disk z-spacing	$[0,\ Z_{K_1}/2]$
8	r_1	Radius of first barrel	$[R_{\text{pipe}},\ R_{\text{ITK}}]$
9	s_r	Barrel radial spacing	$[0,\ (R_{\text{ITK}}{-}R_{\text{pipe}})/2]$
10	$r^d_{\min}$	Disk inner radius	$[R_{\text{pipe}},\ R_{\text{ITK}}]$
11	$r^d_{\max}$	Disk outer radius	$[R_{\text{pipe}},\ R_{\text{ITK}}]$

Fixed constants: $R_{\text{pipe}} = 12.35\,$mm, $R_{\text{ITK}} = 127\,$mm, $Z_{K_1} = 514\,$mm.

Detector Design Constraints

These constraints restrict Thompson sampling to feasible detector designs. The first four constraint classes are affine, whereas the final hit-acceptance constraint is non-affine but differentiable.

VTX-ITK confinement:

$$r_1 + 2s_r + \tfrac{w_b}{2} \le R_{\text{ITK}}, \qquad r^d_{\max} \le R_{\text{ITK}},$$
$$z_1 + 2s_z + \tfrac{w_d}{2} \le Z_{K_1}.$$

Beam-pipe clearance:

$$r_1 \ge R_{\text{pipe}}, \qquad r^d_{\min} \ge R_{\text{pipe}}.$$

Non-overlap and ordering:

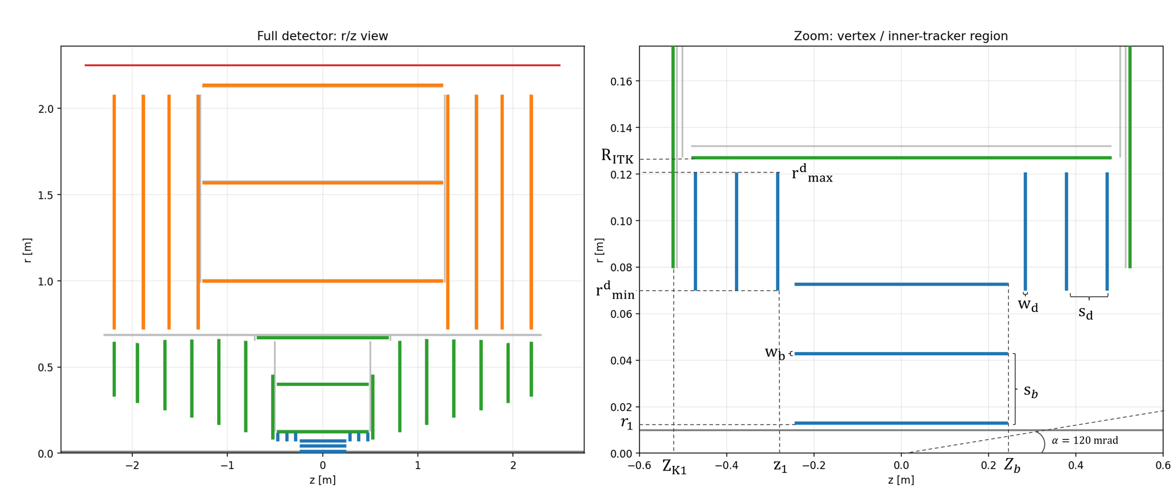
$$s_r \ge w_b, \qquad s_z \ge w_d,$$
$$r^d_{\min} \le r^d_{\max}, \qquad Z_b \le z_1 - \tfrac{w_d}{2}.$$

Forward acceptance ($\kappa = \tan 120\,$mrad):

$$r_1 \ge \kappa\, Z_b, \qquad r^d_{\min} \ge \kappa\,(z_1 + 2s_z).$$

Hit acceptance: every track in $\theta \in [9°, 90°]$ must cross ≥ 6 active layers.

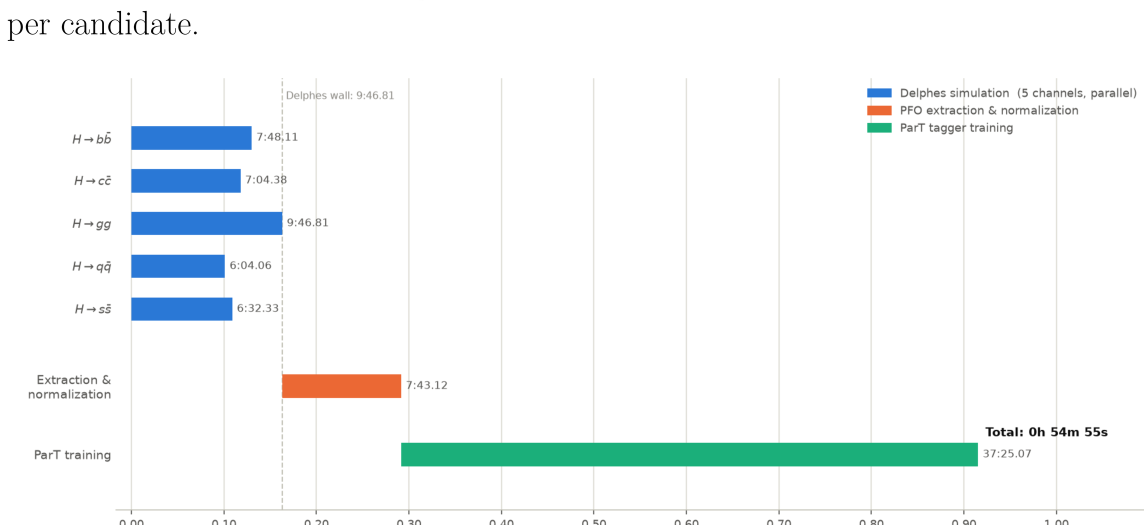
$$N_{\text{hits}}(p_T, \theta) \ge N_{\min} = 6 \quad \forall\,(p_T, \theta) \in \mathcal{A}.$$



Feasible VTX geometry in the R-z plane (zoomed into the vertex/inner-tracker region).

Loop Timing Analysis

One complete optimization iteration, including five-channel Delphes simulation, PFO extraction, and ParT retraining, takes ~ 1 **hour** of wall-clock time for $5 \times 150\,$k events per candidate.

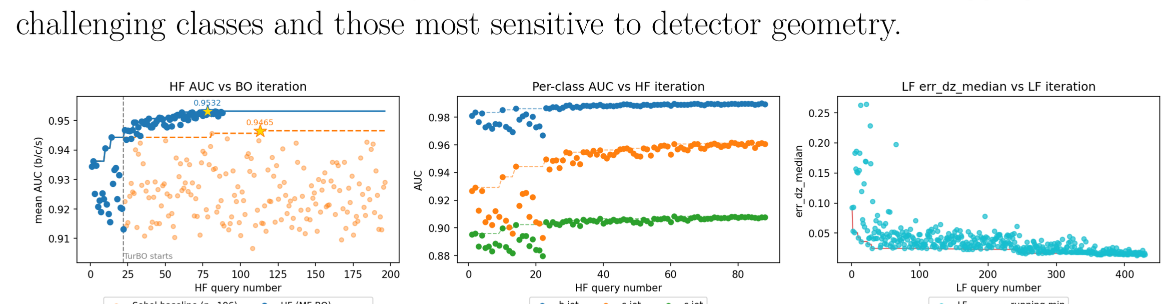


Wall-clock breakdown for one BO iteration using $5 \times 150\,$k events.

- ParT training is the **dominant bottleneck**, accounting for $\sim$68% of the total wall-clock time.
- The highest-leverage strategies for speeding up the loop are reducing the number of HF evaluations, the dataset size, or the number of training epochs.

Results

TuRBO + MF-MES vs. Sobol baseline. The MF optimizer is seeded from $2d = 22$ HF and LF Sobol warm-up points.
The improvement achieved by MF-BO is concentrated in c- and s-tagging, the most challenging classes and those most sensitive to detector geometry.

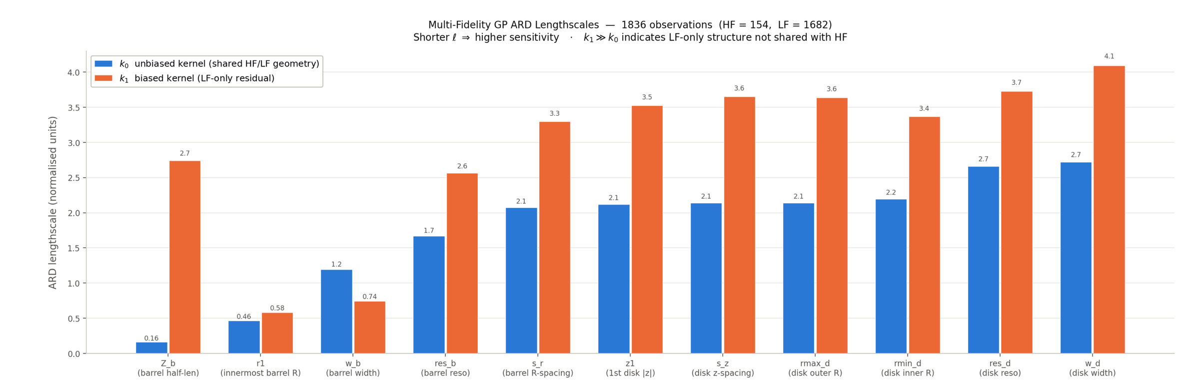


Left: mean AUC versus query number, with the running best shown as a solid line. *Center*: running best AUC for each flavor class. *Right*: running best low-fidelity (LF) objective.

Results (cont.)

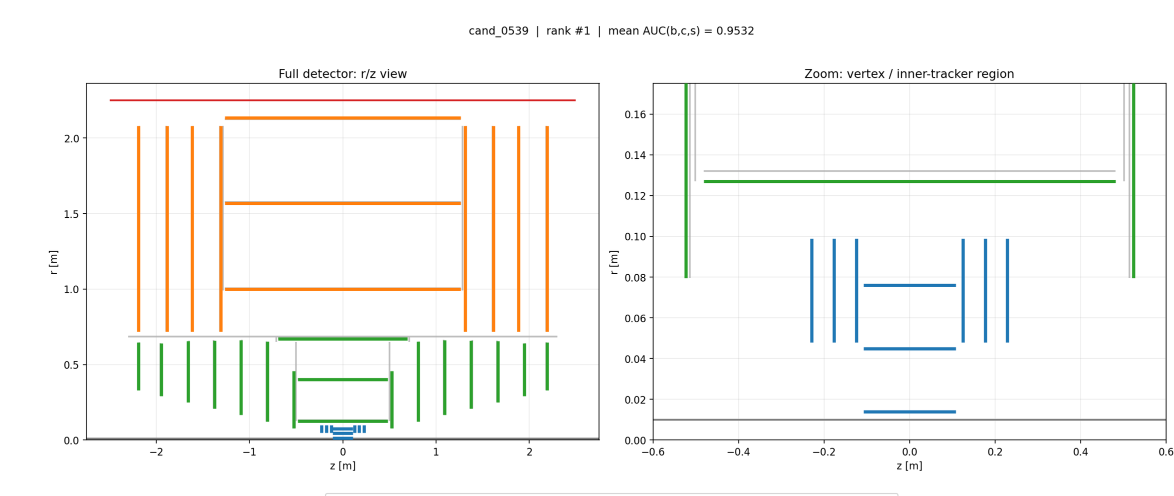
MF-GP ARD lengthscales: after 154 HF and 1,682 LF observations, the ARD lengthscales of the MF-GP kernels k_0 and k_1 reveal the sensitivity of each parameter.

- The barrel parameters have the shortest lengthscales and are therefore the most sensitive, whereas the disk parameters have longer lengthscales.
- Large k_1 lengthscales indicate that the LF-specific component varies slowly across the design space, consistent with broad agreement between the LF approximation and the HF landscape.

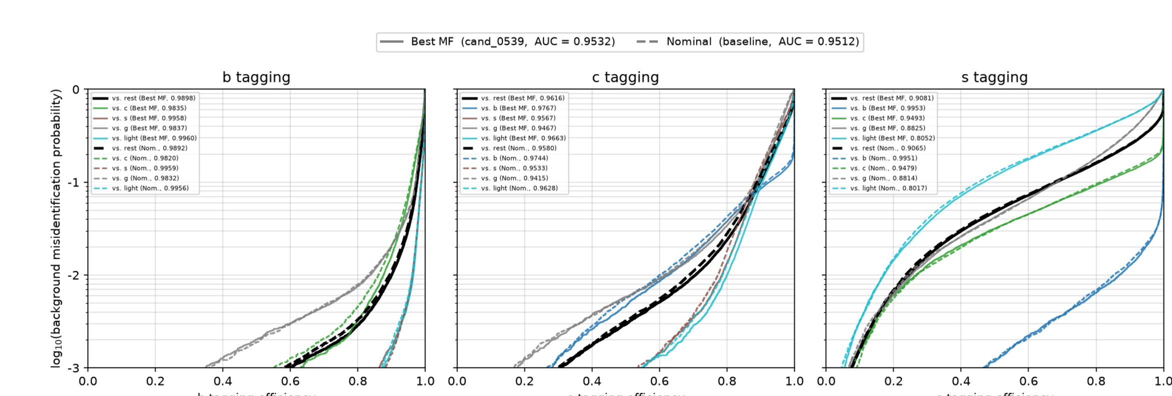


MF-GP ARD lengthscales (normalized units). Bars are sorted short-to-long, left-to-right.

The optimizer converged to a geometry with $\overline{\text{AUC}} = 0.9532$. Full and zoomed r-z views, together with the per-class AUC plots, are shown below.



r-z view of the best candidate detector.



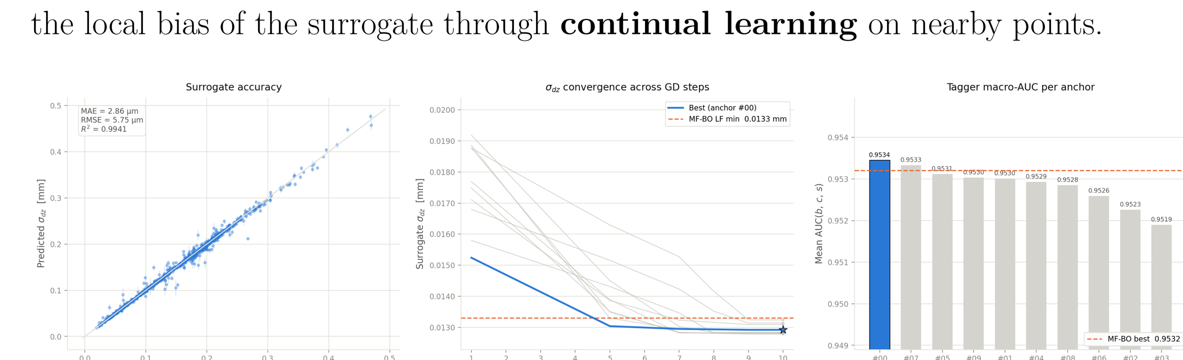
Per-class AUC plots for the best detector candidate, compared with the nominal Delphes baseline.

Takeaways

- Exploiting the low-fidelity objective substantially accelerates the search for the optimum.
- TuRBO is well suited to black-box optimization in high-dimensional spaces.
- ARD analysis of the MF-GP identifies barrel geometry as the dominant performance driver.

Future Work

To complement the MF-BO global search, we trained a **differentiable neural network surrogate** mapping the 11 geometry parameters to σ_{d_z}, achieving a test $R^2 = 0.994$.
We exploited the gradients of the neural network to perform local optimization within trust regions using the SLSQP optimizer.
Starting from $K = 10$ Sobol geometries, we optimized σ_{d_z} for $N = 10$ steps, correcting the local bias of the surrogate through **continual learning** on nearby points.



GD surrogate study. (Left) Surrogate accuracy on held-out test set. (Centre) σ_{d_z} convergence over $K = 10$ anchors. (Right) True tagger macro-AUC per final geometry.

Other extensions under investigation:

- **Coupling precision:** translate macro-AUC gains into Higgs coupling precision (g_{Hbb}, g_{Hcc}, g_{Hss}) via a simplified ZH coupling analysis.
- **Increase search space:** exploit the additivity on the optimization of different subdetectors together.
- **IPC occupancy:** study the optimal geometry under beam-induced background occupancy constraints at the innermost layer.
- **Physics-motivated material budget:** couple the resolution parameters to material budget connecting resolution to cooling requirements and data-transfer load.
- **Full simulation:** add a final full-simulation stage for the top-performing candidates, including seed retraining to validate their improvement over baseline.

References

[1] Wolfgang Kilian, Thorsten Ohl, and Jurgen Reuter. WHIZARD: Simulating Multi-Particle Processes at LHC and ILC. *Eur. Phys. J. C*, 71:1742, 2011.

[2] C. Rimbault, P. Bambade, O. Dadoun, G. Le Meur, F. Touze, M. C. del Alabau, and D. Schulte. GUINEA PIG++ : An Upgraded Version of the Linear Collider Beam Beam Interaction Simulation Code GUINEA PIG. *Conf. Proc. C*, 070625:2728, 2007.

[3] J. de Favereau, C. Delaere, P. Demin, A. Giammanco, V. Lemaître, A. Mertens, and M. Selvaggi. DELPHES 3, A modular framework for fast simulation of a generic collider experiment. *JHEP*, 02:057, 2014.

[4] Huilin Qu, Congqiao Li, and Sitian Qian. Particle Transformer for Jet Tagging, January 2024.

[5] S. Takeno, H. Fukuoka, Y. Tsukada, T. Koyama, M. Shiga, I. Takeuchi, and M. Karasuyama. Multi-fidelity Bayesian Optimization with Max-value Entropy Search and its Parallelization. In *PMLR*, 2020.

[6] D. Eriksson, M. Pearce, J. Gardner, R. D Turner, and M. Poloczek. Scalable Global Optimization via Local Bayesian Optimization. In *NeurIPS*, volume 32, 2019.