Learning to Design Physics Instruments for Varying Budget with Reinforcement Learning

Sara Zoccheddu, Shah Rukh Qasim, Patrick Owen, Nicola Serra

Universität Zürich[UZH]

Motivation

- The **design of advanced physics instruments** is a complex and resource-intensive task—one that requires optimizing numerous parameters to achieve high performance while meeting stringent cost, material, and spatial constraints.
- **Traditional design methods** typically optimize for a fixed set of constraints, requiring repeated manual re-tuning for each new scenario.
- A **flexible approach** is needed—one that delivers optimal detector designs across a range of possible budgets, enabling decision-makers to explore the full landscape of performance versus resource use.

Proposed solution

Demonstrated through calorimeter design

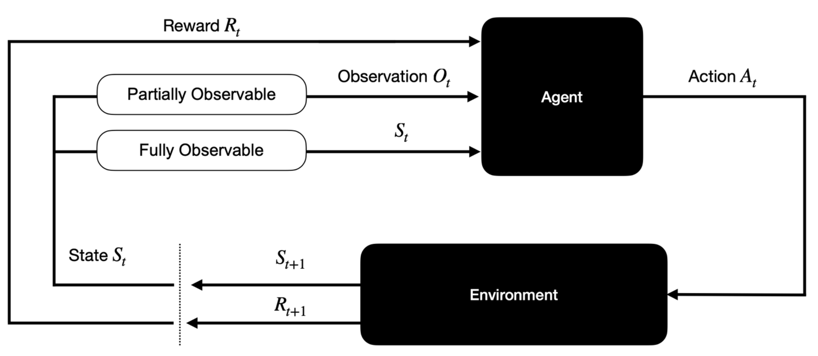
Budget-conditioned RL: the RL agent receives the budget as part of its observation.

This approach produces a **family of optimized calorimeter designs** in a single training run.

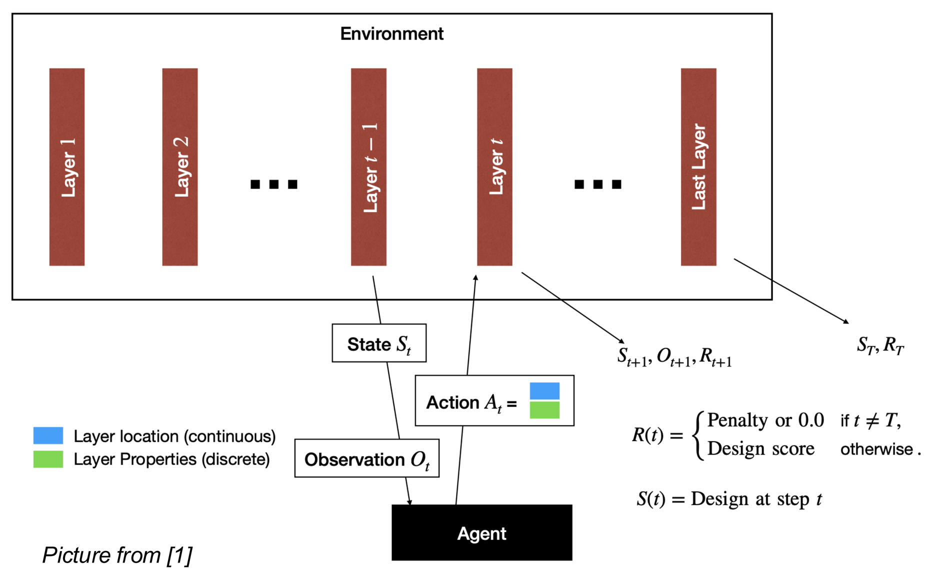
It shows the **trade-off** between performance and resource use.

Methodology

- Agent–environment interaction loop in reinforcement learning.
- Proximal Policy Optimization (PPO).



Picture from [1]



Picture from [1]

- At each episode, a scalar budget is drawn uniformly from a predefined interval.

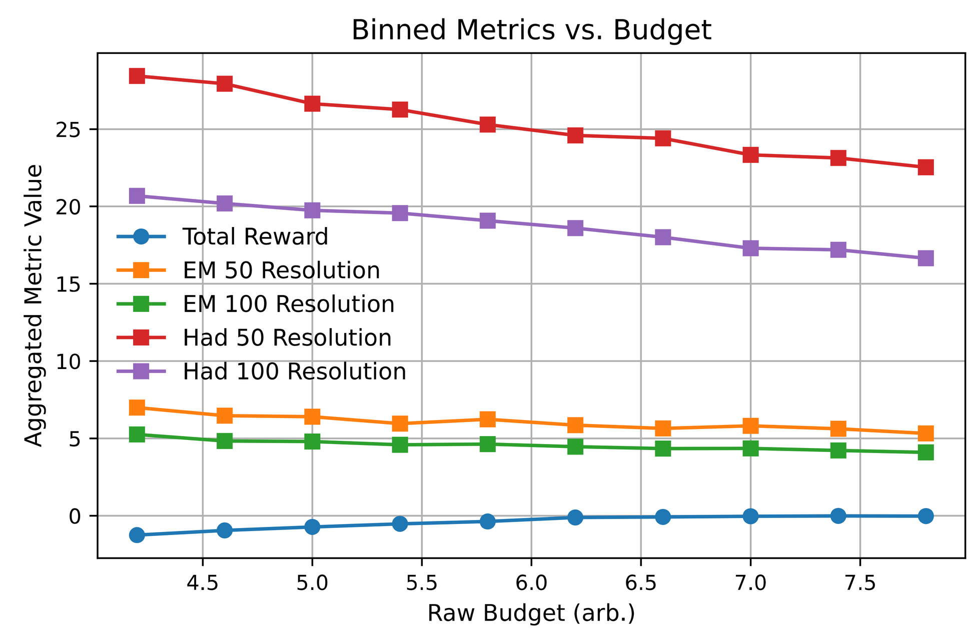
- The sampled budget is then normalized to [0,1] and appended to the observation.

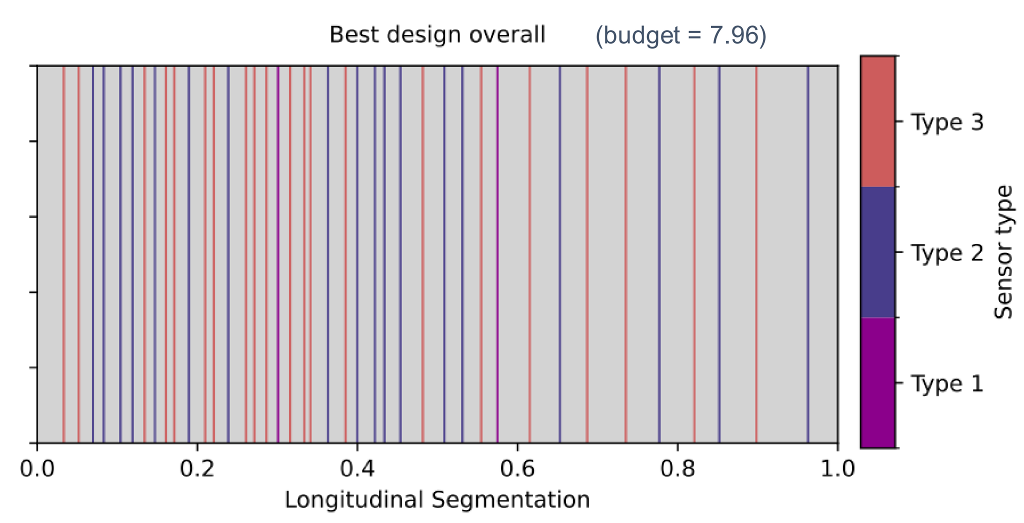
$$O_t = [\,longitudinal\ dimension,\ thickness\ used,\ budget\ sampled\,].$$

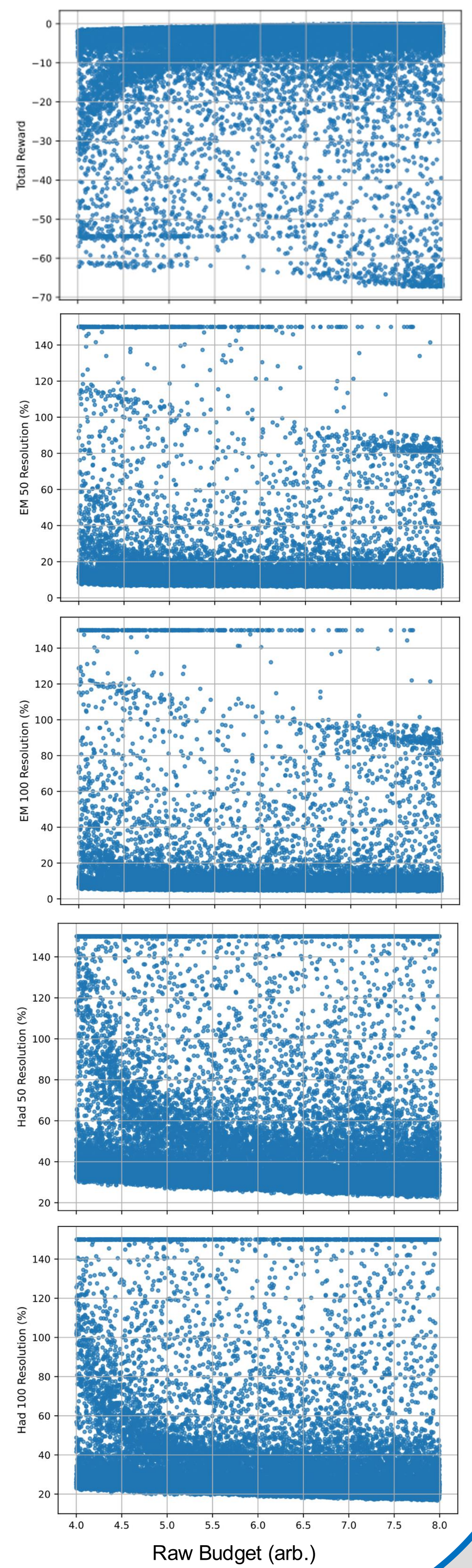
- **Action space:** at each step the agent outputs a mixed (continuous + discrete) action $A_t \in A$:
  - **continuous:** the gap (Δ z) to the next active layer;
  - **discrete:** the sensor type (out of 3 available thicknesses).
- **Agent's policy:** $\pi(\text{action} \mid \text{state, budget\_sampled})$
- **Reward:** it penalizes resolution above target thresholds.

Results

- The **total reward increases** with budget, reflecting improved overall performance when more resources are available.
- Both **electromagnetic** resolutions and **hadronic** resolutions consistently **decrease** (improve) as the budget increases.







Main conclusion

- By directly **conditioning calorimeter design** on the available budget, our reinforcement learning framework enables the generation of optimized calorimeter configurations across the entire range of resource constraints.
- This budget-aware strategy empowers experiment designers to efficiently navigate cost–performance **trade-offs** and make **informed decisions** under evolving budget scenarios.

Future steps:

- Detector design can be flexibly optimized with respect to multiple objectives—such as cost constraints and energy resolution—simultaneously.
- By allowing the RL agent to consider several performance and resource criteria at once, this approach produces a comprehensive set of solutions that reflect real-world trade-offs.
- This multi-objective capability is especially valuable in experimental physics, where project requirements often shift and decision-makers need to balance scientific goals against practical limitations.

References

[1] Qasim, S. R., Owen, P., & Serra, N. (2024). *Physics Instrument Design with Reinforcement Learning*. arXiv. https://arxiv.org/abs/2412.10237