Sustainable HEP 2026 — 5th Edition

Reducing simulation-related emissions in rare-event searches through optimised event biasing

P. Knights[1], **L. J. Milligan[1]**, K. Nikolopoulos[1,2]

[1]University of Birmingham, School of Physics and Astronomy, B15 2TT, UK
[2]University of Hamburg, Institute for Experimental Physics, Hamburg, 22761, Germany

UNIVERSITY OF BIRMINGHAM

Birmingham Gaseous Detectors Laboratory

DM Direct Detection Landscape

UNIVERSITY OF BIRMINGHAM

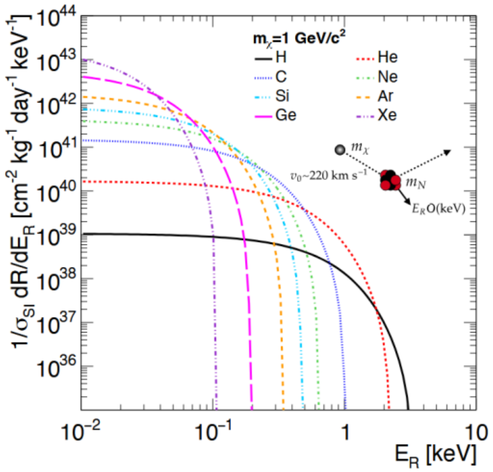
DM direct detection now reaching two extreme frontiers:

- High exposure, **extremely low background**
- Low recoil energies, **very low thresholds**

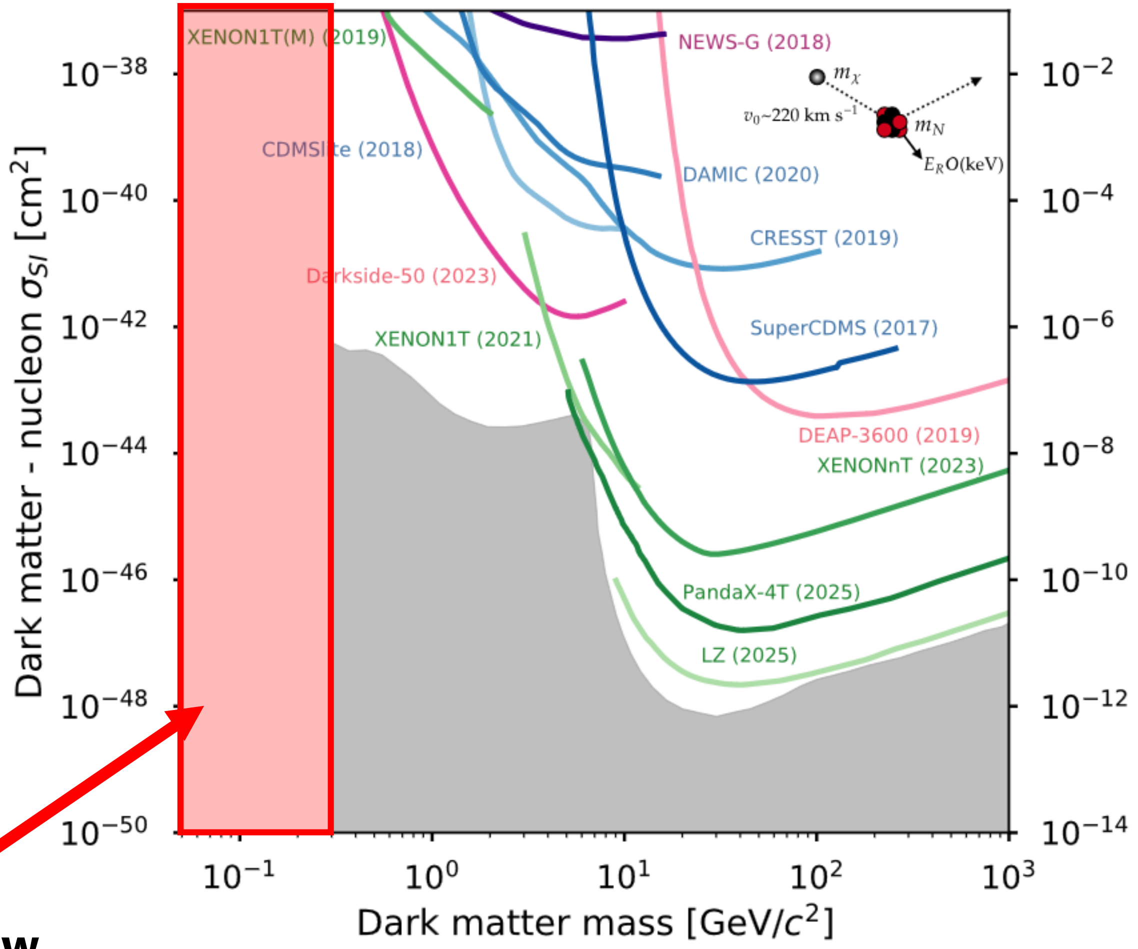
Background simulations **computationally expensive** without event biasing

- Particularly for **shielding studies – ~ 1 in 1 million events survive**

How can we best **optimise our biasing** to reduce computational load, thus **manage our carbon footprint**



Unexplored low mass region – light dark matter



Higher Exposure

Lower Threshold/Lighter Nuclei

Simulation Biasing Schemes

UNIVERSITY OF BIRMINGHAM

Biasing or “variance reduction” techniques aim to reduce the relative uncertainty (RU) of a simulation, for a given computation time

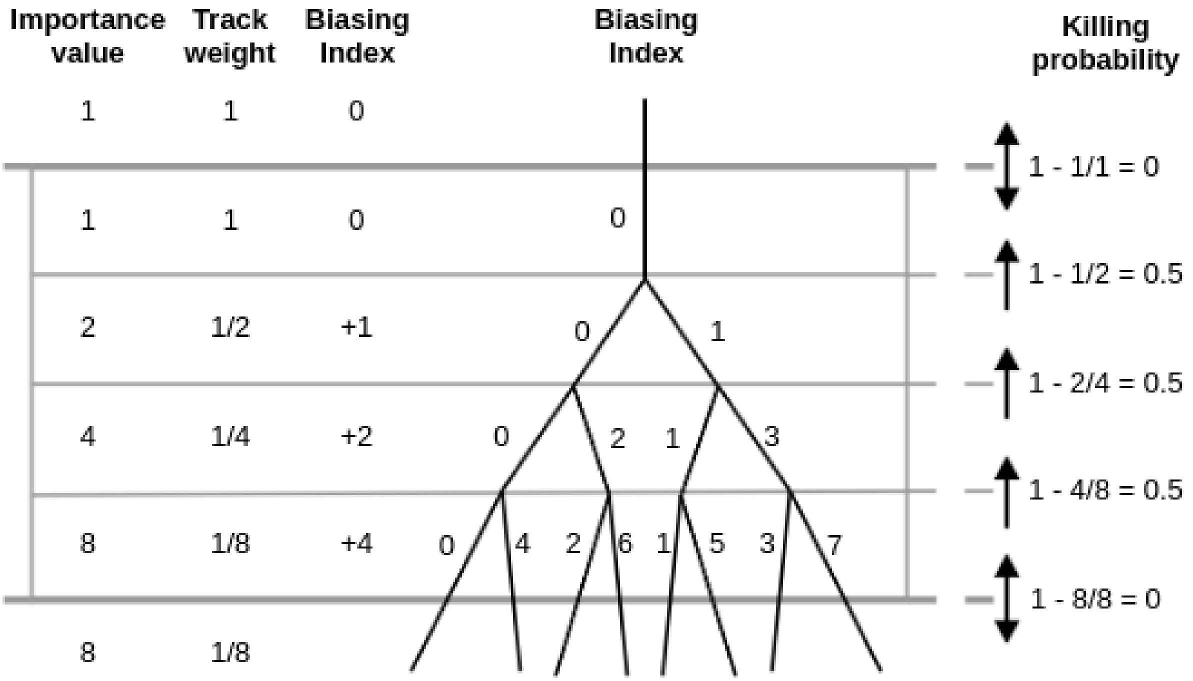
- Modify natural sampling procedure to focus on “important” events
- To compensate for this modification, particles are given statistical weights

$$w_{\text{biased}}\mathbf{pdf}_{\text{biased}} = w_{\text{unbiased}}\mathbf{pdf}_{\text{unbiased}} \qquad n \times w_{\text{biased}} = w_{\text{unbiased}}$$

Range of techniques w/ GEANT4 implementations:

- **Importance splitting –** Geant4 example *biasing/B02*
- Forced collisions – Geant4 example *biasing/GB02*
- Cross-section enhancement – Geant4 example *biasing/GB01*
- Reverse Monte-Carlo – Geant4 example *biasing/reverse/MC01*
- Many more!

Importance Splitting/Killing Biasing



Zatschler et al. NIMA, doi.org/10.1016/j.nima.2025.170766

Questions from user:

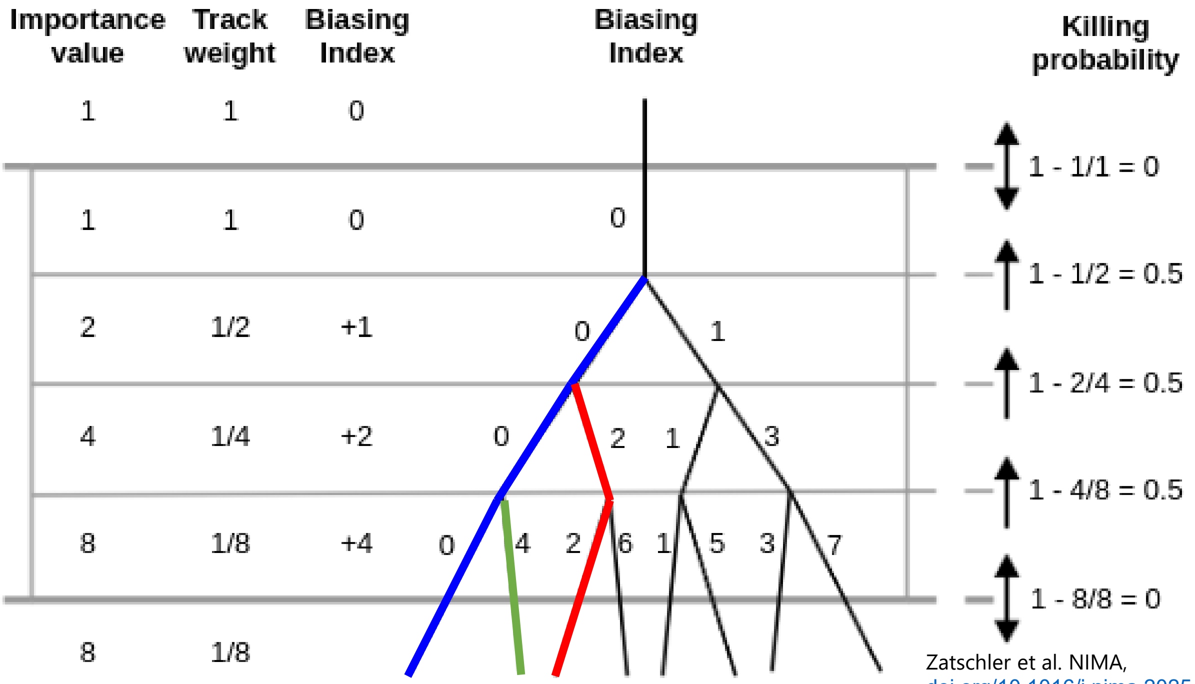
- How to keep account of different track/splitting topologies?
- **How do we optimise layer count? General rule?**

Principles:

- Split volume into **layers assigned importances $\boldsymbol{I}$**
- Particles crossing boundaries forward get **split according to $\boldsymbol{N_{split} = I_{post}/I_{pre}}$**
- Particles going backward get killed with prob. $\boldsymbol{p_{kill} = 1 - I_{post}/I_{pre}}$
- Particles given **statistical weights $\boldsymbol{w = 1/I}$** from final layer
- For m layers, if $\boldsymbol{I = 2^m}$, $\boldsymbol{N_{split} = 2}$, $\boldsymbol{p_{kill} = 0.5}$, $\boldsymbol{w = 1/2^n}$ – most common config.

Building on the Biasing Scheme

UNIVERSITY OF BIRMINGHAM



Zatschler et al. NIMA, doi.org/10.1016/j.nima.2025.170766

Questions from user:

- ~~How to keep account of different track/splitting topologies?~~
- **How do we optimise layer count? General rule?**

Introduce a Biasing Index (B), according to implementation in Zatschler et al., NIMA 1080 (2025) 170766

$$\mathcal{B}_{\text{copy}} = \mathcal{B}_{\text{original}} + I/2$$

- Unique topologies - i.e. those not created by physics processes, can be tracked independently
- Consider collection of hits from **different indices** as **distinct** detector hits
- Sum the **shared weight** from each **distinct** collection/hit to reconstruct background rate

Secondaries not belonging to biased particle type inherit parent particles biasing index

Implementation for DarkSPHERE-30

UNIVERSITY OF BIRMINGHAM

Utilise geometry in DarkSPHERE-30 background simulations as testbed

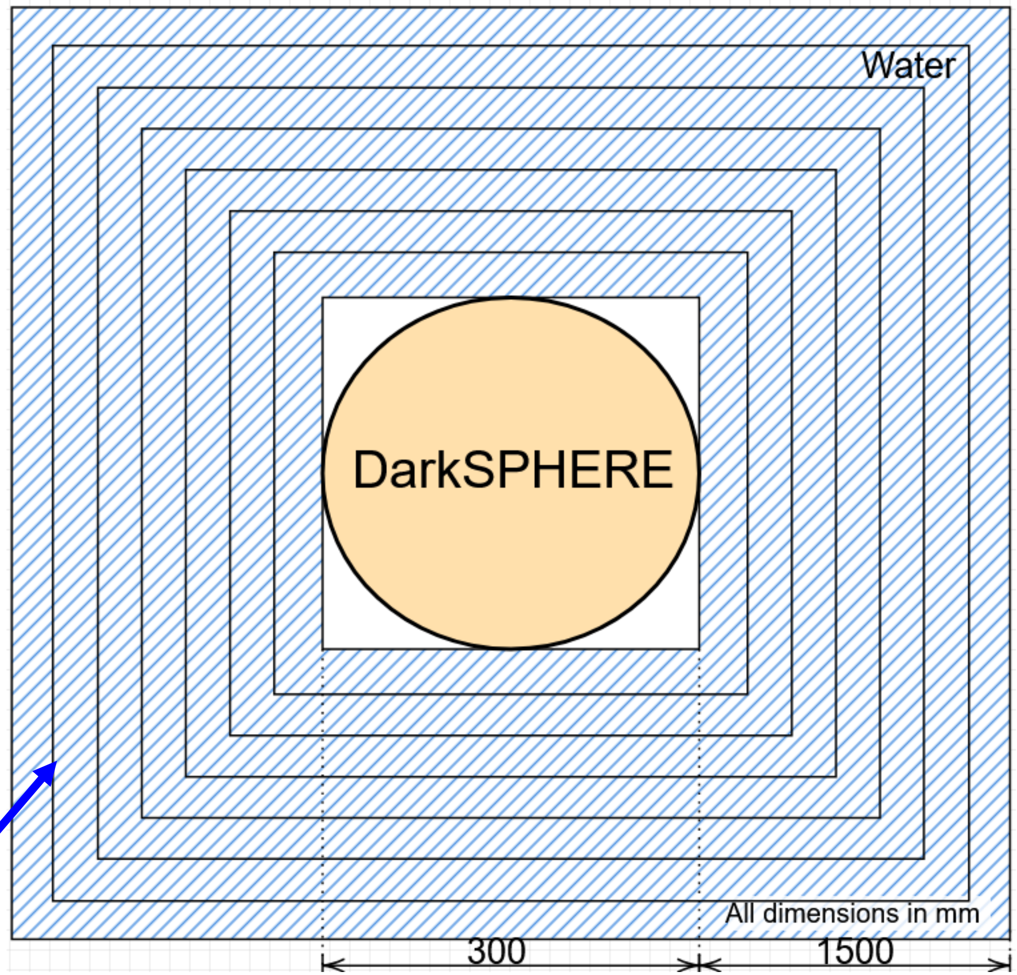
- 14 cm rad. He:C_4H_{10} (10%:90%) at 2 bar, encased in copper (1 cm thick)
- Surrounded in 1.5m thick water shielding
- Cubic air-gap between shielding and sphere

Biasing layers defined such that:

$$t_{\text{shield}} = M \times t_{\text{layer}}.$$

Importances defined such that:

$$I = 2^m,\ N_{\text{split}} = 2 \text{ and } p_{\text{kill}} = 0.5.$$



Equally thick layers fill shielding volume
NOT part of physical volumes

Understanding Optimal Performance

UNIVERSITY OF BIRMINGHAM

Parameter of interest is probability of interaction with detector:

$$Pr_{\text{interaction}} = \frac{\sum w}{N_{\text{primary}}}.$$

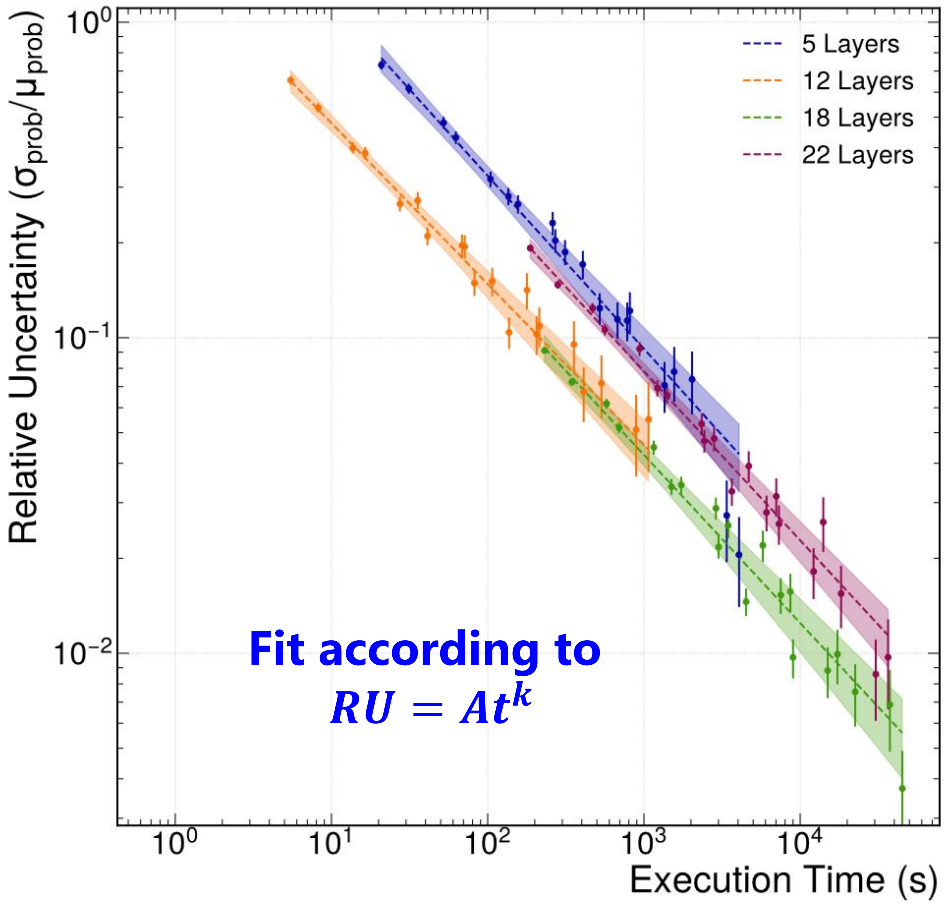
Probability is calculated for differently sized ensembles of simulation results

- Means and SDs used to define relative uncertainty (RU) as basis of comparison between layer amounts:

$$\text{R.U.} = \frac{\sigma_{\text{Pr}}}{\mu_{\text{Pr}}}.$$

RU vs. Execution time curves produced and fit to determine trends for computational efficiency ➜ reduced CO2 footprint

Expectation: at fixed times, RU decreases until reaching a high enough number of layers, then increases



Optimal Performance: Photons

Expectation: at fixed times, RU decreases until reaching a high enough number of layers, then increases

Clearly observe trend by comparing RUs at fixed execution times:
- 300 s for 2750 keV & 100 s for 1461 keV

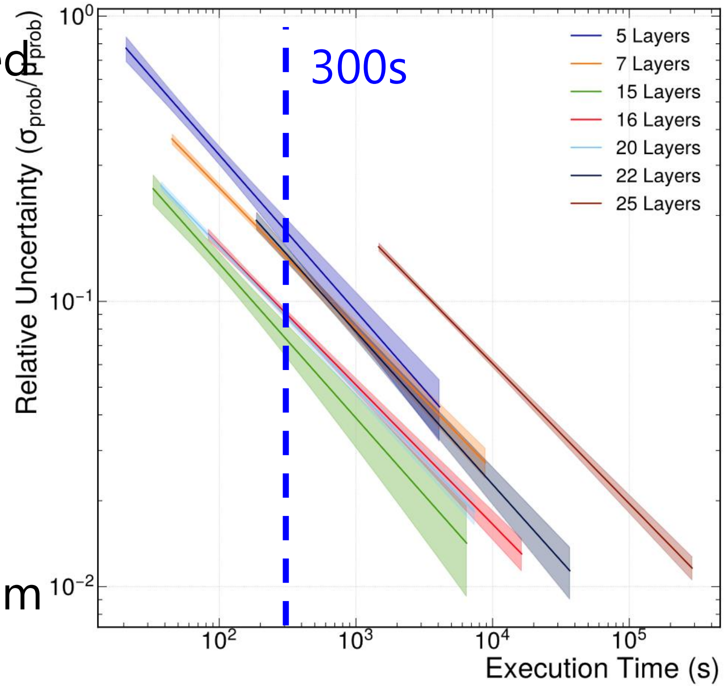
Parabolic trends observed - **turning points** represent **optimal amount of layers**

Link to physics for predictability
- Define $t_{layer} = x \times \lambda^{\gamma}_{H_2O}$
- ($\lambda^{\gamma}_{H_2O}$=24.05 cm 2750 keV, $\lambda^{\gamma}_{H_2O}$=17.14cm 1461 keV)

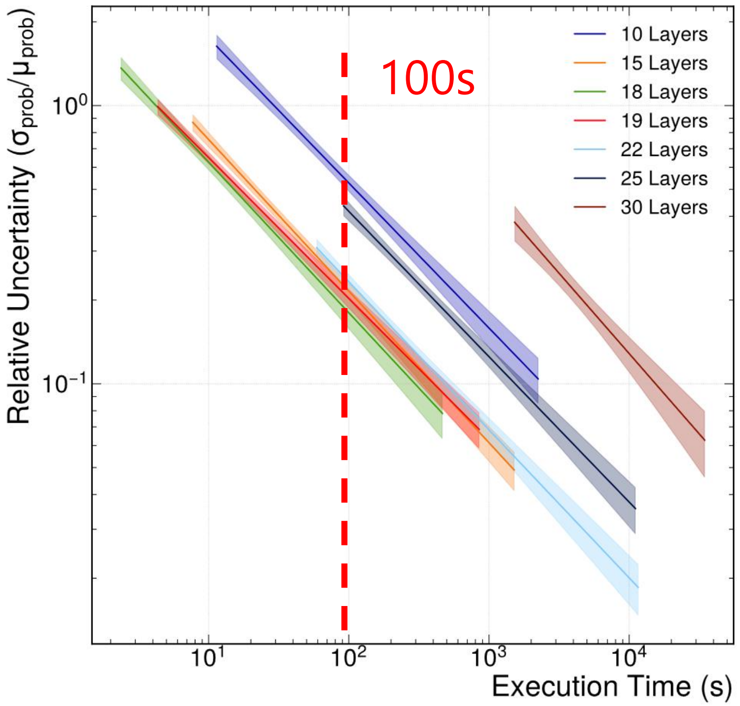
14 layers @ 2750 keV: $(0.45 \pm 0.03)\lambda^{\gamma}_{H_2O}$

18 layers @ 1461 keV: $(0.48 \pm 0.02)\lambda^{\gamma}_{H_2O}$

2750 keV



1461 keV



Optimal Performance: Photons

UNIVERSITY OF BIRMINGHAM

Expectation: at fixed times, RU decreases until reaching a high enough number of layers, then increases

Clearly observe trend by comparing RUs at fixed execution times:

- 300 s for 2750 keV & 100 s for 1461 keV

Parabolic trends observed - **turning points** represent **optimal amount of layers**

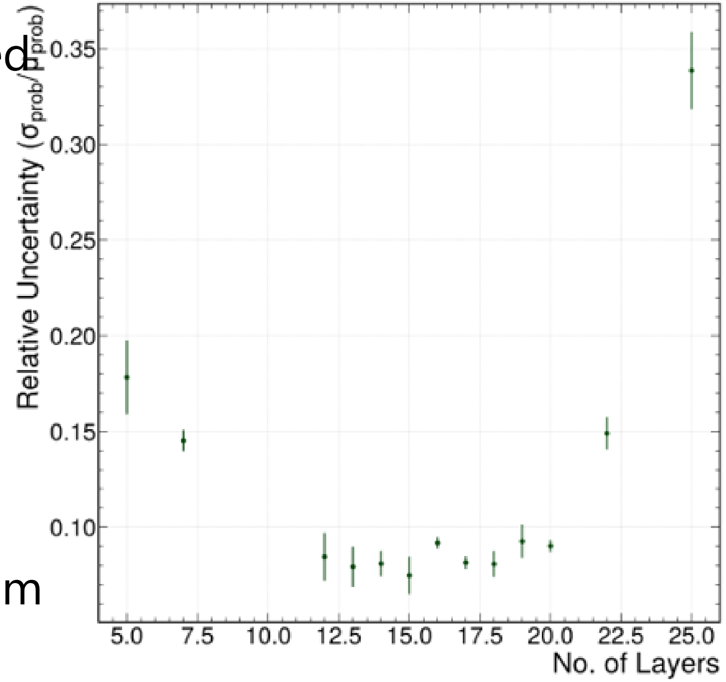
Link to physics for predictability

- Define $t_{layer} = x \times \lambda^{\gamma}_{H_2O}$
- ($\lambda^{\gamma}_{H_2O}$=24.05 cm 2750 keV, $\lambda^{\gamma}_{H_2O}$=17.14cm 1461 keV)

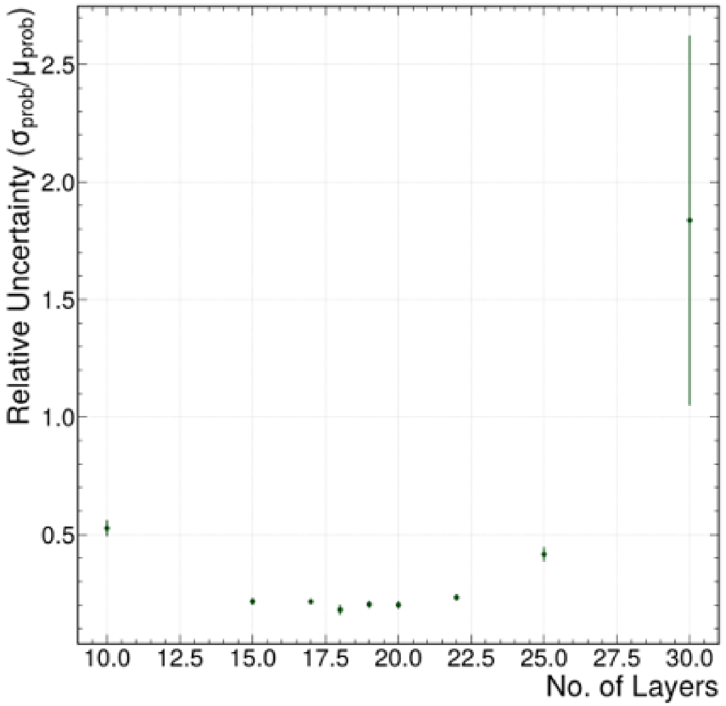
14 layers @ 2750 keV: $(0.45 \pm 0.03)\lambda^{\gamma}_{H_2O}$

18 layers @ 1461 keV: $(0.48 \pm 0.02)\lambda^{\gamma}_{H_2O}$

2750 keV



1461 keV



Note: λ **here refers to mean free path**, not rad. length (X_0) which would be ~36 cm at 2750 keV, e.g.

Optimal Performance: Neutrons

UNIVERSITY OF BIRMINGHAM

Can we extend this guidance to other particle types?

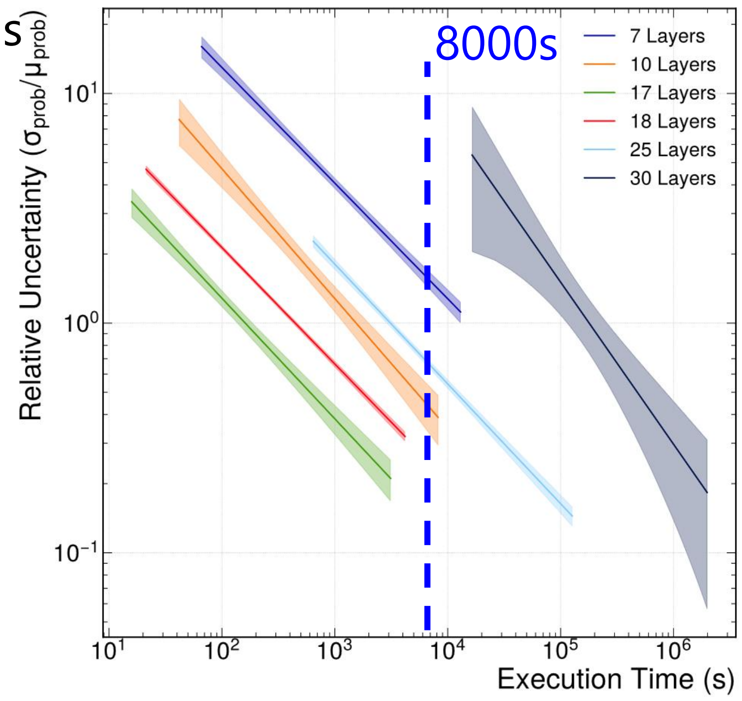
Process used for photons repeated for neutrons

Fits match expectation, and parabolic trend seen at fixed times:
- 8000 s for 10 MeV & 300 s for 5 MeV

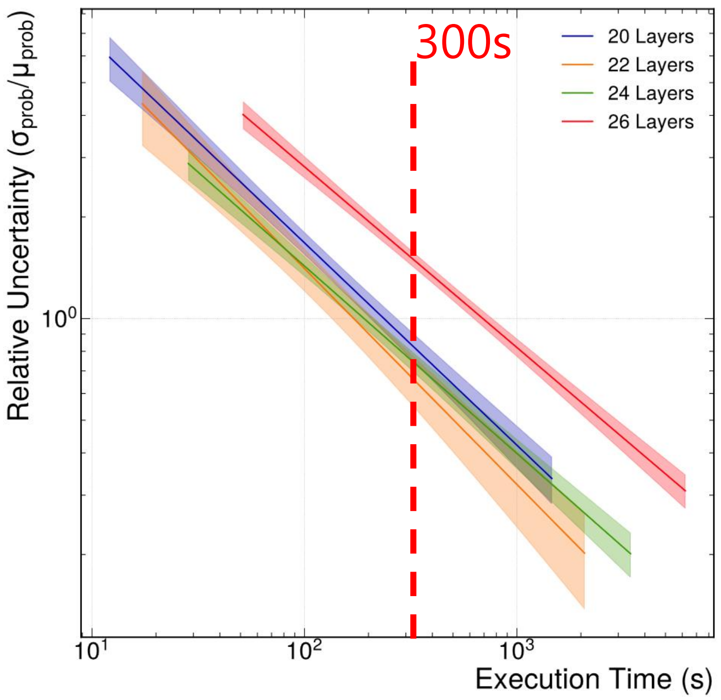
Turning/optimal points ($\lambda^n_{H_2O}$=9.62 cm @ 10 MeV, $\lambda^n_{H_2O}$=7.14 cm @ 5 MeV):
- **18 layers @ 10 MeV: $(0.92 \pm 0.05)\lambda^n_{H_2O}$**
- **23 layers @ 5 MeV: $(0.91 \pm 0.03)\lambda^n_{H_2O}$**

Same behaviour seen for neutrons, for different values

10 MeV



5 MeV



Optimal Performance: Neutrons

UNIVERSITY OF BIRMINGHAM

Can we extend this guidance to other particle types?

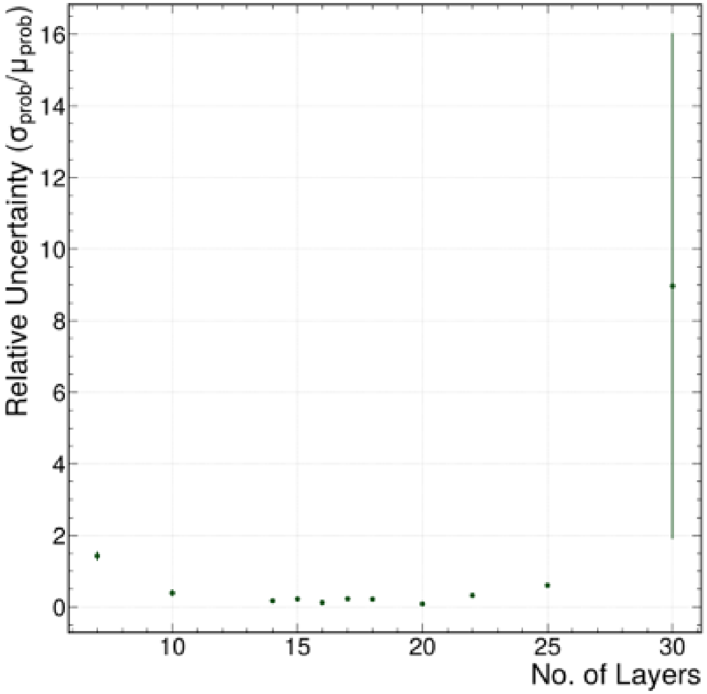
Process used for photons repeated for neutrons

Fits match expectation, and parabolic trend seen at fixed times:
- 8000 s for 10 MeV & 300 s for 5 MeV

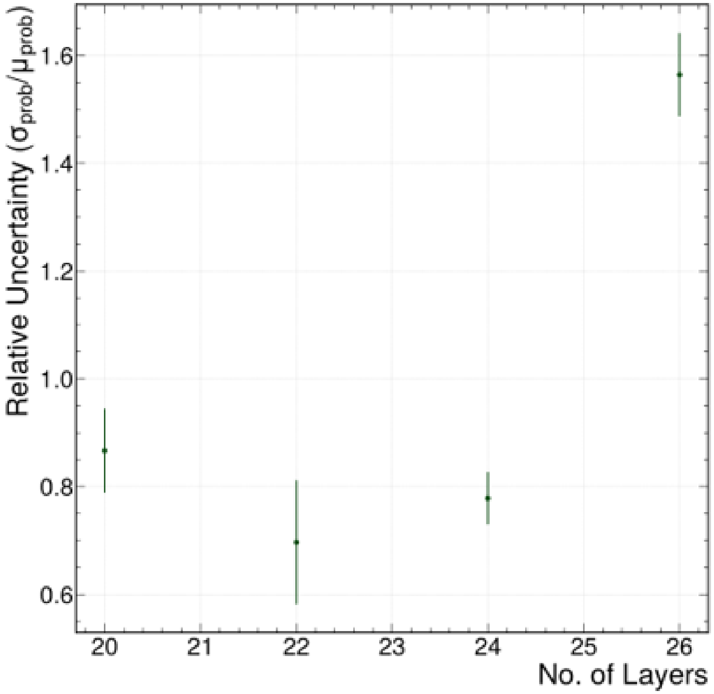
Turning/optimal points ($\lambda^{n}_{H_2O}$=9.62 cm @ 10 MeV, $\lambda^{n}_{H_2O}$=7.14 cm @ 5 MeV):
- **18 layers @ 10 MeV: $(0.92 \pm 0.05)\lambda^{n}_{H_2O}$**
- **23 layers @ 5 MeV: $(0.91 \pm 0.03)\lambda^{n}_{H_2O}$**

Same behaviour seen for neutrons, for different values

10 MeV



5 MeV



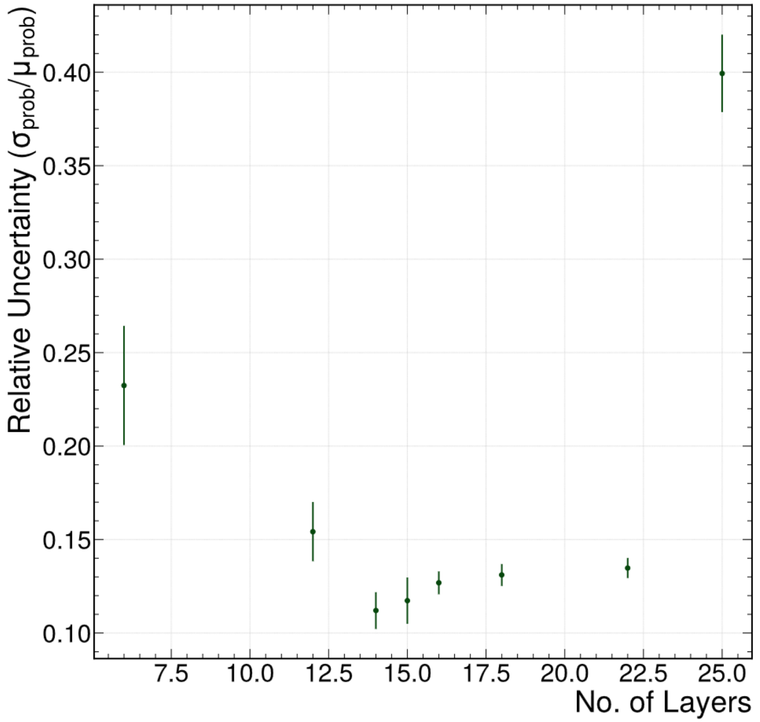
Generalising Optimisation Guidance

Test for generalisation over:
- Shielding material
- Shielding thickness

Photon simulations @ 1461 keV:
- Lead not water, 13.2 cm thick
- Water shielding 120 cm thick

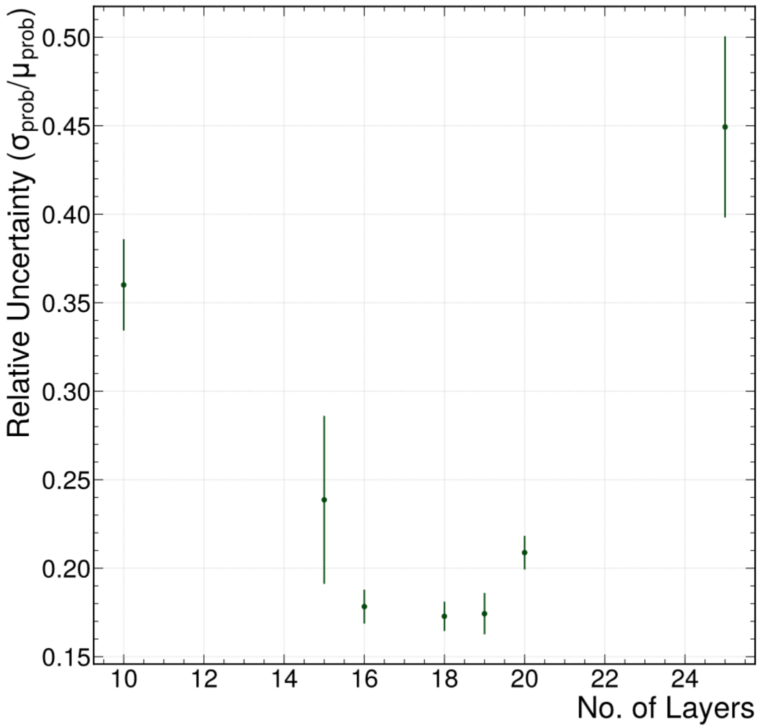
Outcome:
- **18-19 layers for Lead shielding** $(0.44 \pm 0.02)\lambda^{\gamma}_{Pb}$
- **14-15 layers for 120 cm water** $(0.50 \pm 0.03)\lambda^{\gamma}_{H_2O}$

Matches predictions – generalise over shielding material and thickness



Different thickness (120 cm) @ 1461 keV

Predict: optimal @ $\sim 0.48 \times \lambda^{\gamma}_{H_2O}$ or 14-15 layers



Lead (13.2 cm) @ 1461 keV

Predict: optimal @ $\sim 0.48 \times \lambda^{\gamma}_{Pb}$ ($\lambda^{\gamma}_{Pb} = 1.66$ cm) or 18-19 layers

What is the Computational Gain?

Have seen for 2750 keV photon case:
- At **300s ~0.08 RU achieved** in most optimal case
- Compared to **~0.18 at 5 layers, ~0.34 at 25 layers**
- For **unbiased: ~0.5 for 300s**

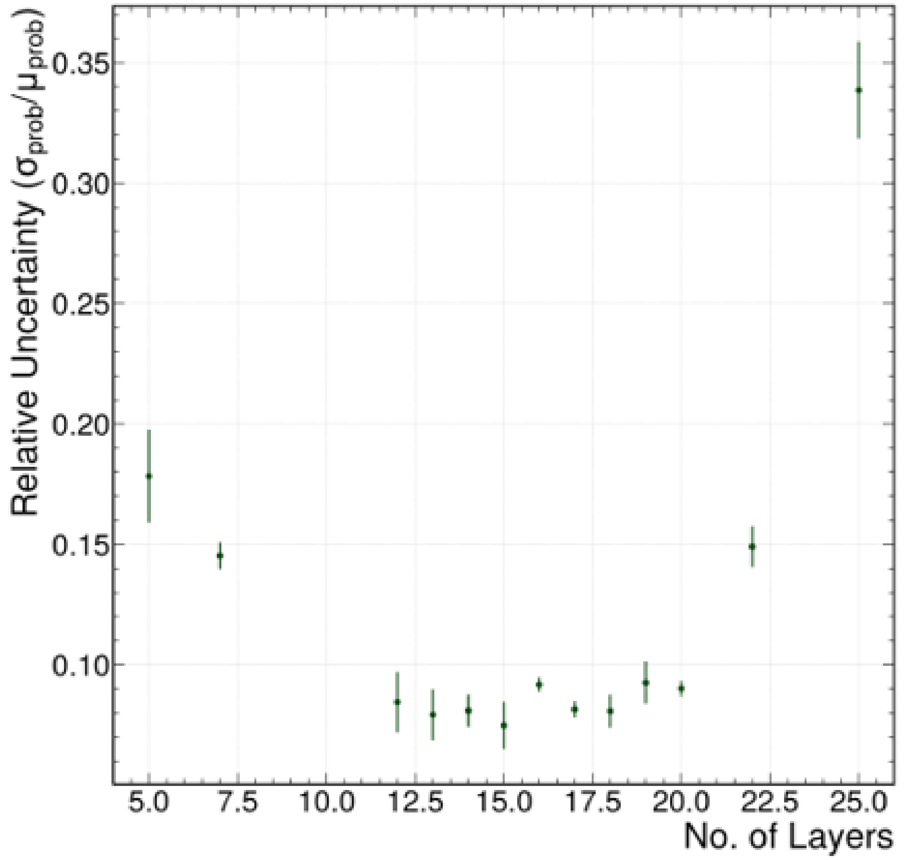
Gain in computational efficiency key for managing CO2 footprint

What is the computational gain of optimal case?

Set 0.08 as fixed RU:
- Optimal case takes 300 s to reach 0.08 RU
- 5 layers takes ~1268 s
- 25 layers takes ~5544 s
- **Unbiased takes ~30000 s**

How does this translate to a reduced CO2 footprint?

2750 keV



Computing Footprint in HEP

UNIVERSITY OF BIRMINGHAM

Computing footprint a large proportion of professional footprint

- Relatively higher for researchers engaged in rare-event searches

Can estimate footprint:

$$\text{Footprint }[\text{gCO}_2\text{e}] = \boxed{f_{\text{PUE}}} \times \boxed{n_{\text{WPC}}} \times \boxed{f_{\text{conv}},}$$

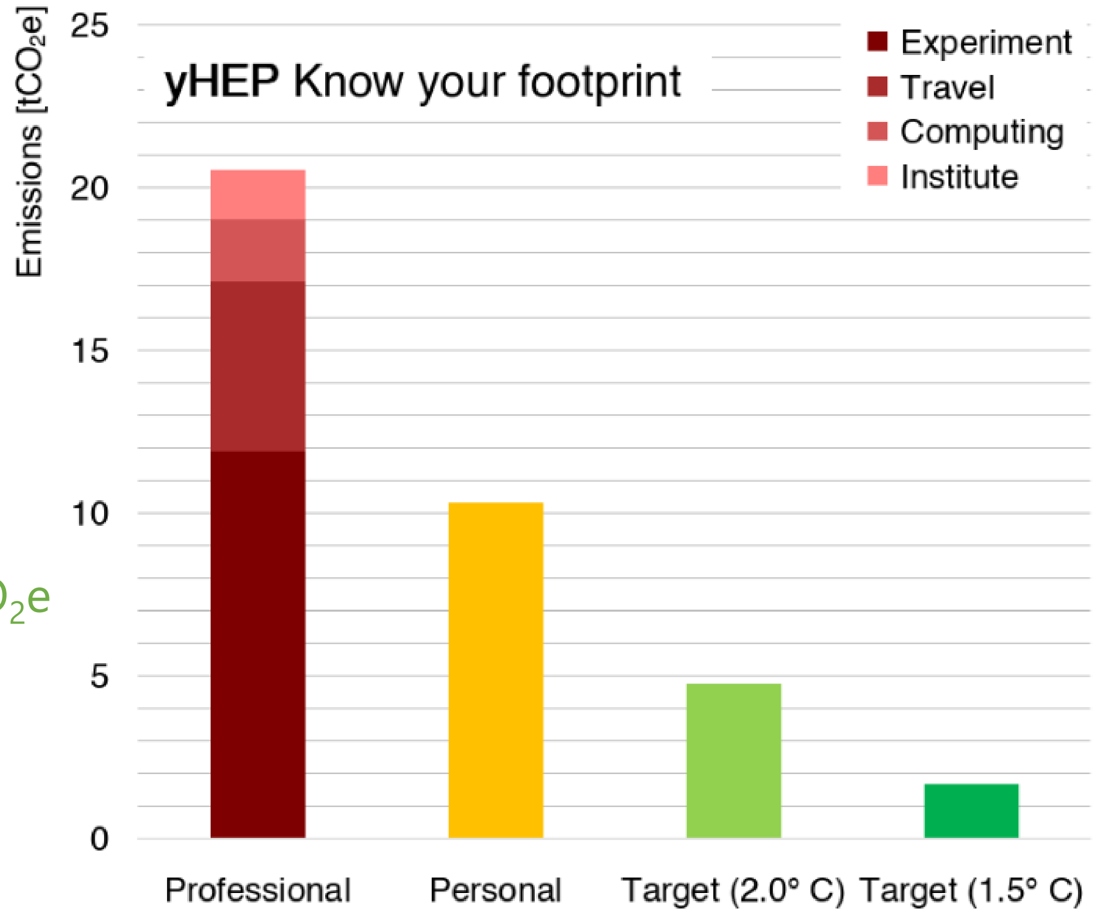
Power Usage Effectiveness of a HPC cluster

Conversion of kWh to gCO_2e (grid dependent)

$$n_{\text{WPC}} = \boxed{p_{\text{CPU-core}}} \times \boxed{t_{\text{core-h}},}$$

Power consumption of CPU core in kW

Workload of core in hours



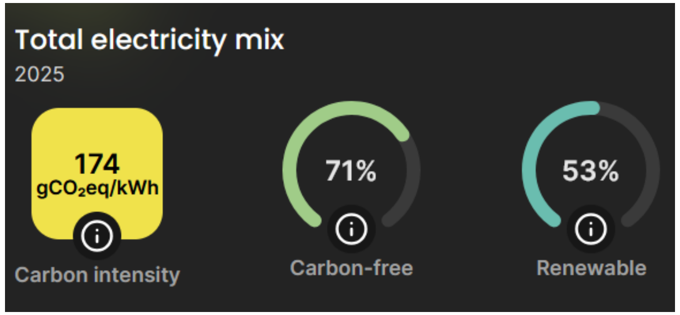
V.S Lang et al. arXiv:2403.03308, doi:10.1038/s44168-025-00232-7

Reduced Footprint due to Optimised Biasing

UNIVERSITY OF BIRMINGHAM

The computational gain from optimal biasing over an unbiased, under/over-biased simulation is large
- Expect large reduction in carbon footprint too!

Utilise Birmingham cluster for comparison:
- $f_{PUE} = 1.5$ (global average)
- $f_{conv} = 174$ gCO_2e/kWh
- $p_{CPU-core} = 6.25$ W
- t_{core-h} will vary between 3 cases, based on time taken to reach fixed RU



Time taken to reach fixed RU of 0.081 used for comparison

$$\text{Footprint [gCO}_2\text{e]} = 1.63 \times t_{\text{core-h}}.$$

Note: footprint is calculated on per core basis

Biasing Level	Footprint [gCO_2e/core]	Factor Increase on Optimal
Optimal (14L)	0.135	N/A
Under-biased (5L)	0.571	4.23
Over-biased (25L)	2.510	18.59
Unbiased	13.58	$\sim 10^2$

Large decrease in footprint if optimal layers used!

Summary

DM searches are requiring increasingly low backgrounds to achieve high exposure and low thresholds

Shielding simulations increasingly intractable without event biasing
- Need **optimised biased simulations, improve computational efficiency,** and **manage carbon footprint**

Commonly used biasing scheme: **Importance Splitting/Killing**
- **Specific guidance on optimal layer configurations is lacking**

DarkSPHERE-30 detector as a test-bed, more specific guidance found:
- At high enough numbers of layers, **relative uncertainty reverses expected trend**
- **Optimal point is turning point in this trend**, corresponding to **certain layer thicknesses**:
    - **$\sim 0.45\lambda^{\gamma}_{H_2O}$ for photons & $\sim 0.91\lambda^{n}_{H_2O}$ for neutrons**
- **Generalisable to different shielding materials and thicknesses**

Using optimal layer thickness can yield a **factor of ~100 reduction in computing carbon footprint**
- **Poorly optimised biasing can also lead to larger footprints**

Full publication to be submitted (on arXiv) very soon!!!