AI ASSISTED DISCOVERY OF PHYSICAL VALIDITY BOUNDS IN COLLECTIVE BEAM EFFECTS SIMULATIONS

O. H. Ramachandran[1], G. Ha[1,2], C-K. Huang[3], X. Lu[1,2], J. Power[2], Ji Qiang[4] and Alec Tilton[2]

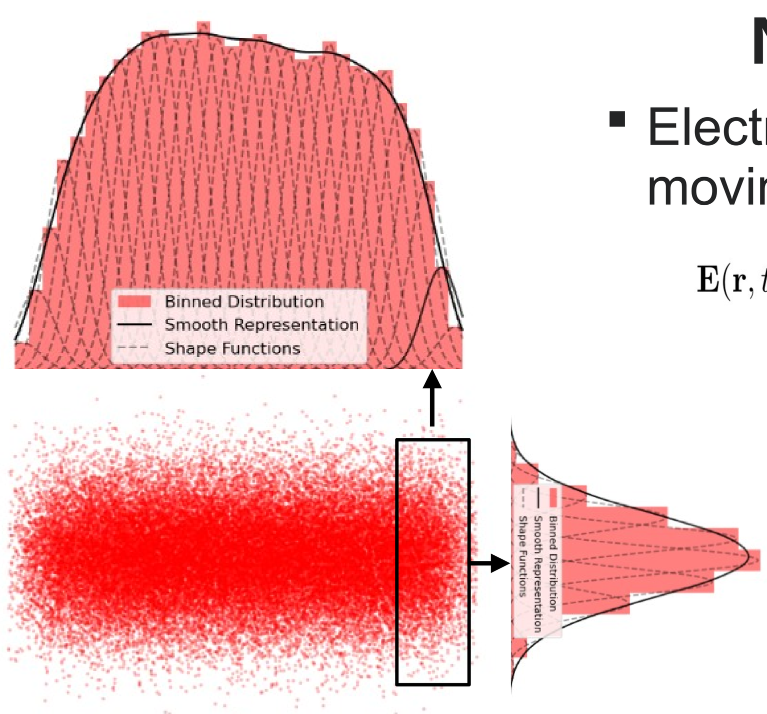
1 Argonne National Laboratory, Lemont, IL 60439, USA
2 Northern Illinois University, DeKalb, IL 60115, USA
3 Los Alamos National Laboratory, Los Alamos, NM 87545, USA
4 Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA

ABSTRACT

- Phase space degradation due to Coherent Synchrotron Radiation (CSR) is a challenge in the generation of high brightness beams.
- Existing simulation tools (and analysis in the literature) typically limit themselves to 1D, where the beam transverse size is small compared to the bunch length.
- In prior work [1], we have developed a 3D CSR code that is capable of efficiently calculating the fields generated by a bunch in a bending magnet while including the effects of entry/exit wakes and shielding by parallel plate walls.
- In this work, we outline the development of an agentic AI harness where an LLM (GPT-5.2) is used to intelligently sample a defined parameter space following instructions in natural language.
- The obtained dataset is used to understand the limits of the Derbenev criterion as a measure of validity for 1D CSR for different beam conditions/shielding configurations.

FORMULATION

Numerical Formulation



- Electric/Magnetic fields are calculated on a moving mesh following Jefimenko's equations:

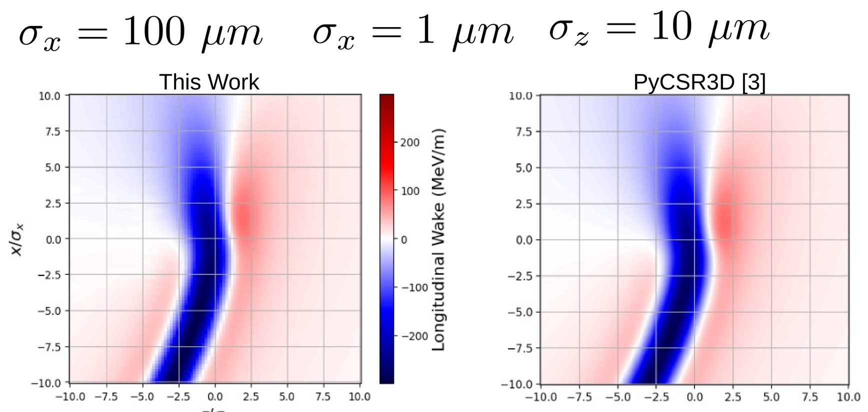
$$\mathbf{E}(\mathbf{r},t)=\frac{1}{4\pi\epsilon_0}\int d\mathbf{r}'\left[\frac{\mathbf{R}}{|\mathbf{R}|}\rho(\mathbf{r}',t_r)+\frac{\mathbf{R}}{c|\mathbf{R}|^2}\partial_t\rho(\mathbf{r}',t_r)-\frac{1}{c^2|\mathbf{R}|^2}\partial_t\mathbf{J}(\mathbf{r}',t_r)\right]$$

- The bunch profile in space is described using smooth Gaussian shape functions

$$\rho(\mathbf{r}',t)=Q\sum_i S_i\left(\mathbf{r}'(t)\right)$$

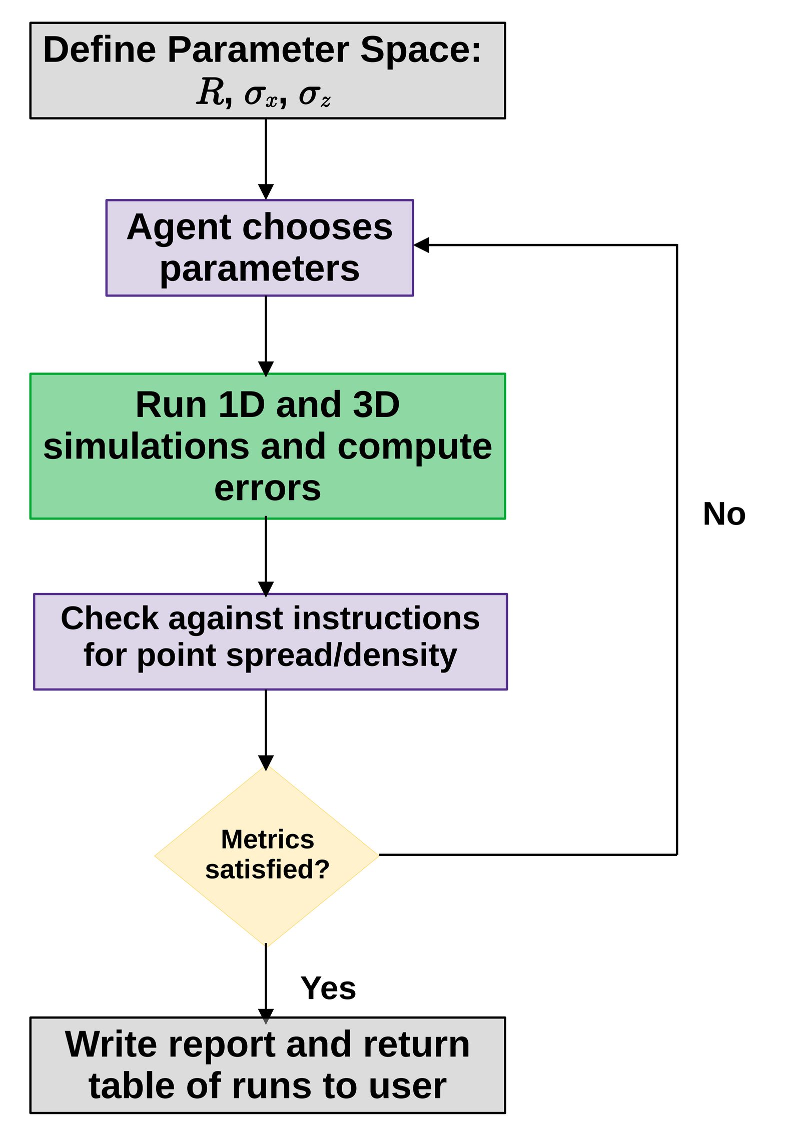
$$\mathbf{J}(\mathbf{r}',t)=Q\partial_t'\mathbf{r}'(t')\sum_i S_i\left(\mathbf{r}'(t)\right)$$

Fig: Steady-state wake compared to PyCSR3D for a Gaussian*3 with $\sigma_x=100\ \mu m \quad \sigma_x=1\ \mu m \quad \sigma_z=10\ \mu m$



- Shielding by parallel plates is achieved with layers of image charges.
- The particle evolution through the magnet is continually binned, resulting in some self-consistency.
- The current prototype code is **GPU-accelerated** through the CuPy[4] package.

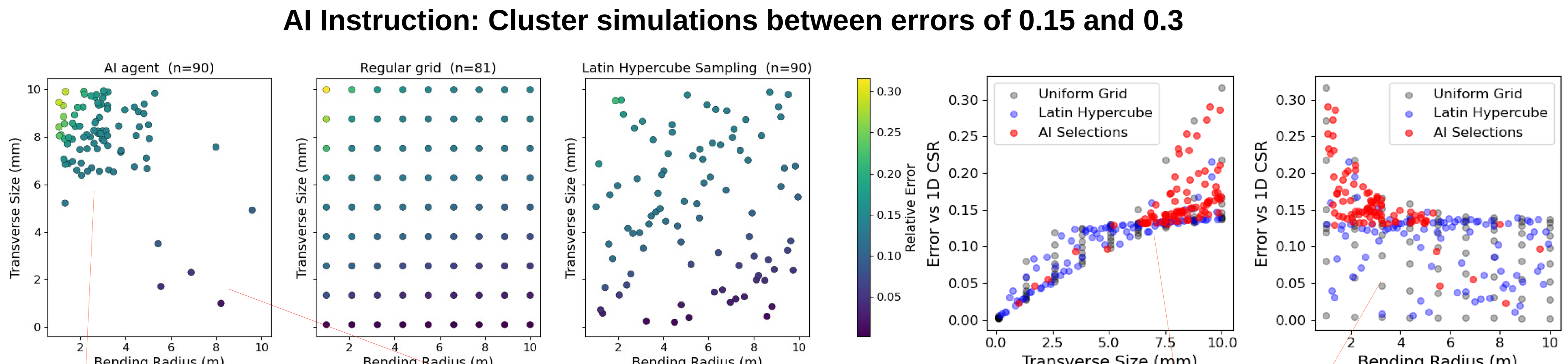
LLM-Augmented Parameter Search



- The LLM augments our parameter search by allowing the user to specify chains of instructions to the AI on how it should sample the search space.
- Initially Latin Hypercube Sampling is used to broadly sample the search space. Then, the LLM is free to propose new parameters and run comparisons.
- All simulation runs (input and output metrics) are stored in CSV files and are imported into the LLM's context.
- The LLM used was GPT-5.2 [2]. In general, we find that smaller models (say Llama/Mistral with 8b parameters or lower) tend to hallucinate once the context gets large (after 20-30 simulations).

TARGETED SIMULATIONS WITH AN AI AGENT

AI Instruction: Cluster simulations between errors of 0.15 and 0.3

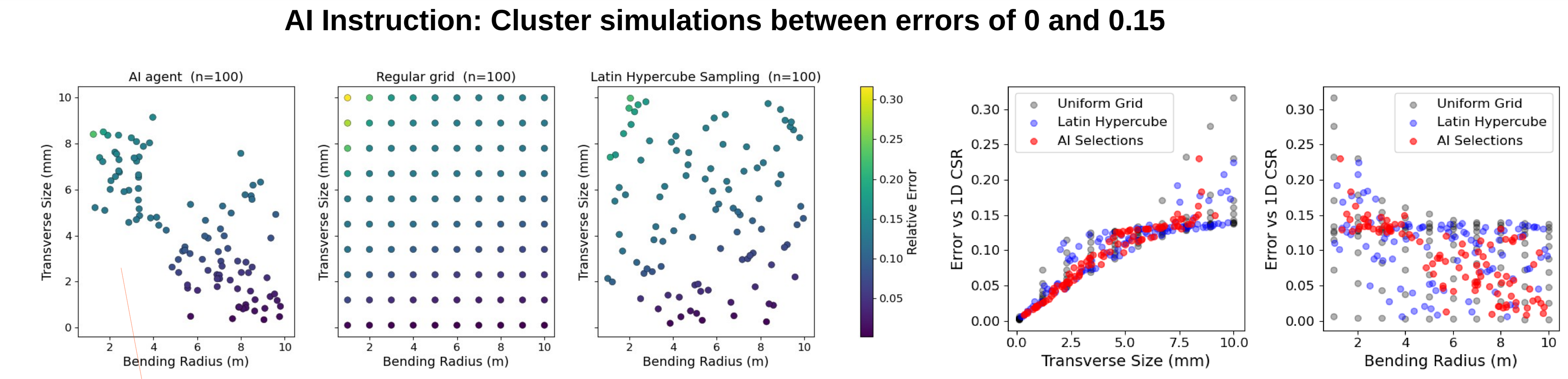


Once it figures out a region with the right constraints, it can autonomously focus there

The AI initially samples the entire space with 10 samples following a Latin Hypercube

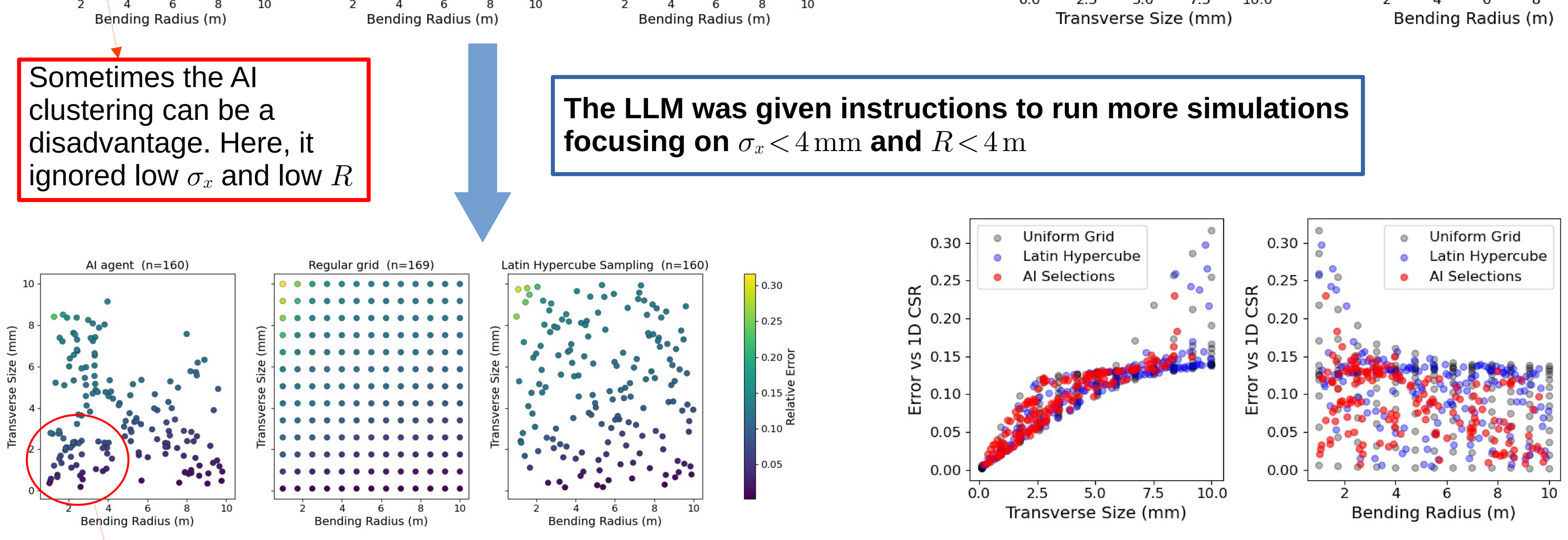
Notice the clustering in the AI agent compared to the regular/hypercube sampling.

AI Instruction: Cluster simulations between errors of 0 and 0.15



Sometimes the AI clustering can be a disadvantage. Here, it ignored low σ_x and low R

The LLM was given instructions to run more simulations focusing on $\sigma_x < 4\,\mathrm{mm}$ and $R < 4\,\mathrm{m}$



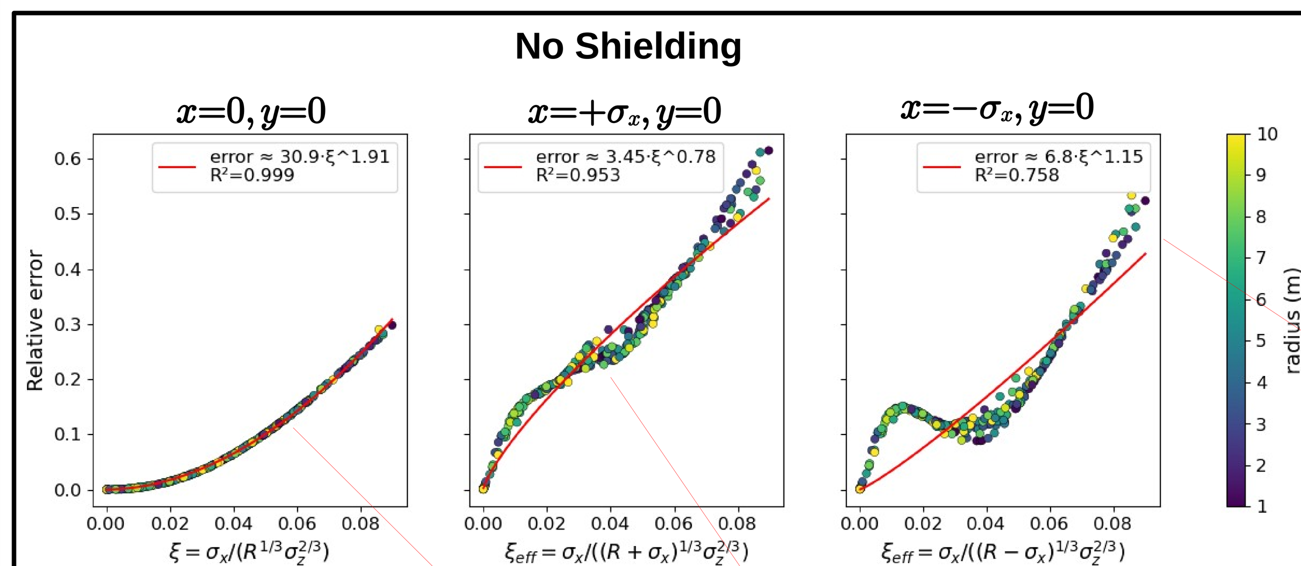
With that additional instruction, the AI filled in the missing region of the search space.

In general, this task can be set up algorithmically, but the AI agent gives us a means to simply write English instructions and mostly obtain the desired behavior, saving a lot of time.

ANALYSIS OF DERBENEV CRITERION

Next, we analyzed the data and tried to fit power laws of the form $\epsilon = A\xi^p$, where $\xi = \sigma_x/(R\sigma_z^2)^{1/3}$ is the standard Derbenev [5] criterion.

No Shielding

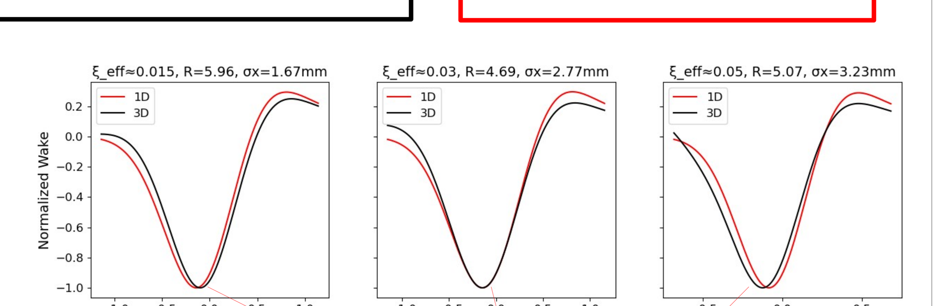


The key takeaways here are that Derbenev is generally good without shielding, but does not capture strong retardation effects. With shielding, a different power law scaling needs to be derived.

When offset towards the center of curvature of the magnet, a simple power law stops working due to offsets caused by the 3D retardation condition.

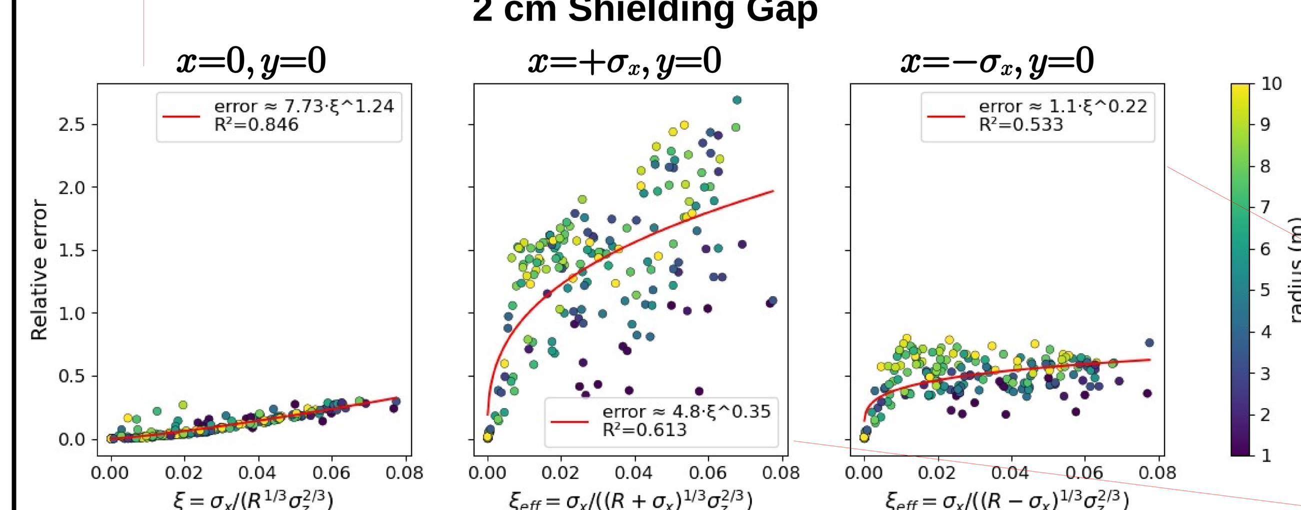
With a shielding gap, the Derbenev criterion is still a good power law base through the center of the bunch, though with a lower R^2

With no shielding, the Derbenev criterion completely captures the error through the centerline of the bunch. At $x=+\sigma_x$, a modified version works almost as well.



The shifts in the wake between 1D and 3D are caused by the more complex retardation condition in 3D. This is what causes the saddle at $\xi_{\mathrm{eff}}\approx 0.04$

2 cm Shielding Gap



Off-axis, the power law fit is much worse. Note that higher R for a given ξ clusters at the top of the spread. This probably means the physics follows a different power law base.

REFERENCES

- Ramachandran et al., in Proc. IPAC2026, Normandy, France
- OpenAI. (2025). ChatGPT (GPT-5.2) [Large language model]. https://chat.openai.com/chatgpt
- C. Mayes, "ChristopherMayes/PyCSR3D: PyCSR3D Version 0.2.0," v0.2.0, Zenodo, Sep. 2021. doi: 10.5281/zenodo.5496096.
- R. Okuta, et al., in Proc. Workshop on Machine Learning Systems (LearningSys) at NIPS 2017, 2017.
- Y. S. Derbenev, et al., DESY Report TESLA-FEL 95-05, Hamburg, 1995.

U.S. DEPARTMENT OF ENERGY
Argonne National Laboratory is a U.S. Department of Energy laboratory managed by UChicago Argonne, LLC.