

u^b

^b
UNIVERSITÄT
BERN

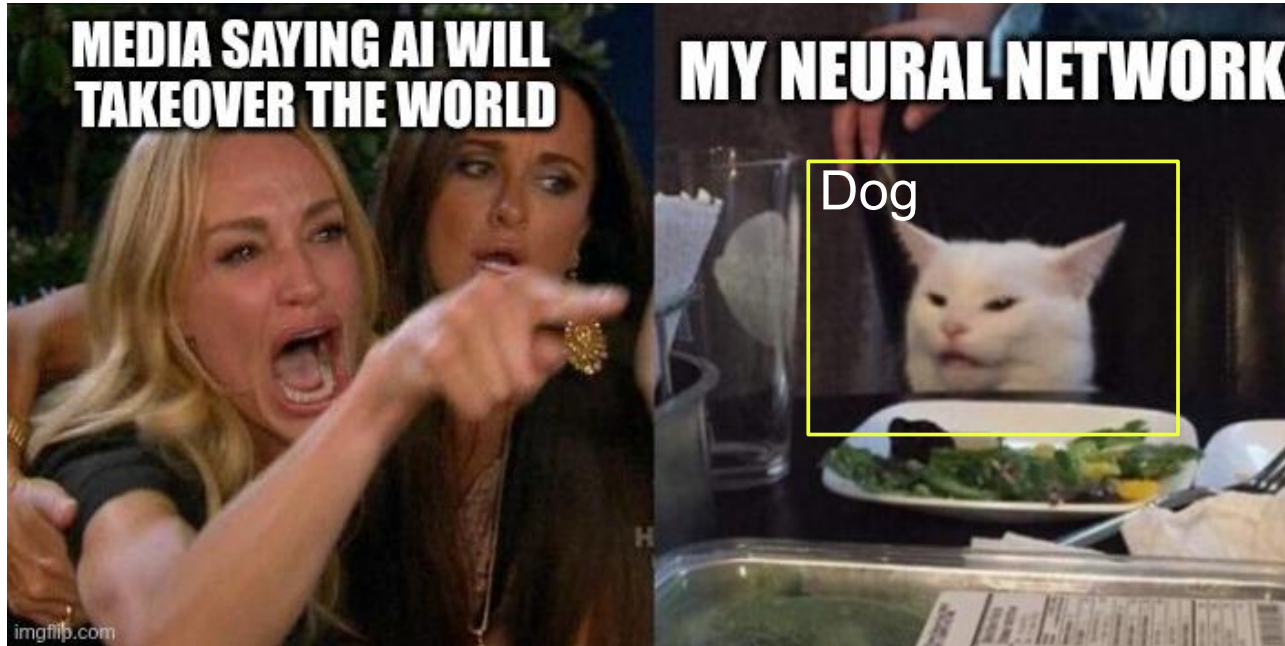
AEC
ALBERT EINSTEIN CENTER
FOR FUNDAMENTAL PHYSICS

SIREN training on FSD LUT

Dominik Wermelinger



Feedback & Ideas highly welcome!

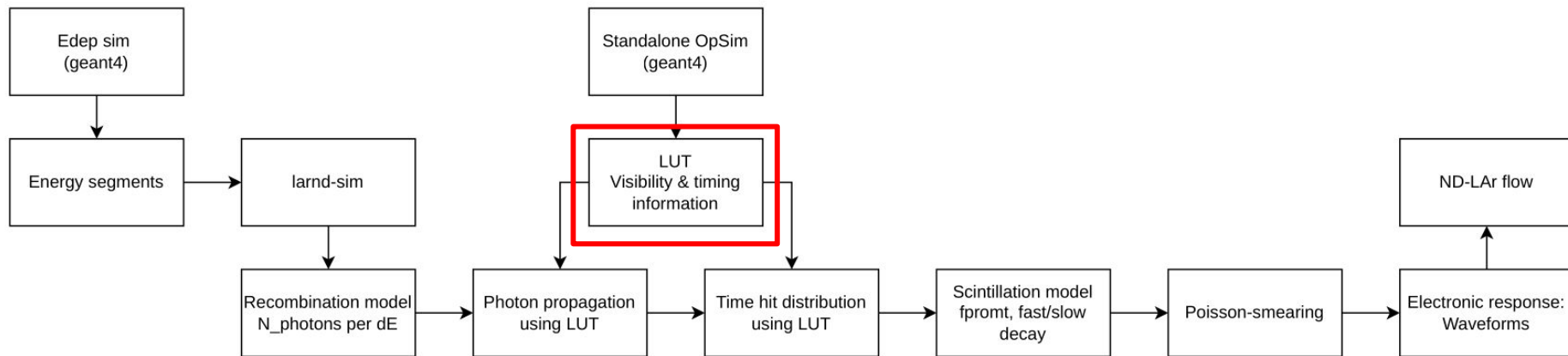


LUT's usage in simulations

Why Light Lookup Tables (LUT)?

- Photons simulations in LArTPC's computationally expensive (~20-30k photons/MeV)
- Precompute **visibility** and **TOF** of voxelized detector with MC simulation -> LUT

(x,y,z) -> (visibility, time distribution of SiPM hits)



SIREN - Implicit Neural Representations with Periodic Activation Functions

Siren:

- Train on LUT from simulation, but also on data! -> Data-driven LUT
- LUT of FSD is 5GB, Siren models ~ 10-100 MB

Resources and pointers:

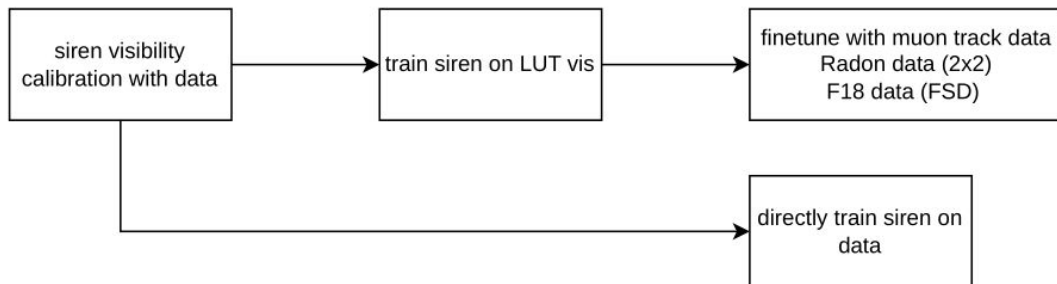
- Original Siren paper: (<https://arxiv.org/pdf/2006.09661>)
- Icarus paper (<https://arxiv.org/pdf/2211.01505>)
- Siren repo (<https://github.com/CIDeR-ML/siren-lartpc>)

Computing:

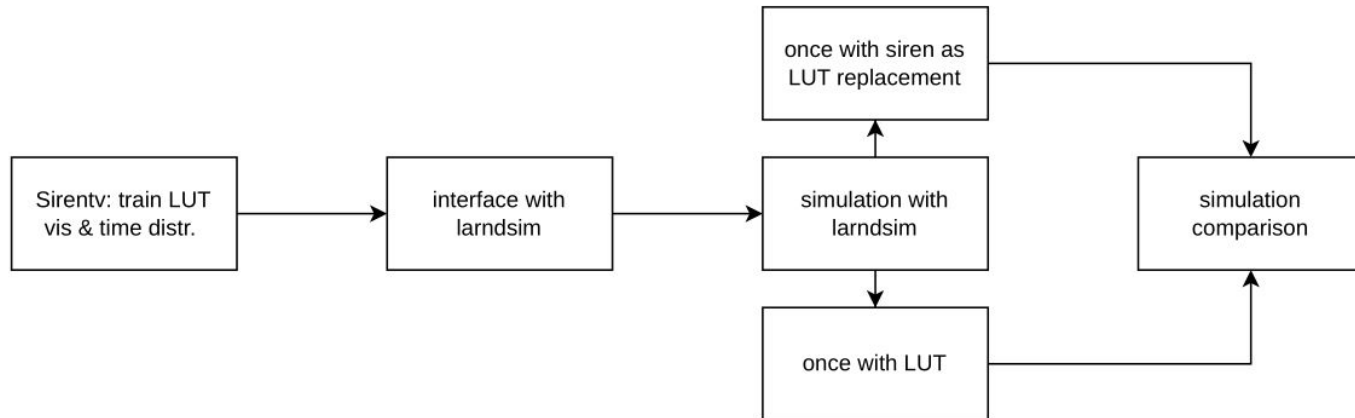
- Ran on 1x Nvidia RTX 2080 Ti GPU (local lhep-cluster)

Motivation and way forward

First stage:

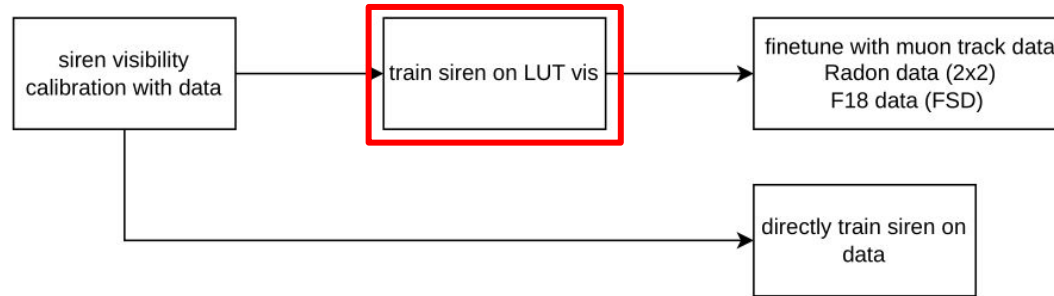


Second stage:

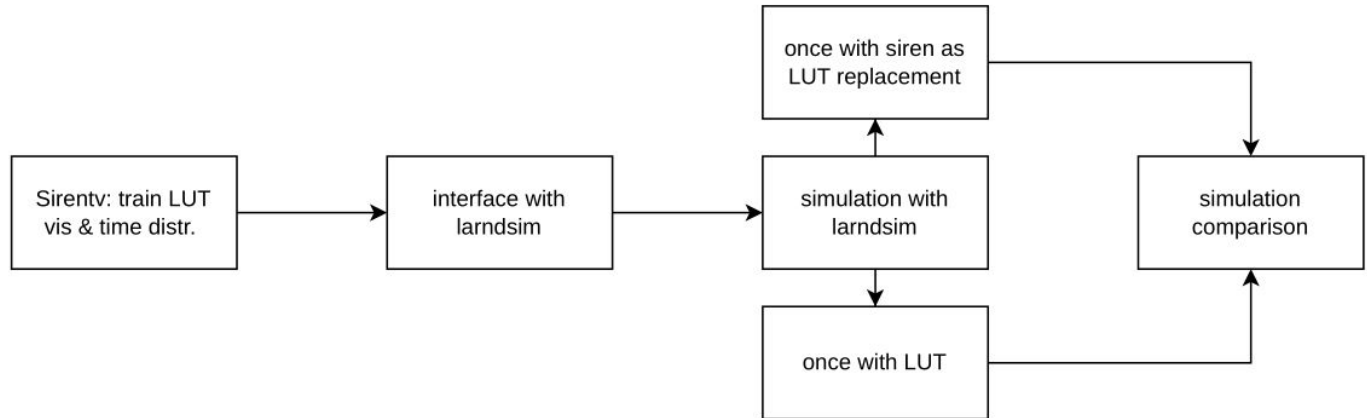


Motivation and way forward

First stage:



Second stage:



FSD LUT - Overview

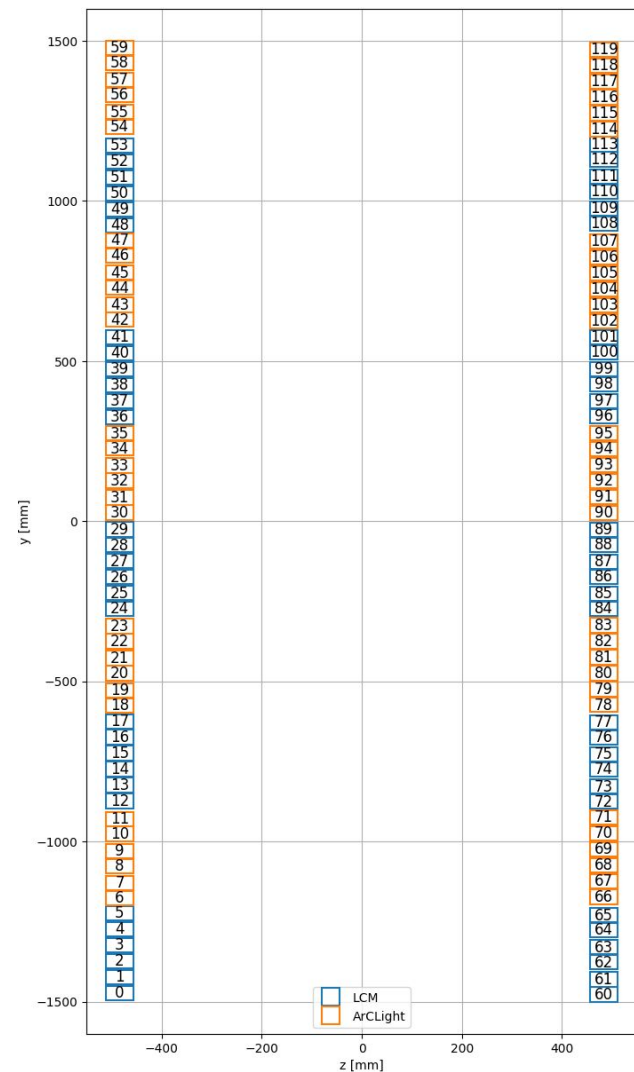
Contents of FSD LUT

- Voxel size $\sim 1.5 \times 1.5 \times 1.5 \text{ cm}^3$
- Per voxel 10^7 photons emitted
- $(n_x, n_y, n_z, n_{ch}) = (32, 200, 64, 120)$

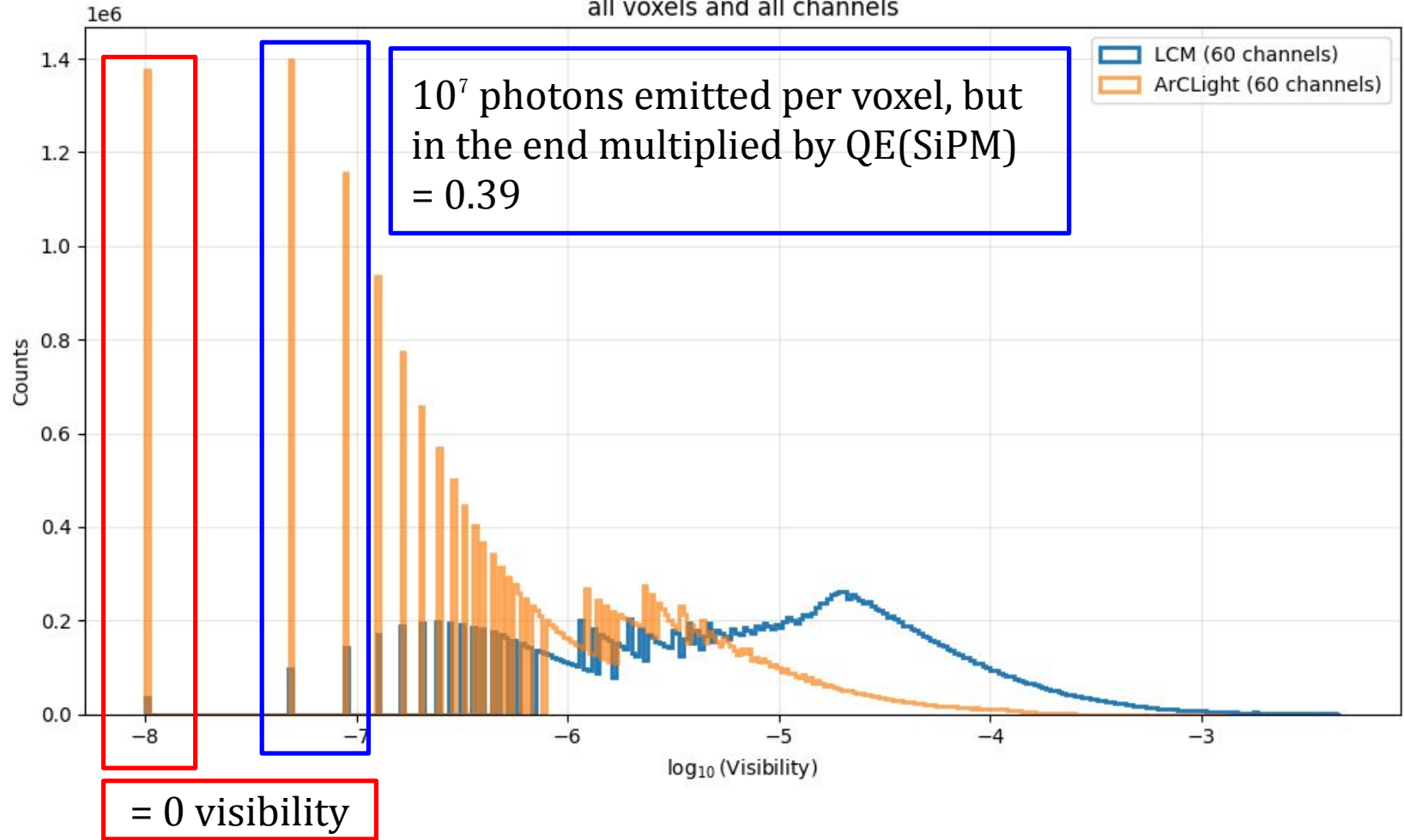
For each of the 120 channels:

- visibility (fraction of photons registered by a given SiPM)
- t_0 (shortest travel time)
- t_{0_mean} (mean arrival time)
- time_distr (time distribution of hits)

Full FSD LUT is **$\sim 5 \text{ GB}$** of data, only visibilities are **$\sim 300 \text{ MB}$**

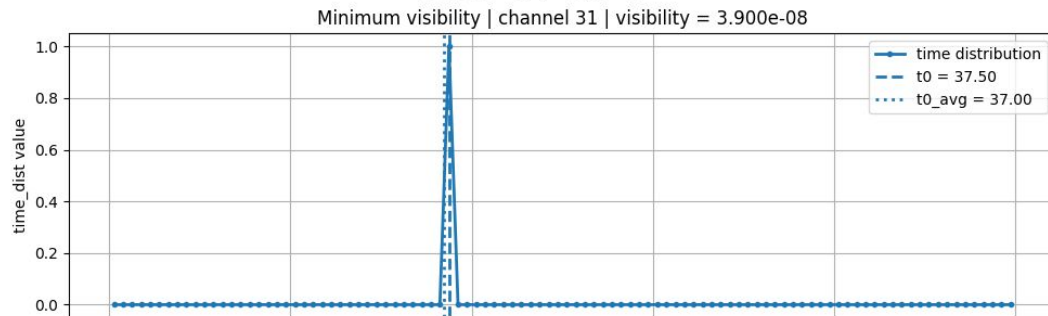


Visibility distribution
all voxels and all channels

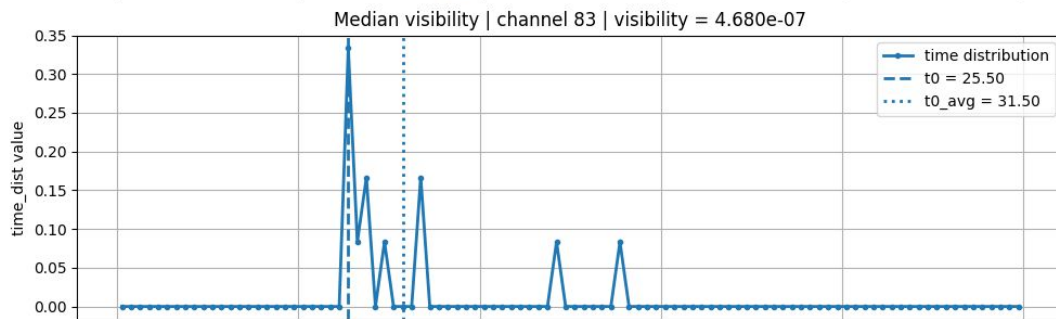


LUT - time distribution

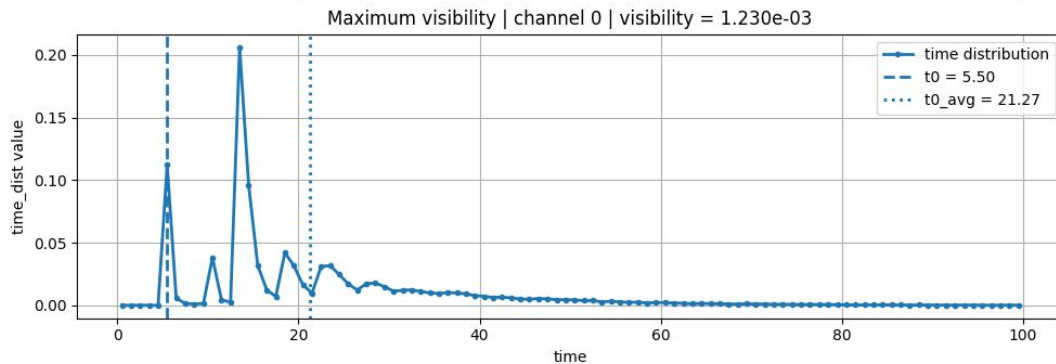
min visibility of voxel



median visibility of
voxel



max visibility of voxel

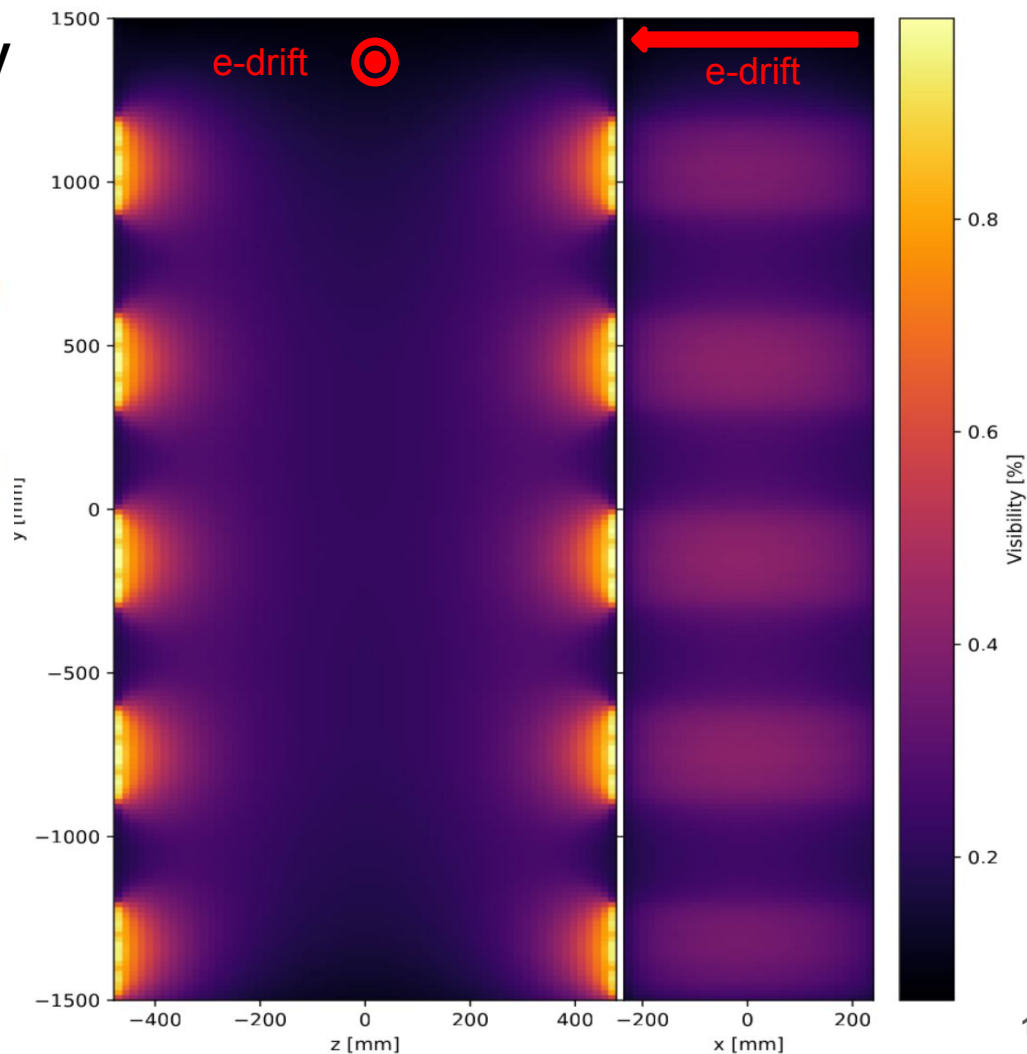


FSD LUT - mean projected visibility

$$ZY(z, y) = \frac{1}{N_x} \sum_{x=1}^{N_x} \sum_{n_{ch}=1}^{120} \text{vis}(x, y, z, n_{ch})$$

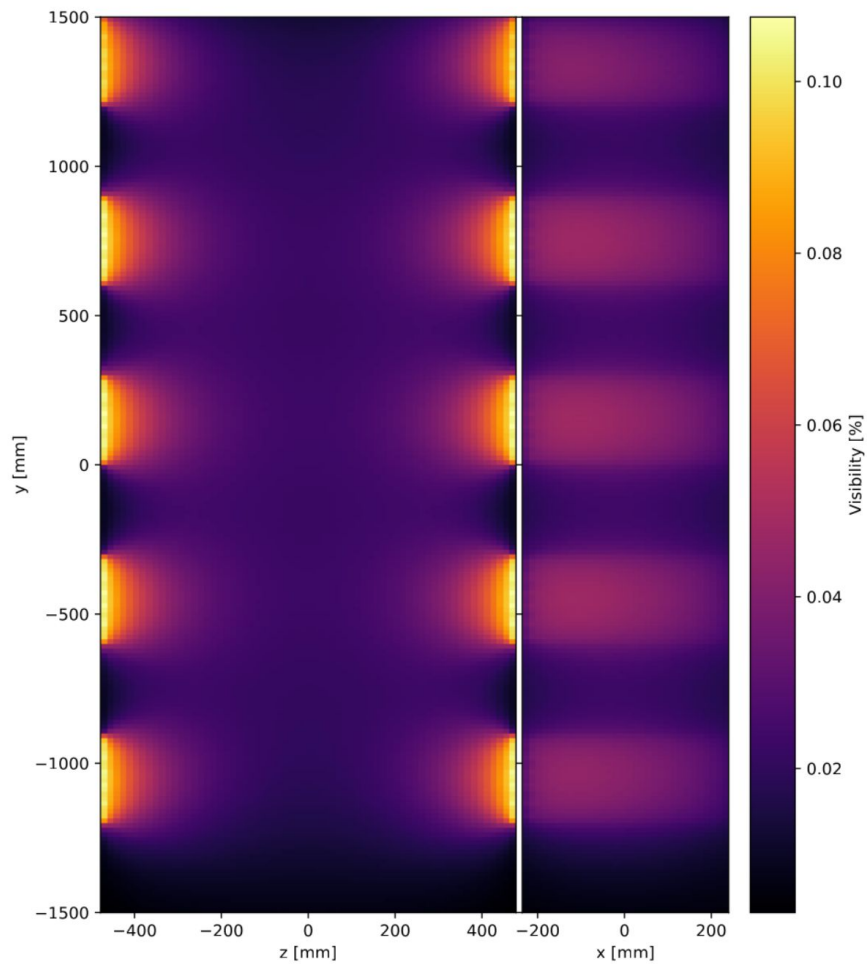
$$XY(x, y) = \frac{1}{N_z} \sum_{z=1}^{N_z} \sum_{n_{ch}=1}^{120} \text{vis}(x, y, z, n_{ch})$$

“How well is every voxel seen by all channels”

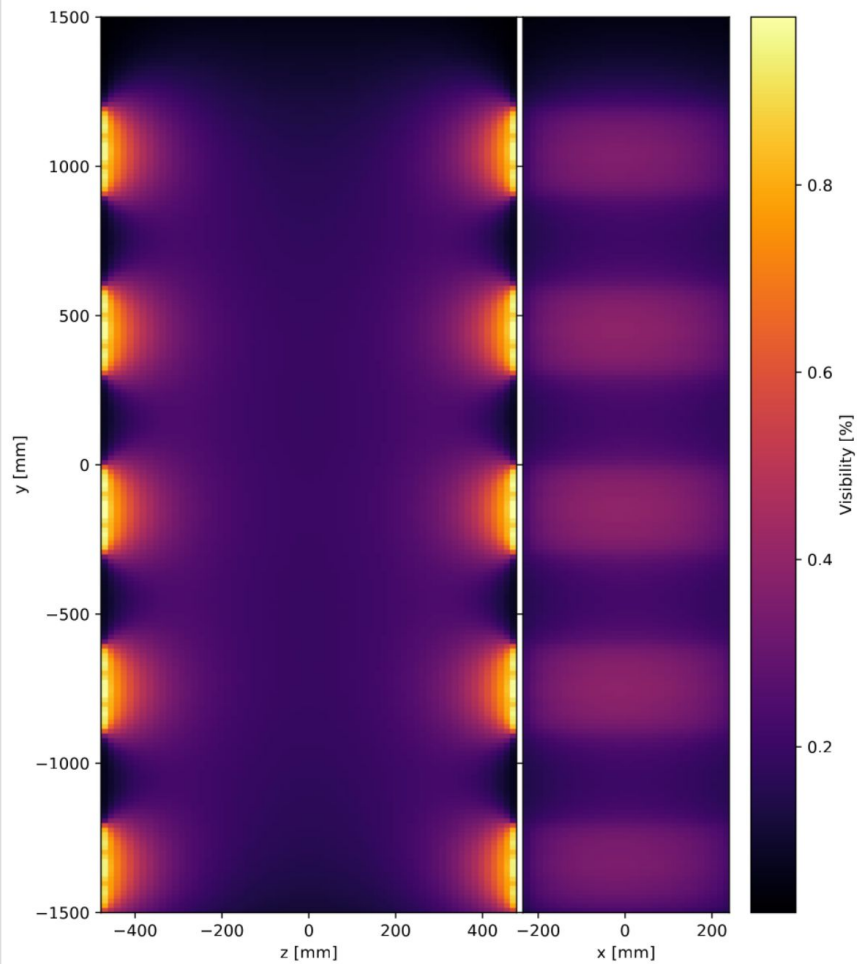


FSD - LUT

ACL



LCM



Box Plots convention

Boxplots:

$$P_{25}(\text{vis}) < \text{vis} \leq P_{75}(\text{vis})$$

Whiskers:

$$\text{lower whisker} = P_5(\text{vis})$$

$$\text{upper whisker} = P_{95}(\text{vis})$$

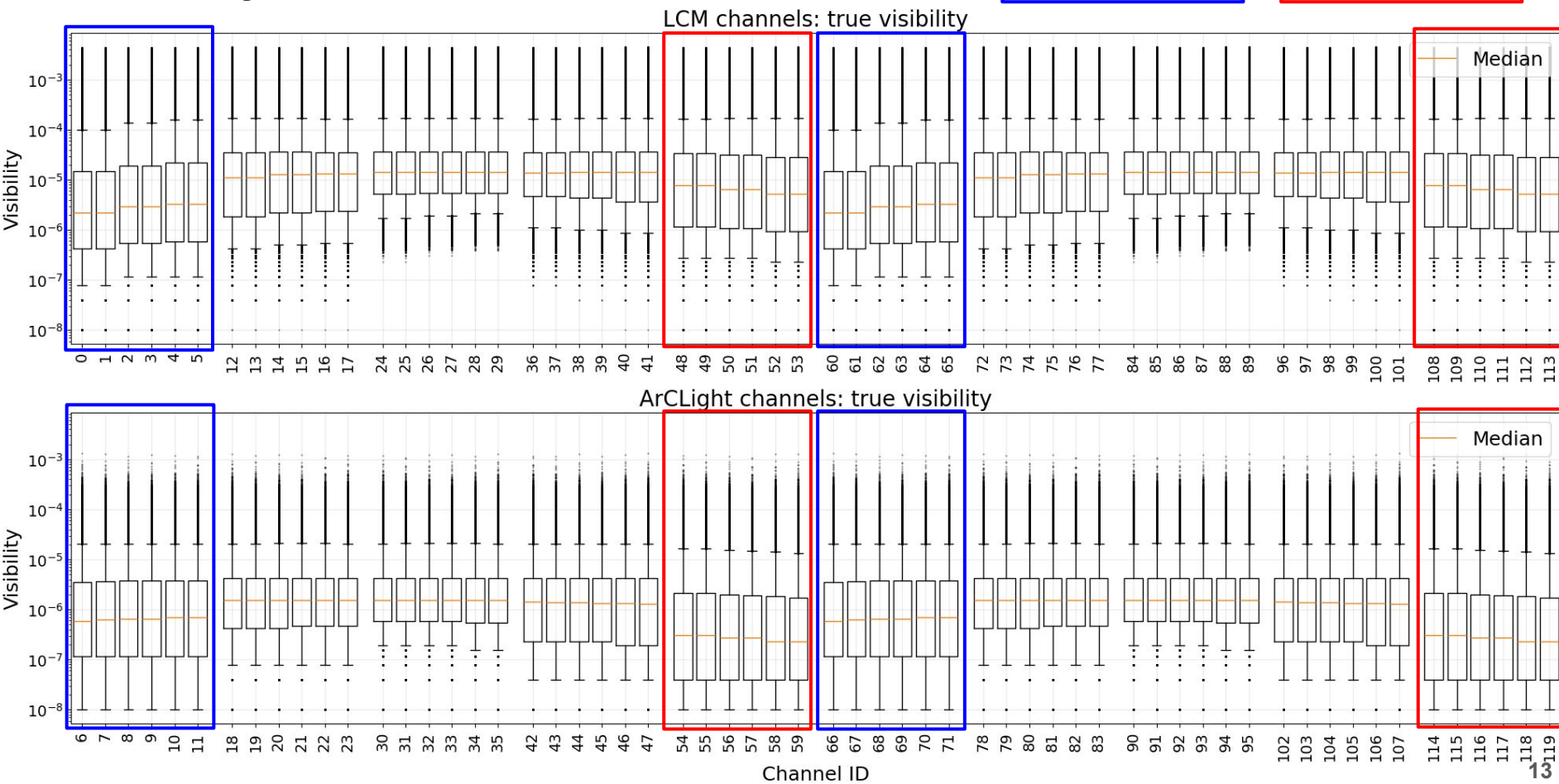
Outliers:

$$P_5(\text{vis}) < \text{vis} < P_{95}(\text{vis})$$

Visibility distribution for each channel

most bottom LT

most top LT



First results of a SIREN training on FSD LUT

Hyperparameters:

- Number of hidden layers $n_L = 10$
- Number of features $n_F = 512$
- “sine frequency” $w_0 = 30$
- validation fraction **5%**
- max training epochs **10'000**
- vis log transformation **eps = 1e-7**

First layer

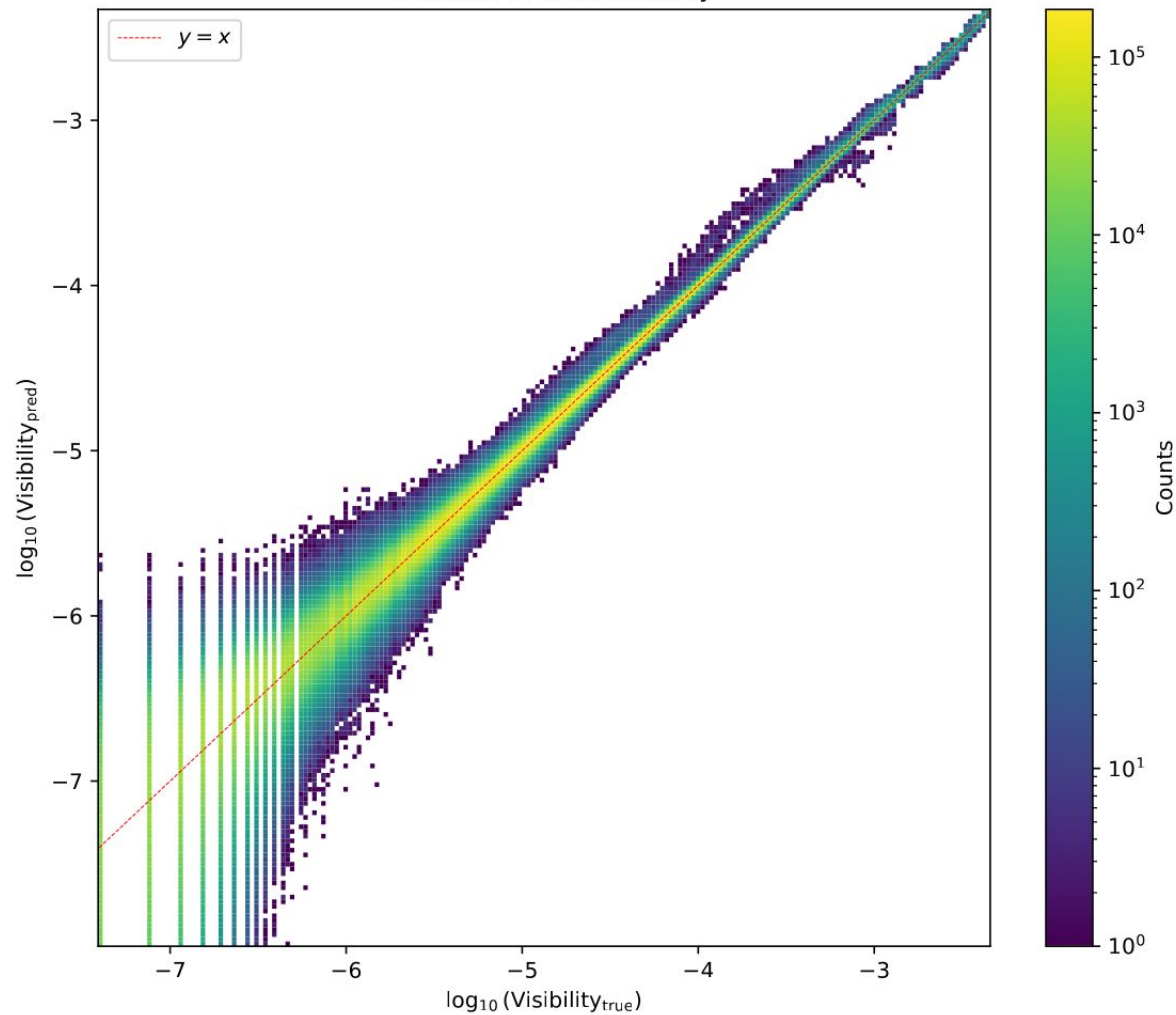
$$h_i^{(0)} = \sin \left(\omega_0 \left(\sum_{j=1}^3 W_{ij}^{(0)} x_j + b_i^{(0)} \right) \right)$$

Loss: weighted-MSE

$$v = \frac{\log_{10}(v_{raw} + \epsilon) - \log_{10}(\epsilon)}{\log_{10}(v_{max} + \epsilon) - \log_{10}(\epsilon)}$$

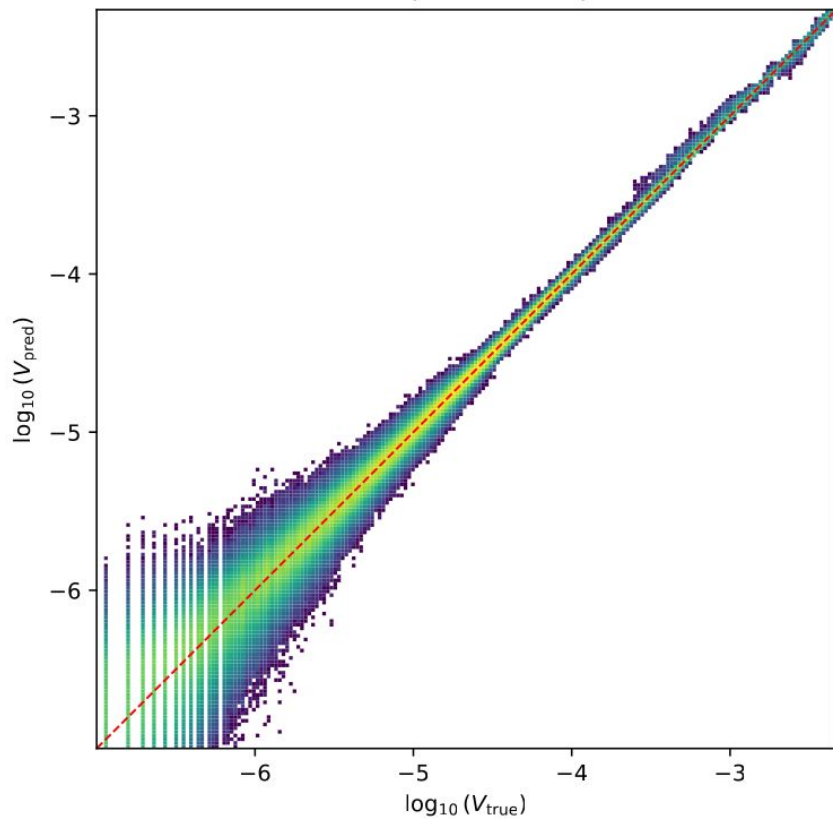
$$L = \frac{1}{NC} \sum_{i=1}^N \sum_{j=1}^C w_{ij} (\hat{v}_{ij} - v_{ij})^2$$

Predicted vs True Visibility

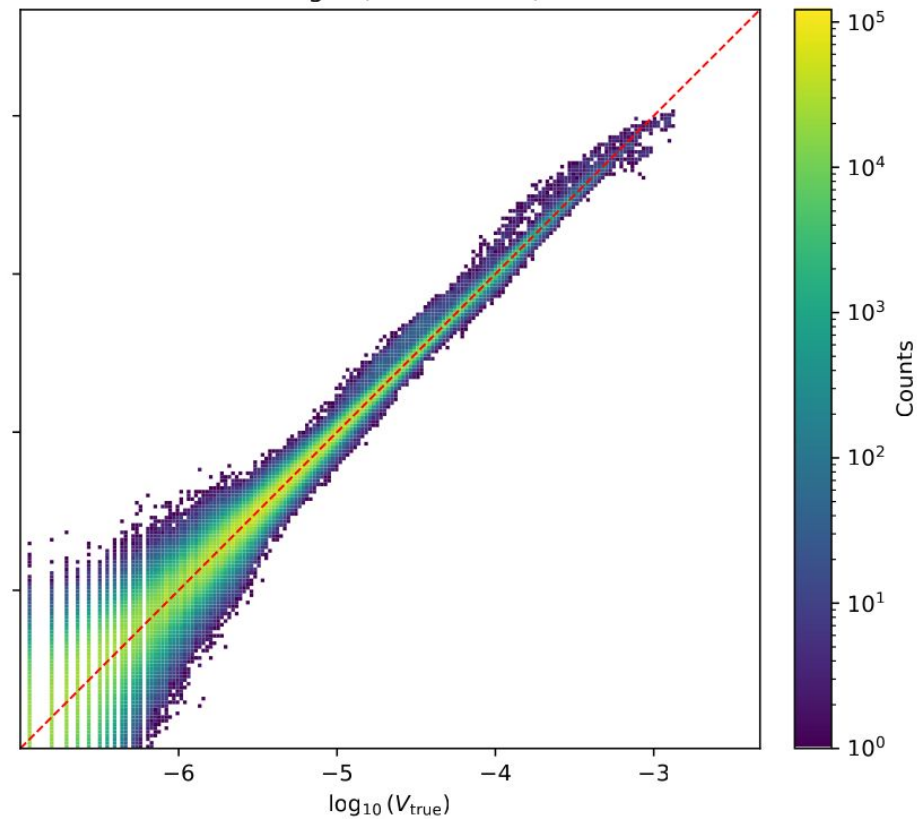


Predicted vs True Visibility

LCM (60 channels)



ArCLight (60 channels)



Model metrics

Absolute bias (per channel & voxel)

$$b_{\text{abs}} = |v_{\text{pred}} - v_{\text{true}}|$$

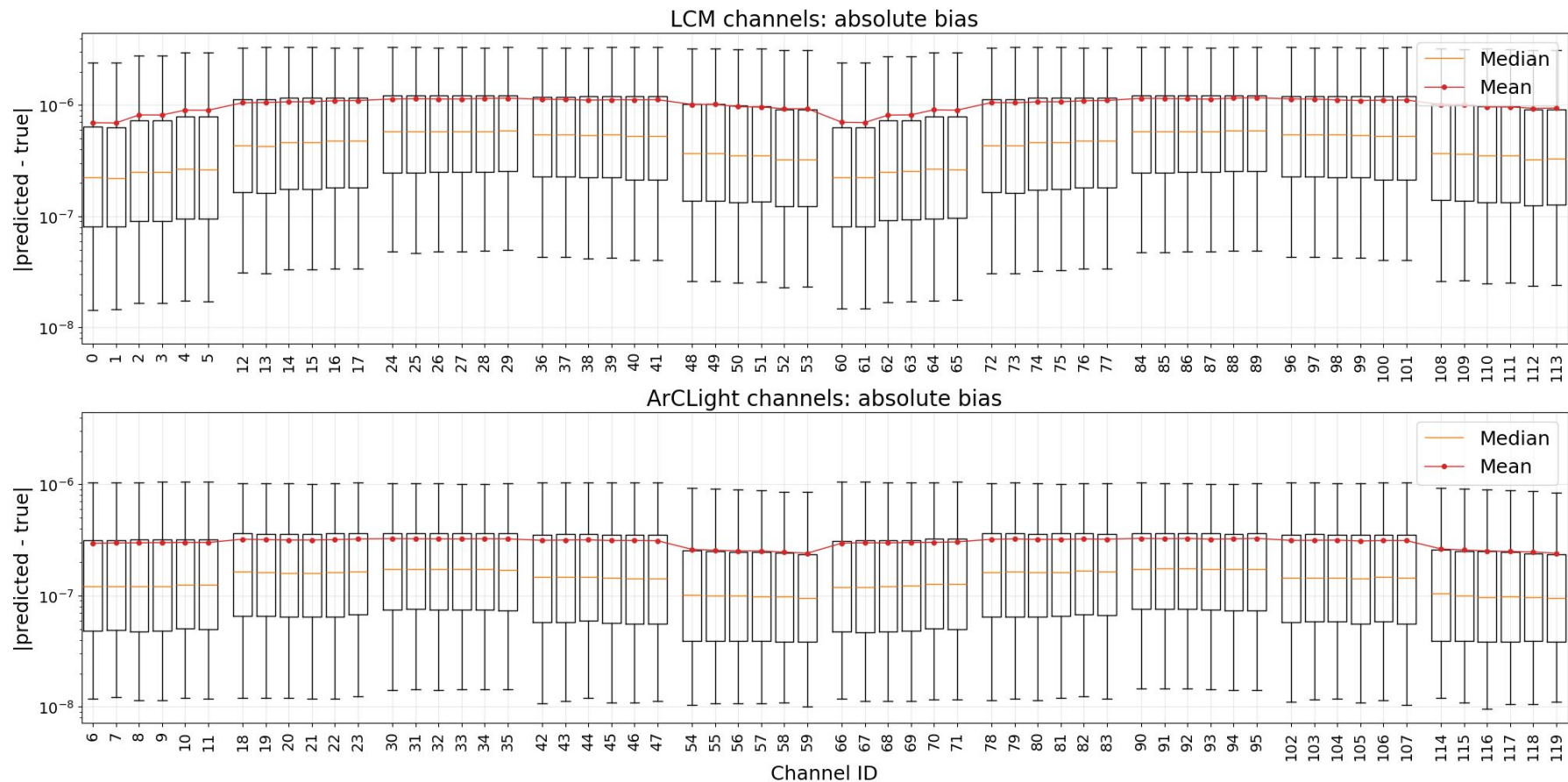
Relative bias (per channel & voxel)

$$b_{\text{rel}} = \frac{|v_{\text{pred}} - v_{\text{true}}|}{v_{\text{true}}}$$

Icarus paper (over all channels & voxels)

$$b_{\text{abs}} = \langle |\Psi_j(\mathbf{x}_i) - v_{ij}| \rangle,$$
$$b_{\text{rel}} = \left\langle 2 \left| \frac{\Psi_j(\mathbf{x}_i) - v_{ij}}{\Psi_j(\mathbf{x}_i) + v_{ij}} \right| \right\rangle$$

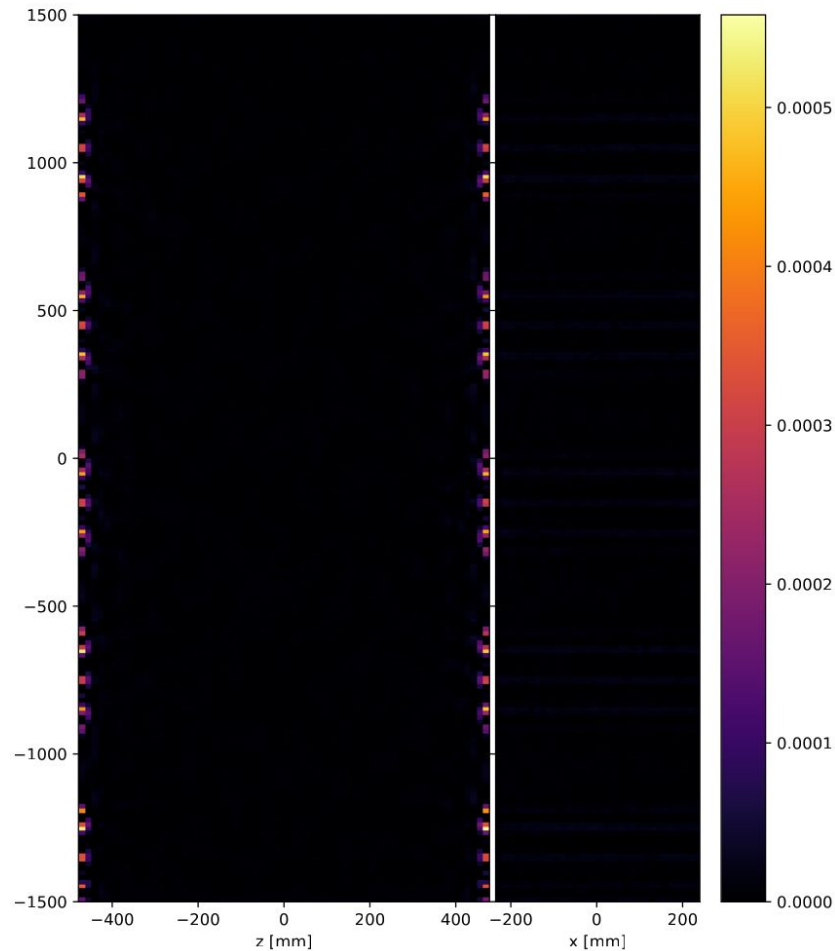
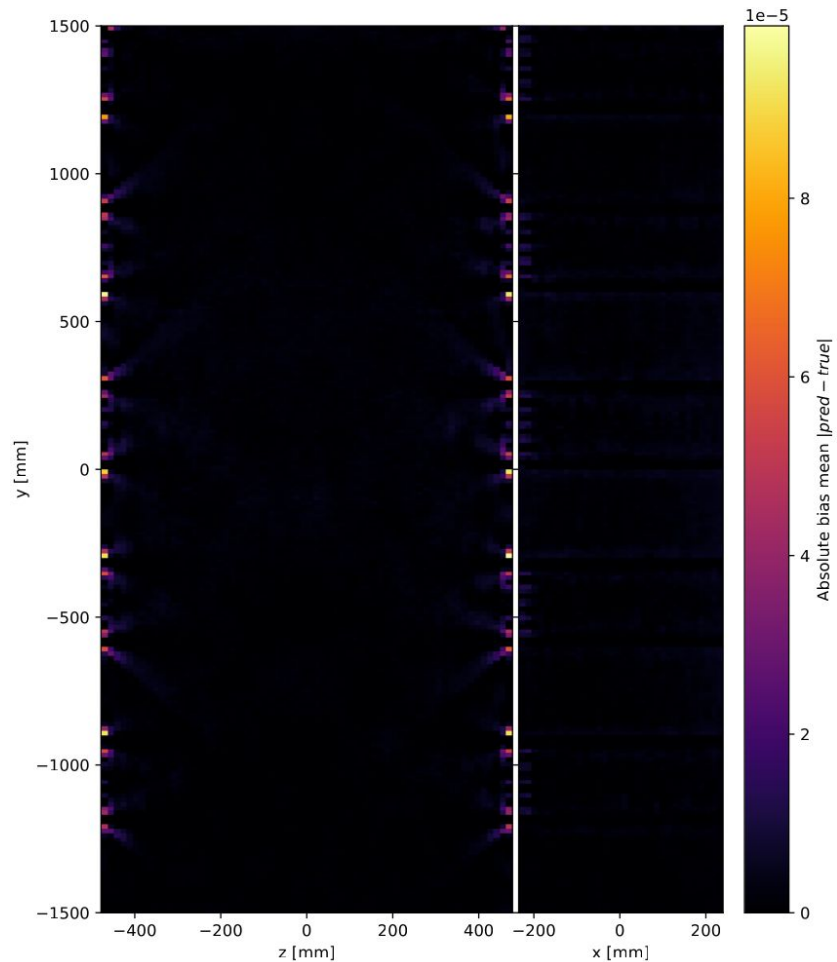
Absolute bias distribution for each channel



ACL

$$P_{zy}(z, y) = \frac{1}{N_x} \sum_{x=1}^{N_x} \left[\frac{1}{N_{ch}} \sum_{ch=1}^{N_{ch}} |\text{pred} - \text{true}| \right]$$

LCM



Model metrics

Absolute bias

$$b_{\text{abs}} = |v_{\text{pred}} - v_{\text{true}}|$$

Relative bias

$$b_{\text{rel}} = \frac{|v_{\text{pred}} - v_{\text{true}}|}{v_{\text{true}}}$$

- Can explode for small true visibilities
- Smallest visibility is 1 photon hit -> $v_{\text{min}} = 10^{-7}$
- Assuming binomial uncertainty and small probability

$$\sigma_p/p \sim (p(1-p)/\text{sqrt}(N))/p \sim 1/\text{sqrt}(N)$$

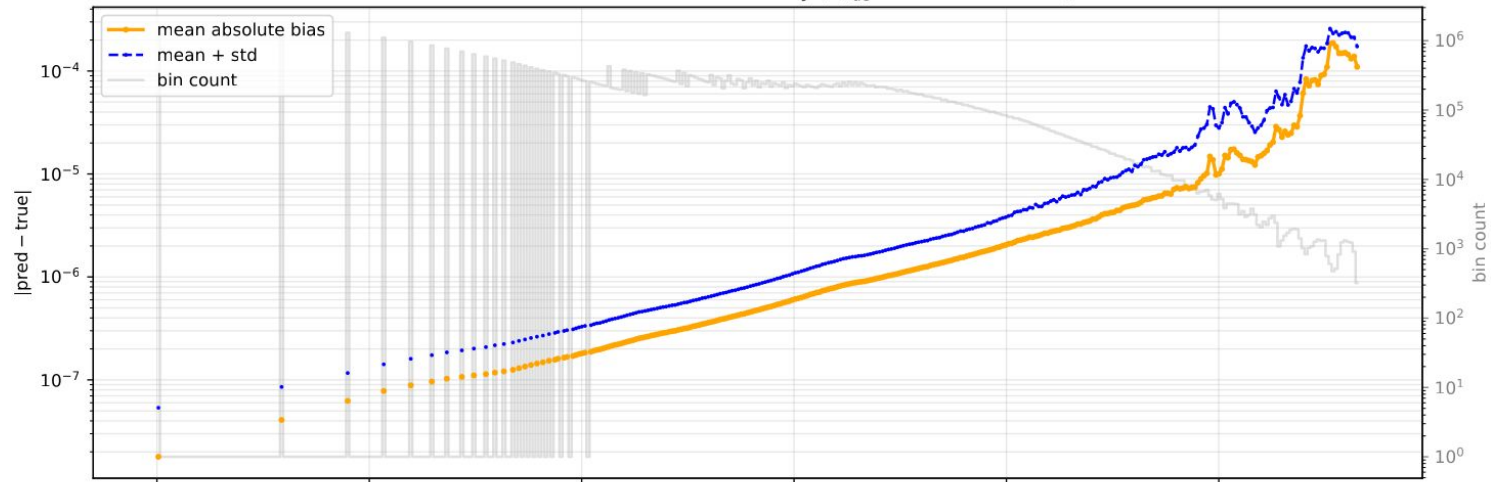
-> Uncertainties on visibilities

$$\sigma_{\text{vis}} (v = 10^{-7}) \sim 100 \%$$

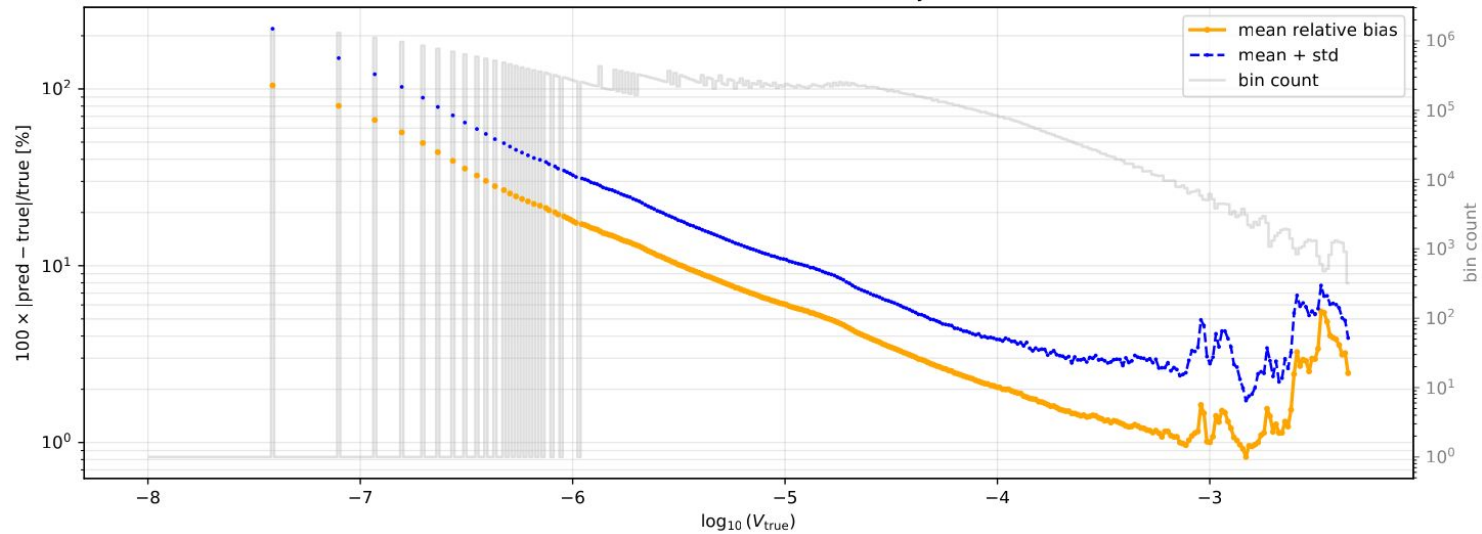
$$\sigma_{\text{vis}} (v = 10^{-6}) \sim 32 \%$$

$$\sigma_{\text{vis}} (v = 10^{-5}) \sim 10 \%$$

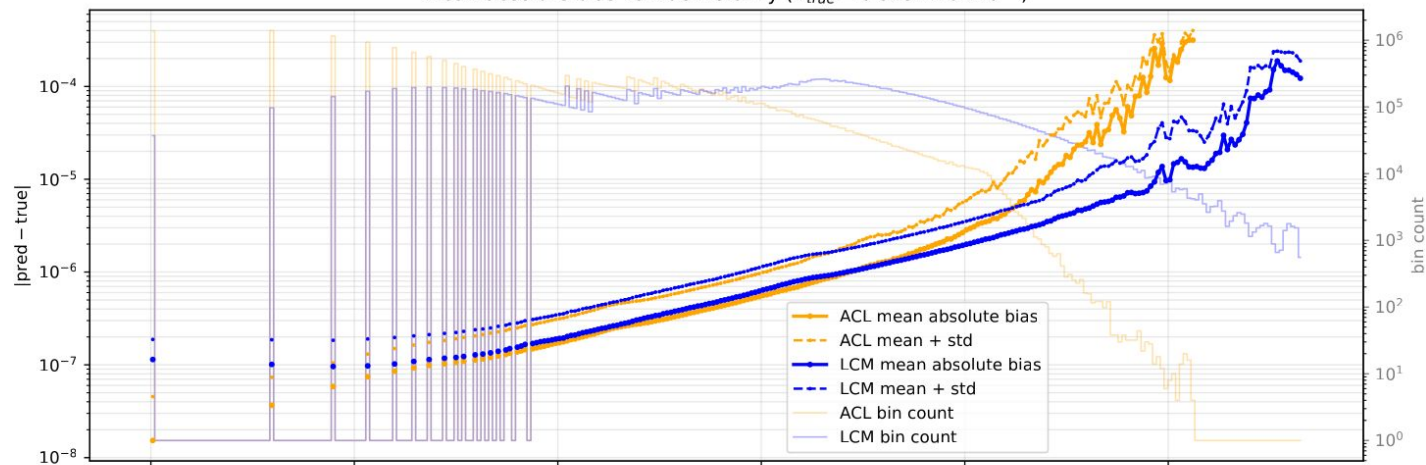
Mean absolute bias vs true visibility ($V_{\text{true}} = 0$ shown at 10^{-8})



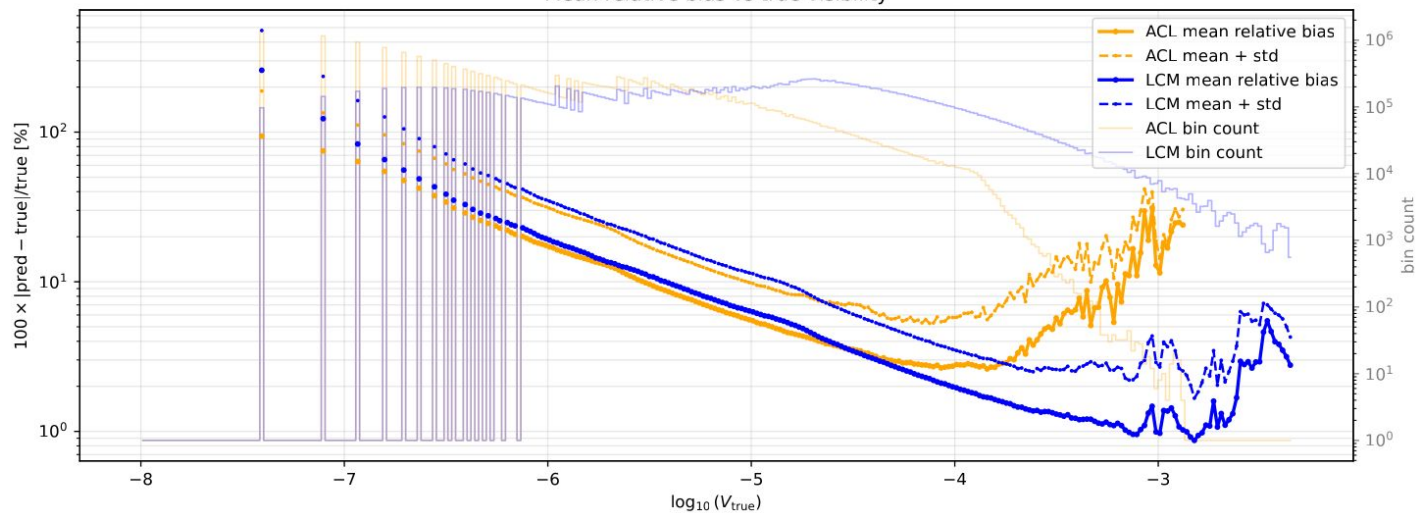
Mean relative bias vs true visibility



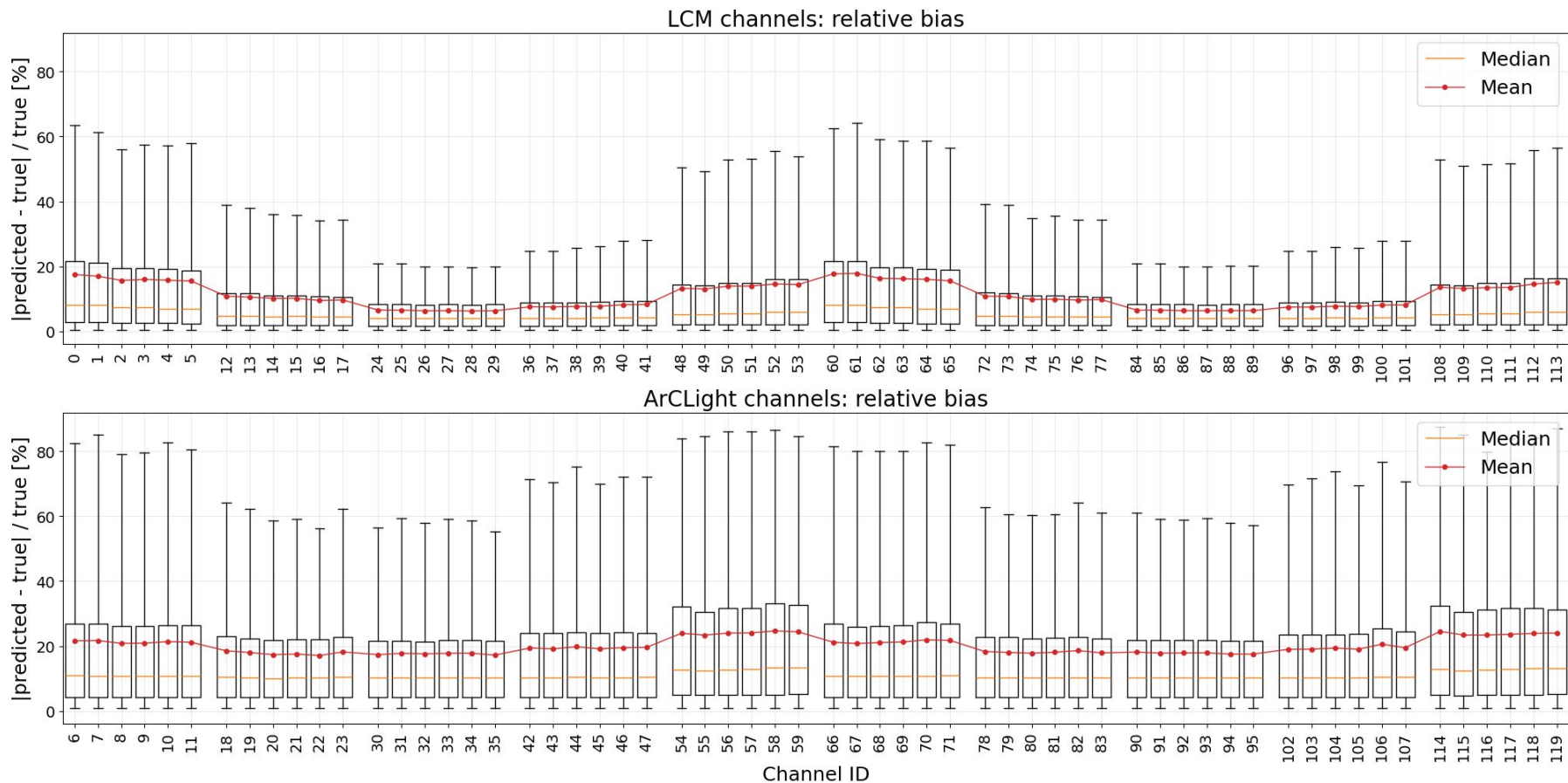
Binned bias vs true visibility: ACL vs LCM (300 bins)
 Mean absolute bias vs true visibility ($V_{true} = 0$ shown at 10^{-8})



Mean relative bias vs true visibility



Relative bias: true vis > 1e-7

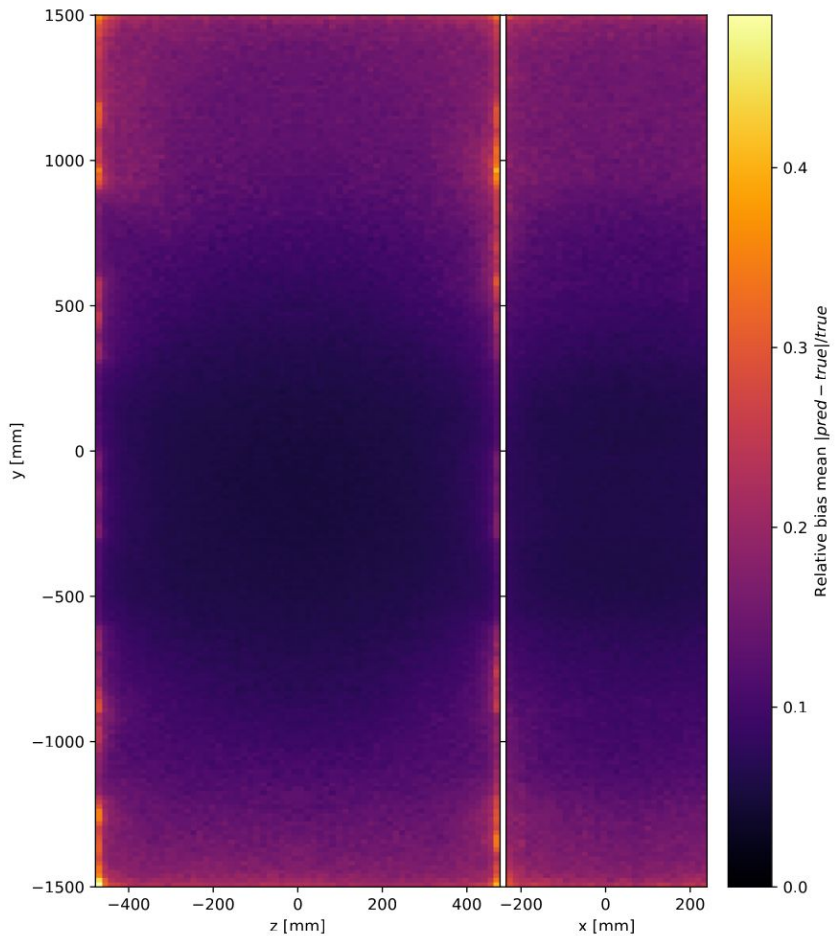
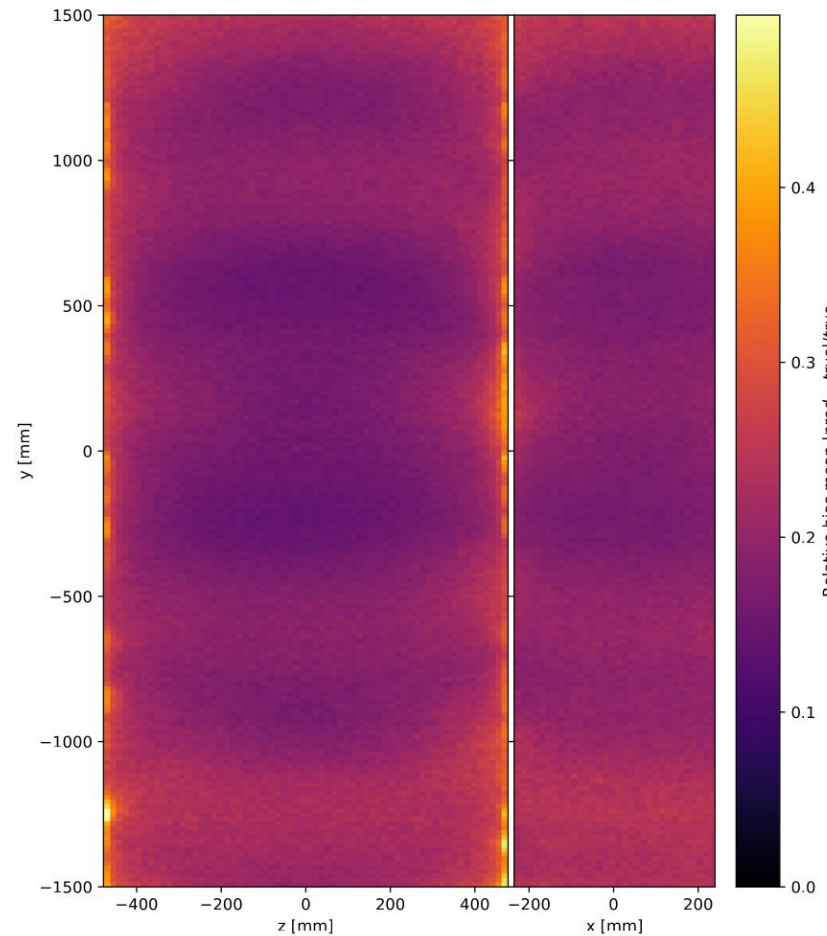


true vis > 1e-7

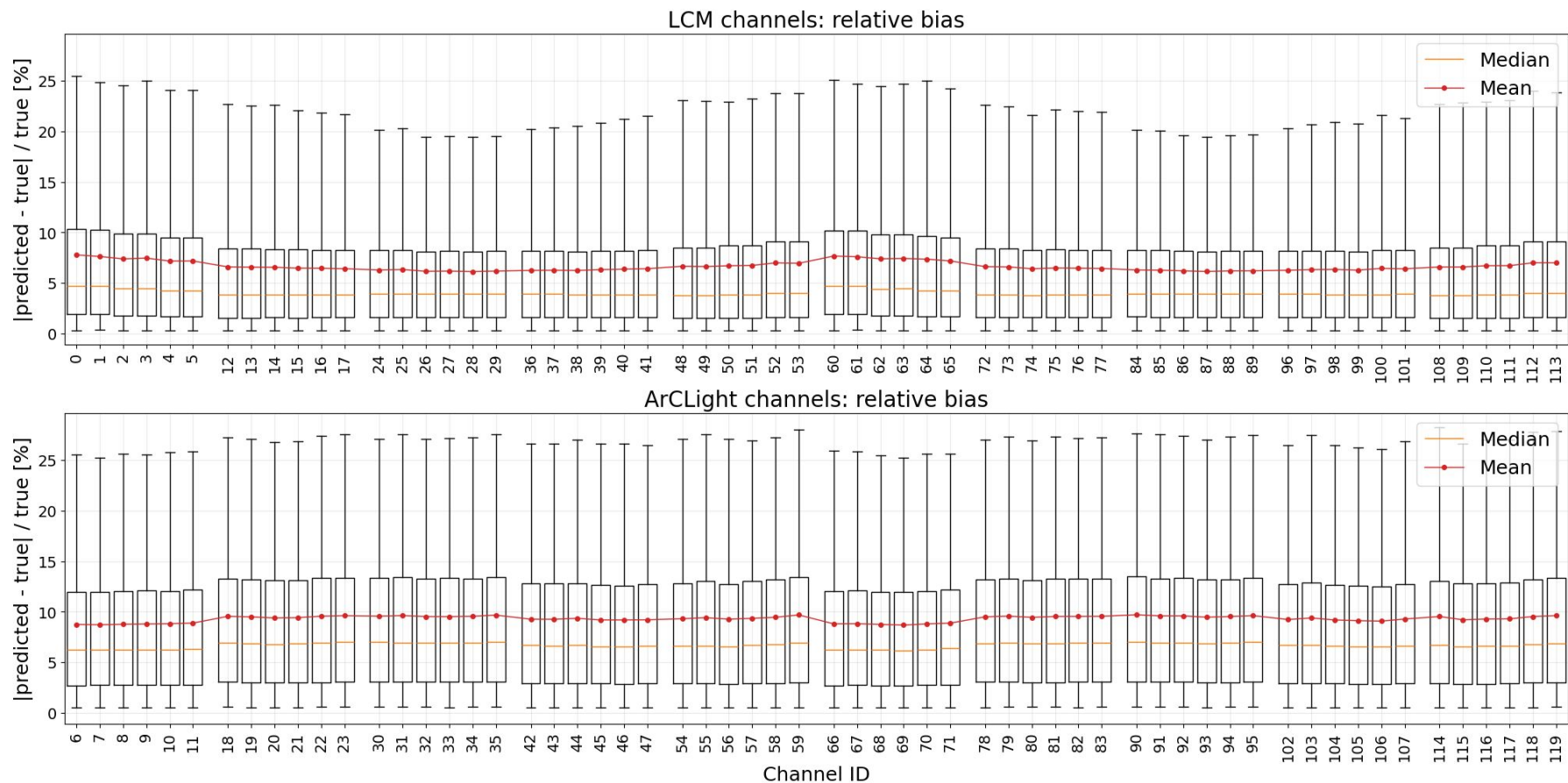
ACL

$$P_{zy}(z, y) = \frac{1}{N_x} \sum_{x=1}^{N_x} \left[\frac{1}{N_{ch}} \sum_{ch=1}^{N_{ch}} \frac{|\text{pred} - \text{true}|}{\text{true}} \right]$$

LCM



Relative bias: true vis > 1e-6

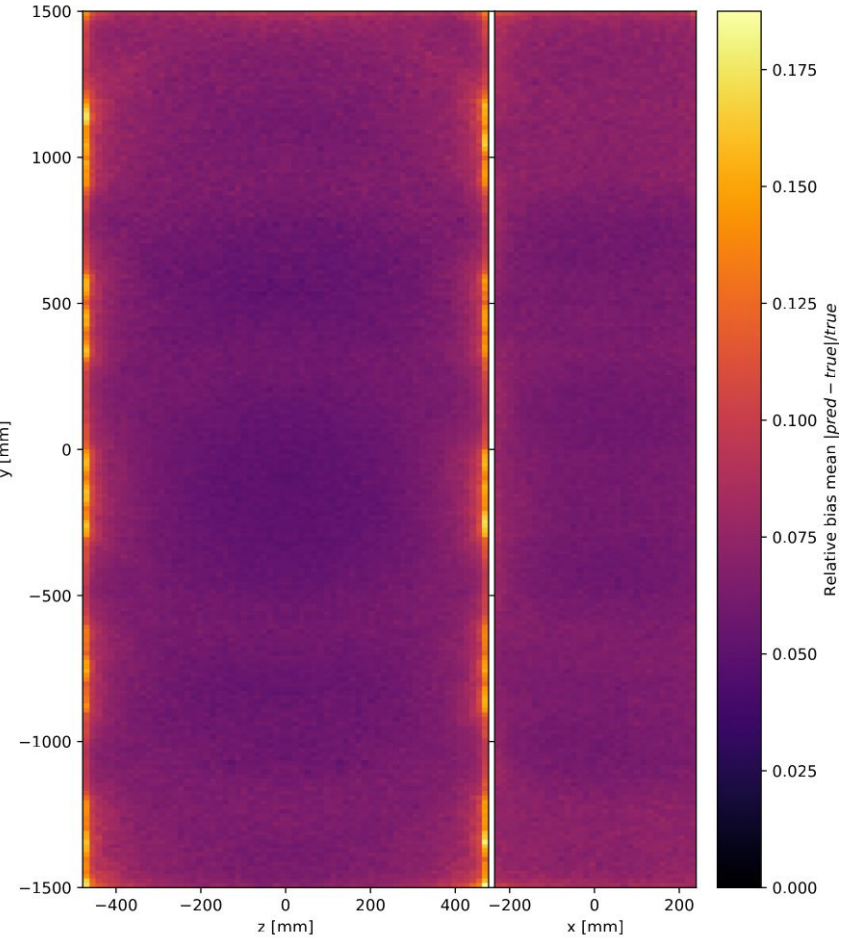
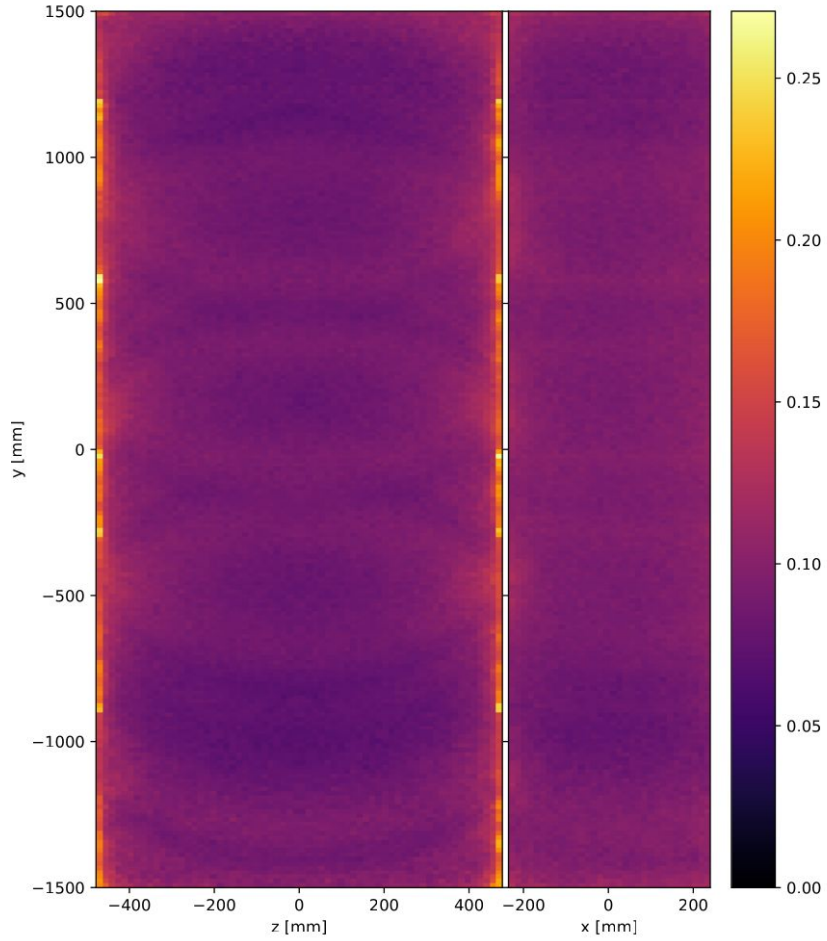


true vis > 1e-6

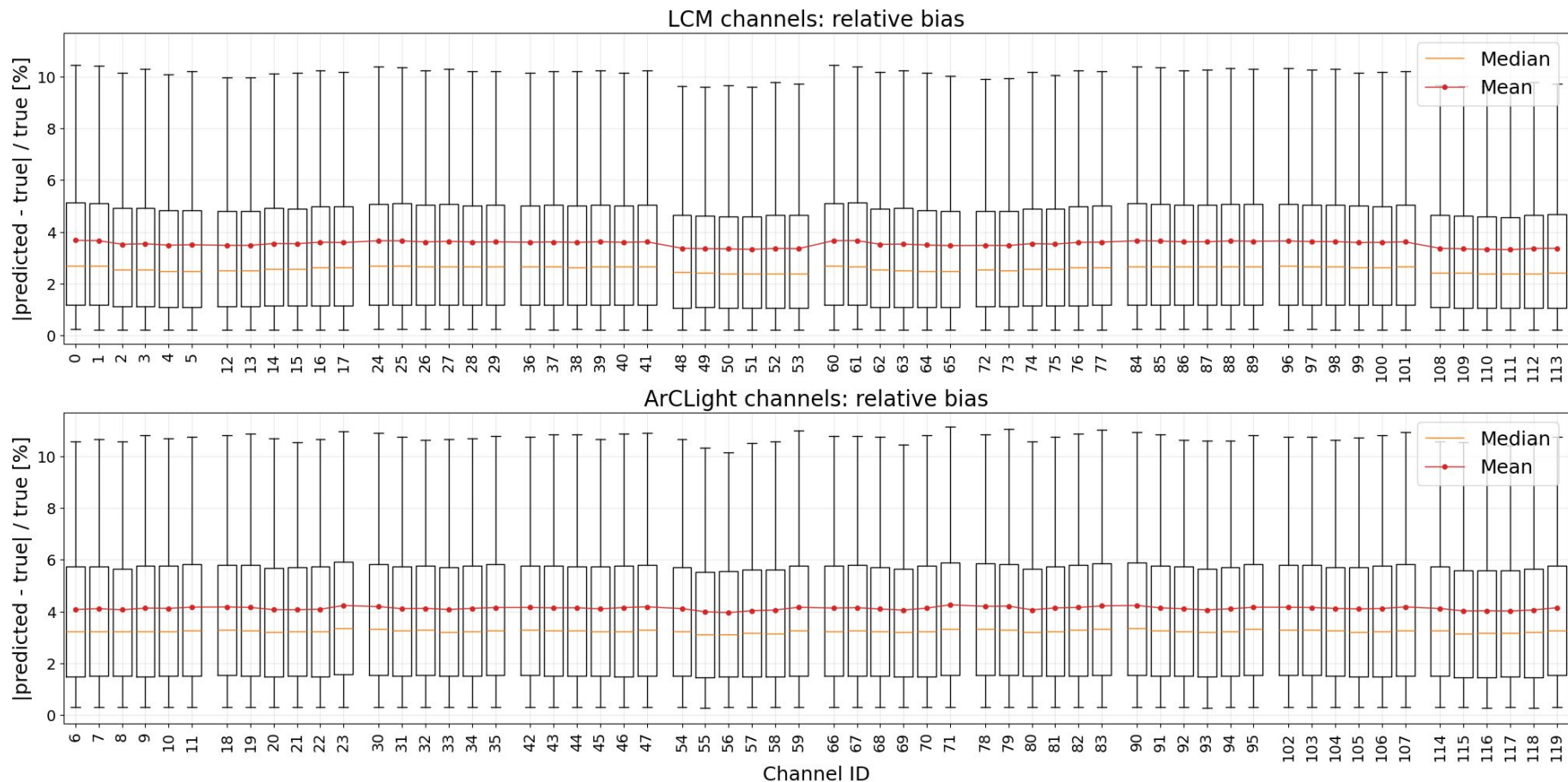
ACL

$$P_{zy}(z, y) = \frac{1}{N_x} \sum_{x=1}^{N_x} \left[\frac{1}{N_{ch}} \sum_{ch=1}^{N_{ch}} \frac{|\text{pred} - \text{true}|}{\text{true}} \right]$$

LCM



Relative bias: true vis > 1e-5

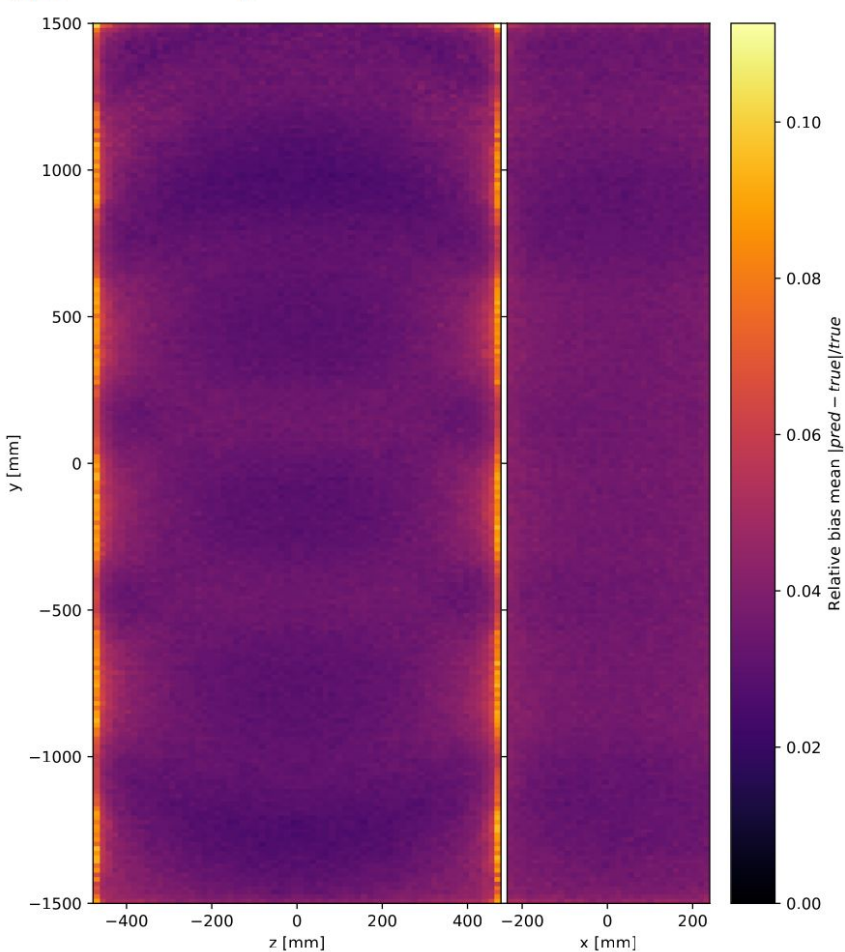
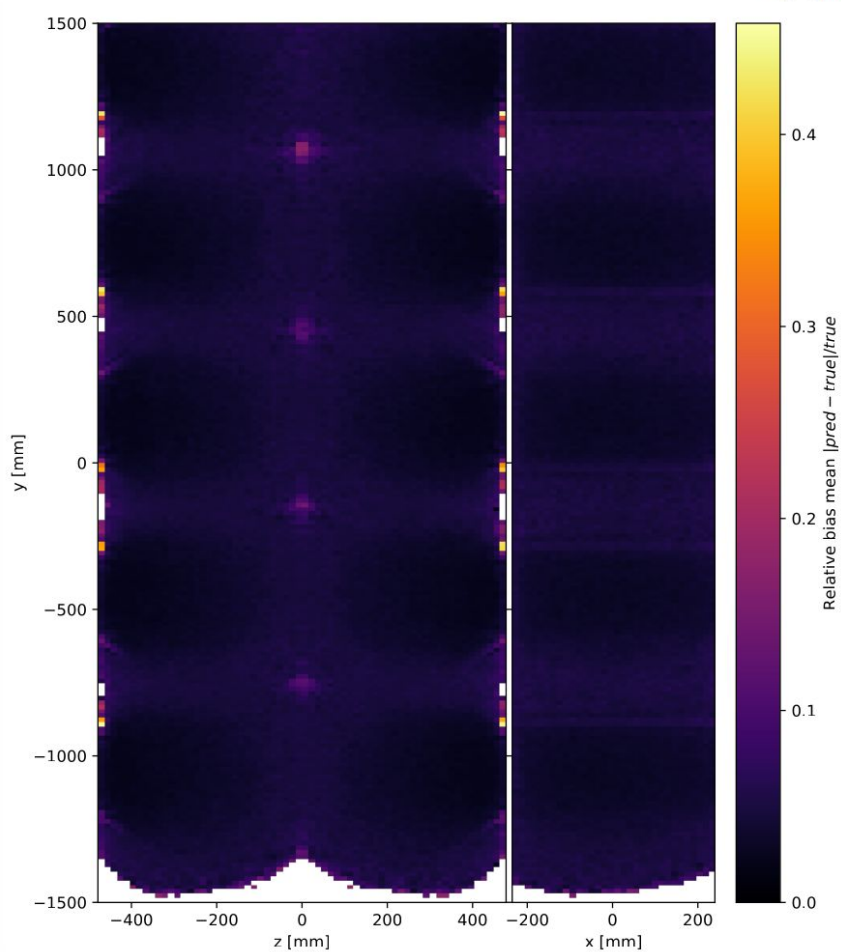


true vis > 1e-5

ACL

$$P_{zy}(z, y) = \frac{1}{N_x} \sum_{x=1}^{N_x} \left[\frac{1}{N_{\text{ch}}} \sum_{\text{ch}=1}^{N_{\text{ch}}} \frac{|\text{pred} - \text{true}|}{\text{true}} \right]$$

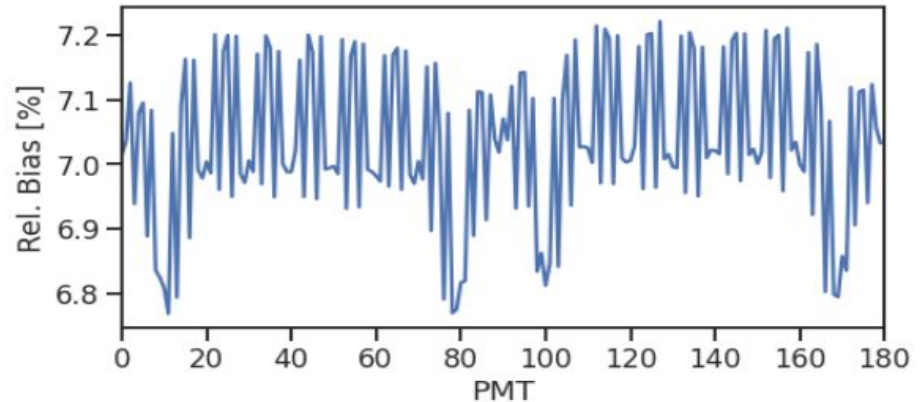
LCM



Comparison mean relative bias vs thresholds

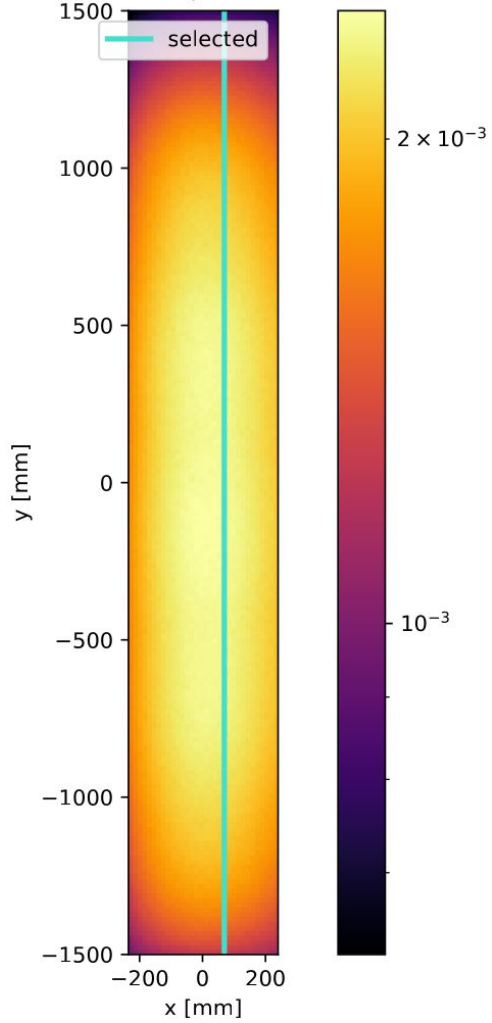
Threshold for true vis	Uncertainty of vis_thr	Mean LCM channels [%]	Mean ACL channels [%]
1e-7	$\sigma_{\text{vis}} (v = 10^{-7}) \sim 100 \%$	10.83	19.73
1e-6	$\sigma_{\text{vis}} (v = 10^{-6}) \sim 32 \%$	6.59	9.33
1e-5	$\sigma_{\text{vis}} (v = 10^{-5}) \sim 10 \%$	3.54	4.12

Relative bias plot per channel (Icarus paper): <https://arxiv.org/pdf/2211.01505>

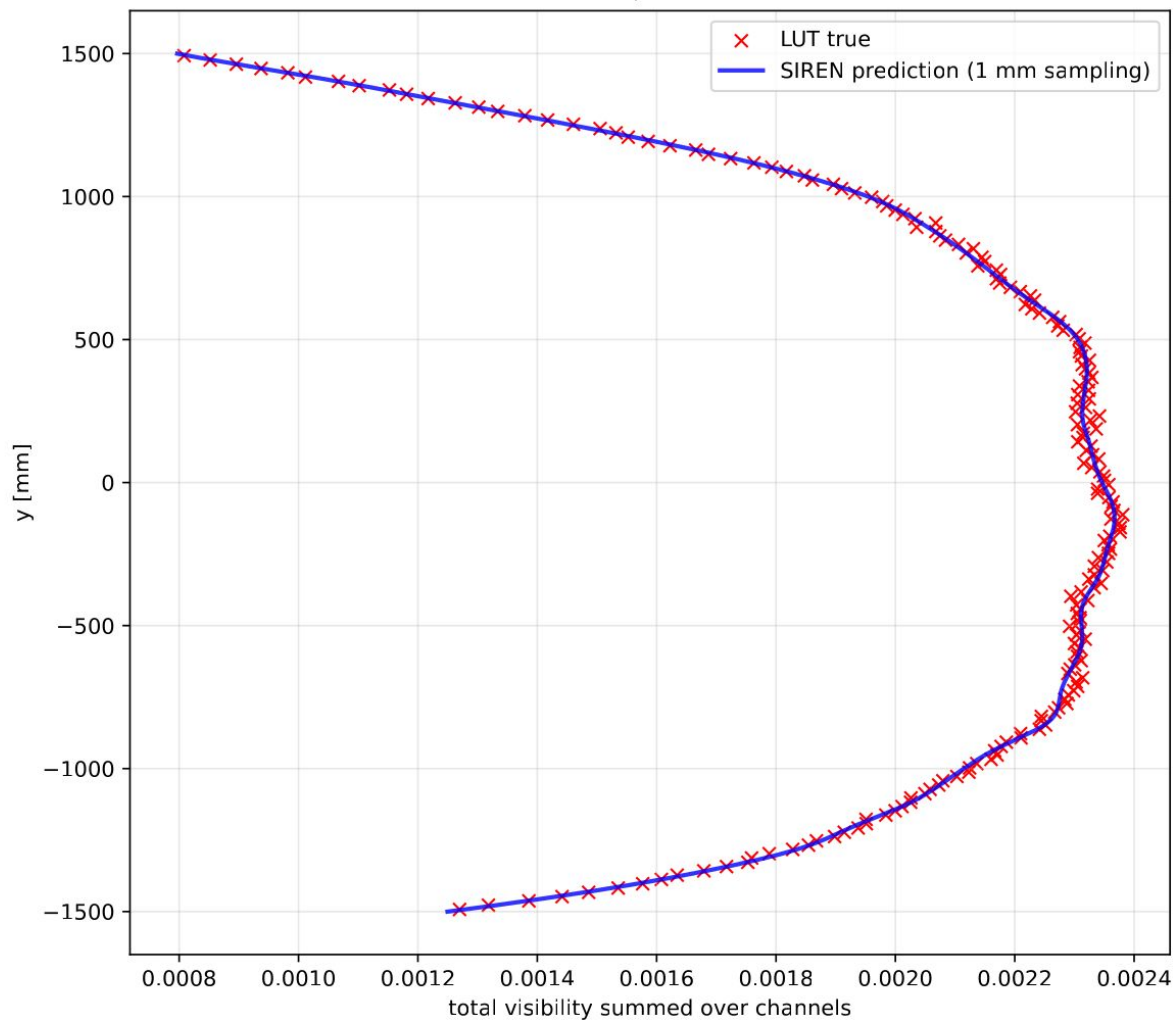


* Icarus paper choose a true vis threshold of 4.5×10^{-5} , but the visibility range was larger

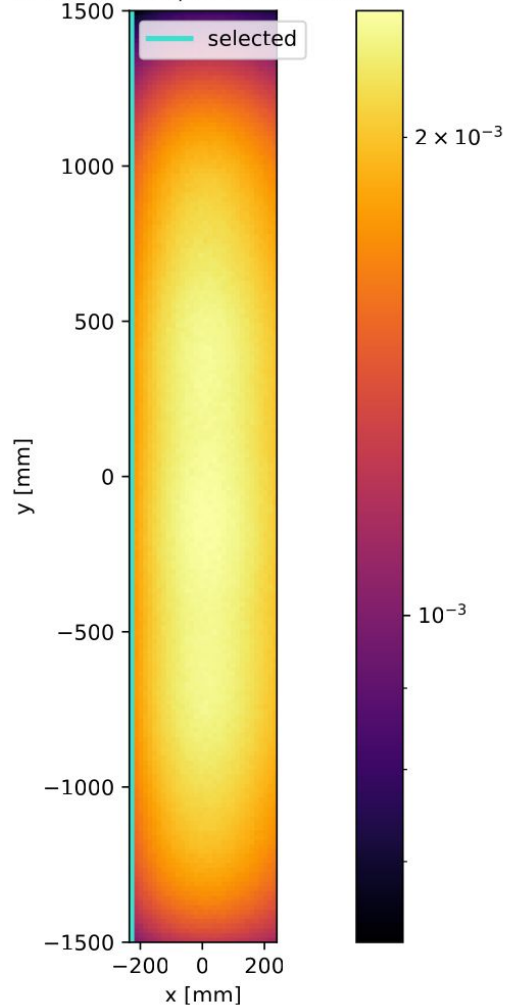
True total visibility at
 $z = 7.48$ mm, $x = 69.17$ mm



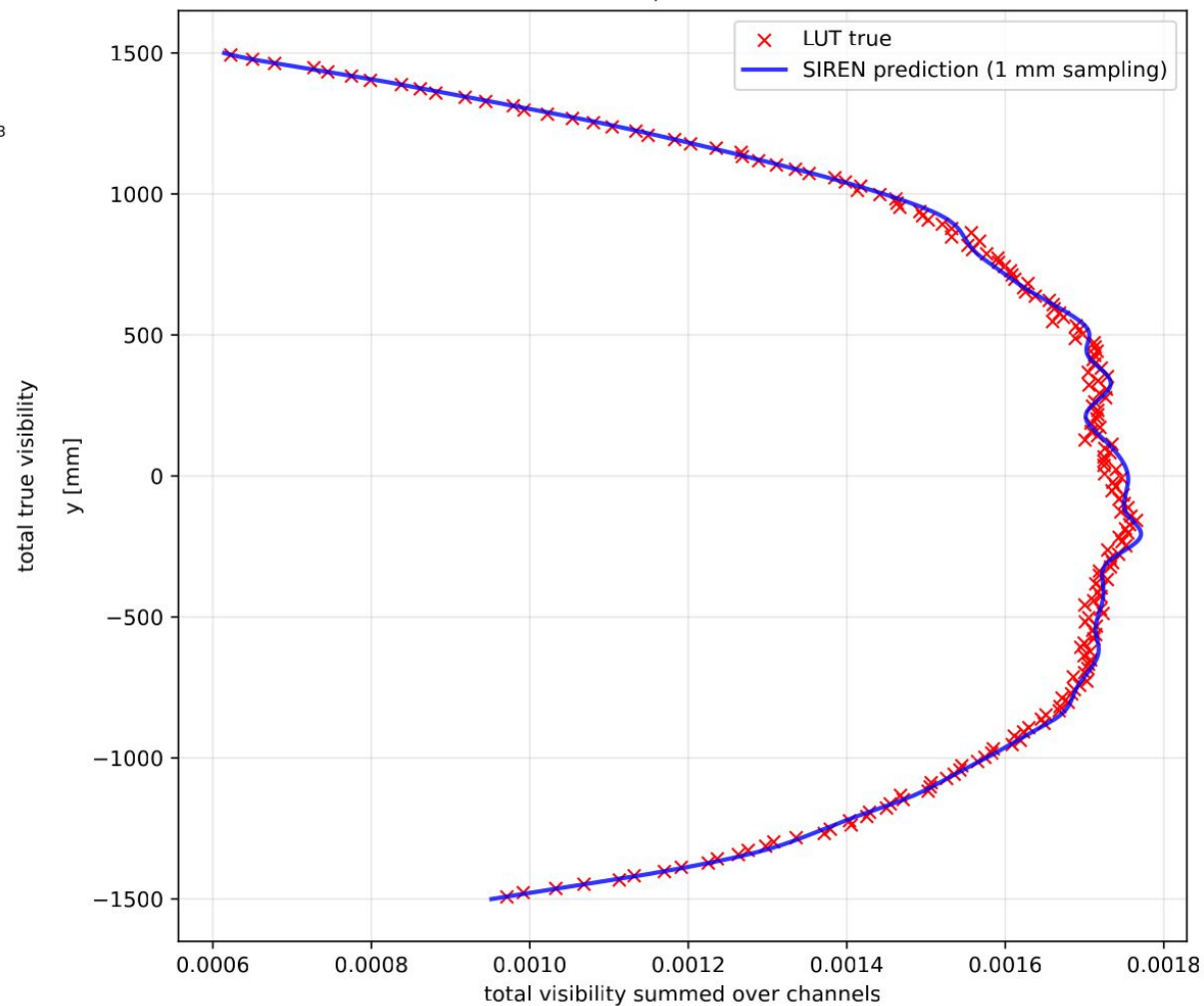
Selected voxel line
 $z = 7.48$ mm, $x = 69.17$ mm



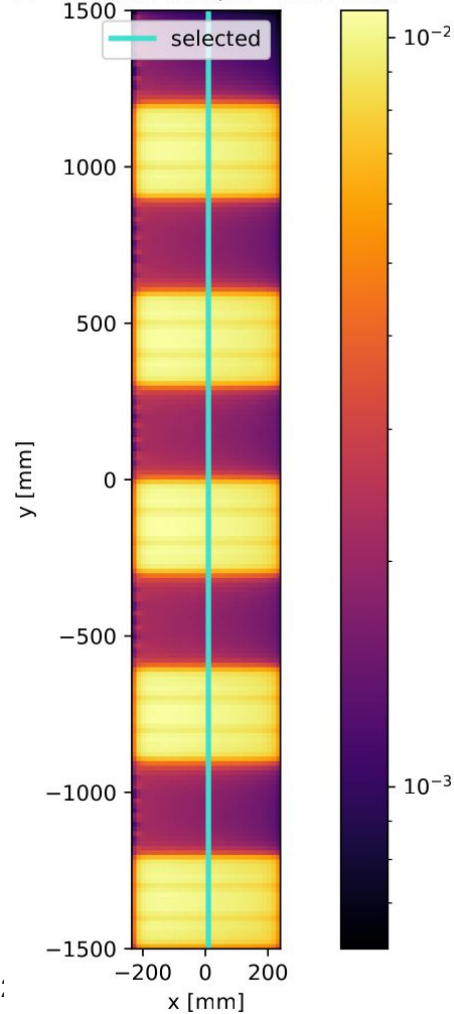
True total visibility at
 $z = 7.48$ mm, $x = -227.90$ mm



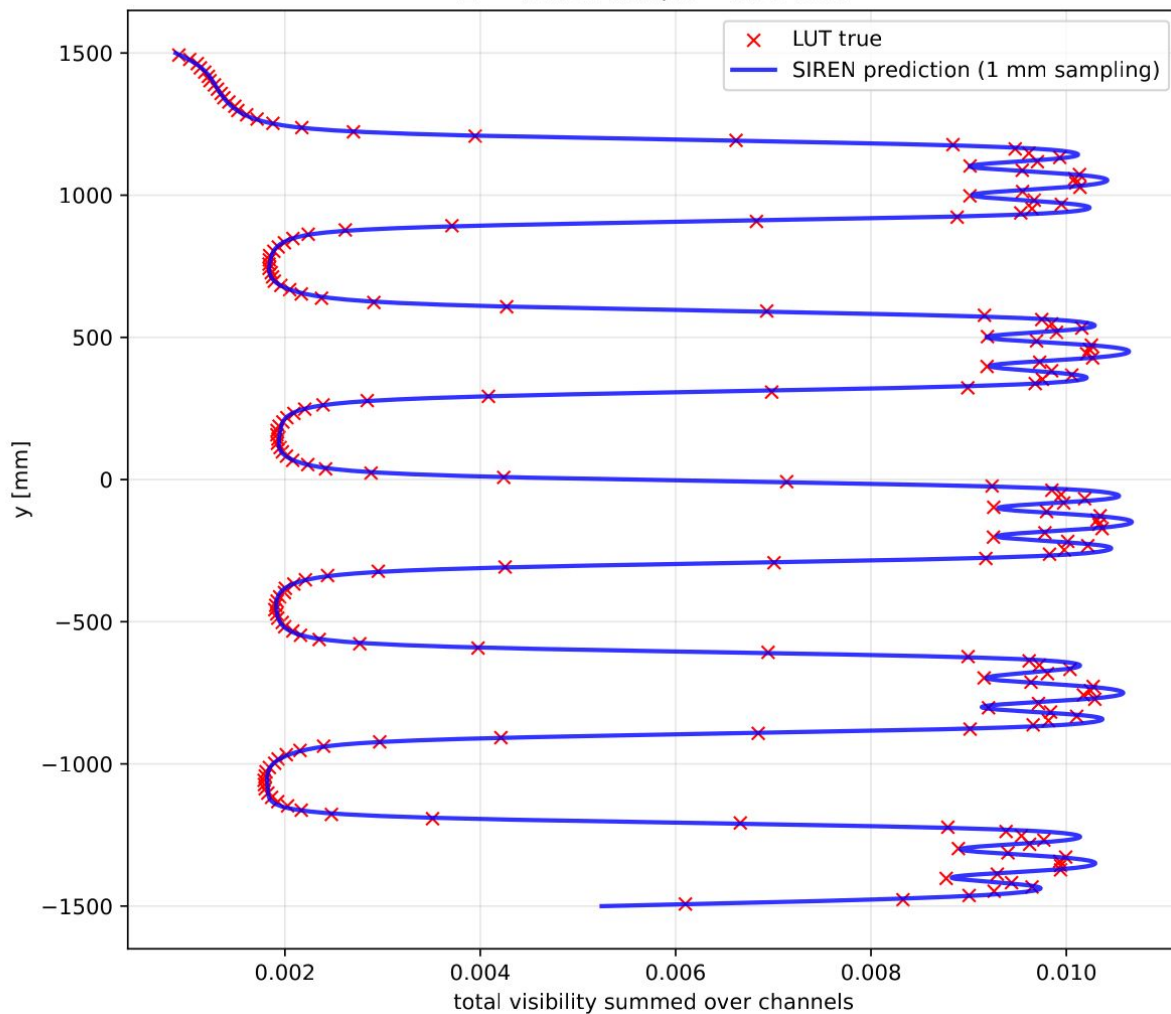
Selected voxel line
 $z = 7.48$ mm, $x = -227.90$ mm



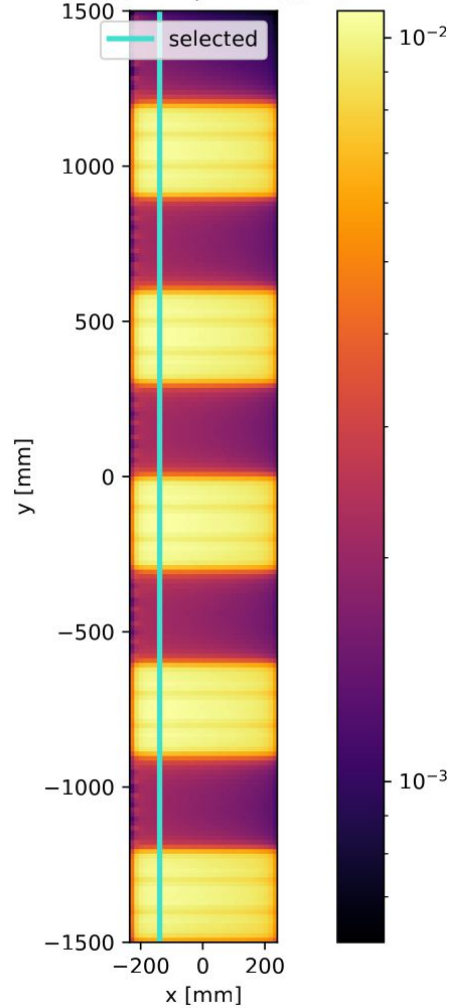
True total visibility at
 $z = 471.09$ mm, $x = 9.75$ mm



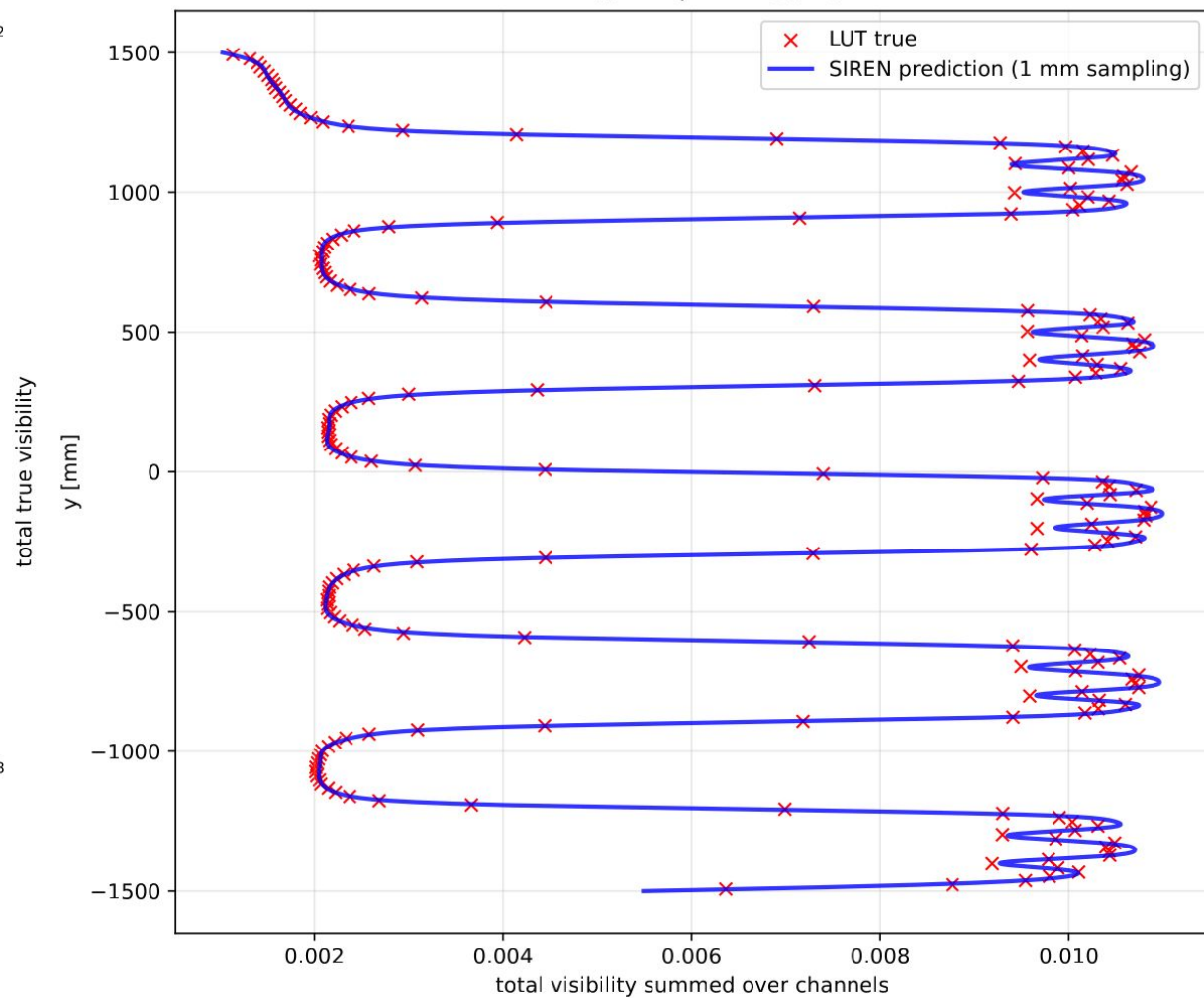
Selected voxel line
 $z = 471.09$ mm, $x = 9.75$ mm



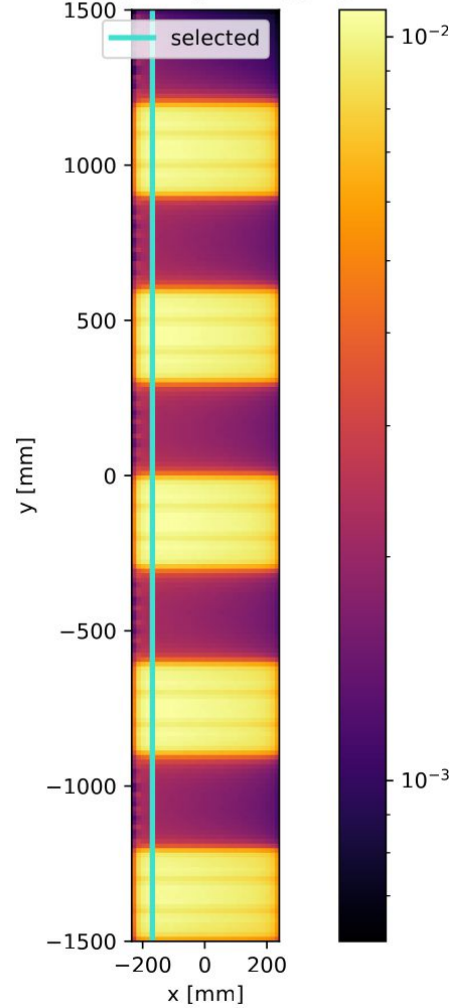
True total visibility at
 $z = 471.09$ mm, $x = -138.78$ mm



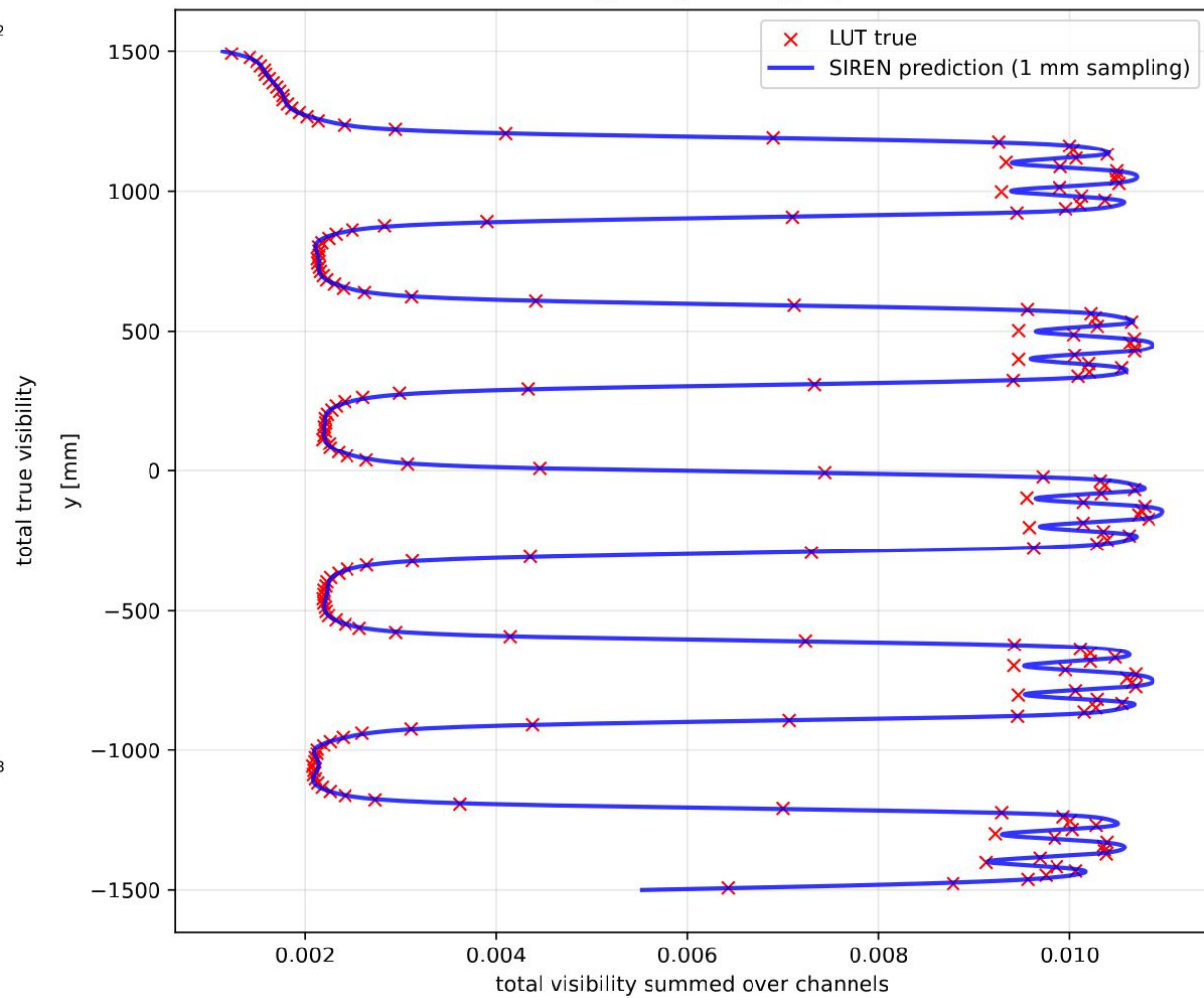
Selected voxel line
 $z = 471.09$ mm, $x = -138.78$ mm



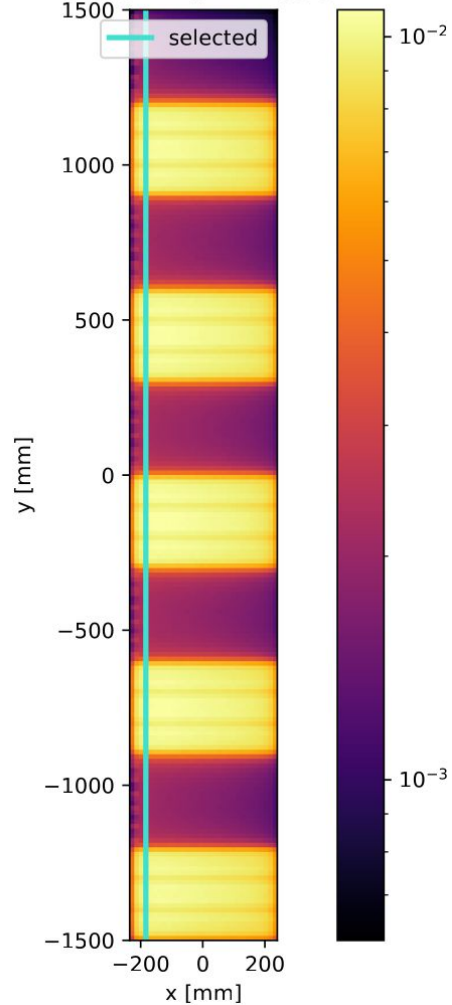
True total visibility at
 $z = 471.09$ mm, $x = -168.48$ mm



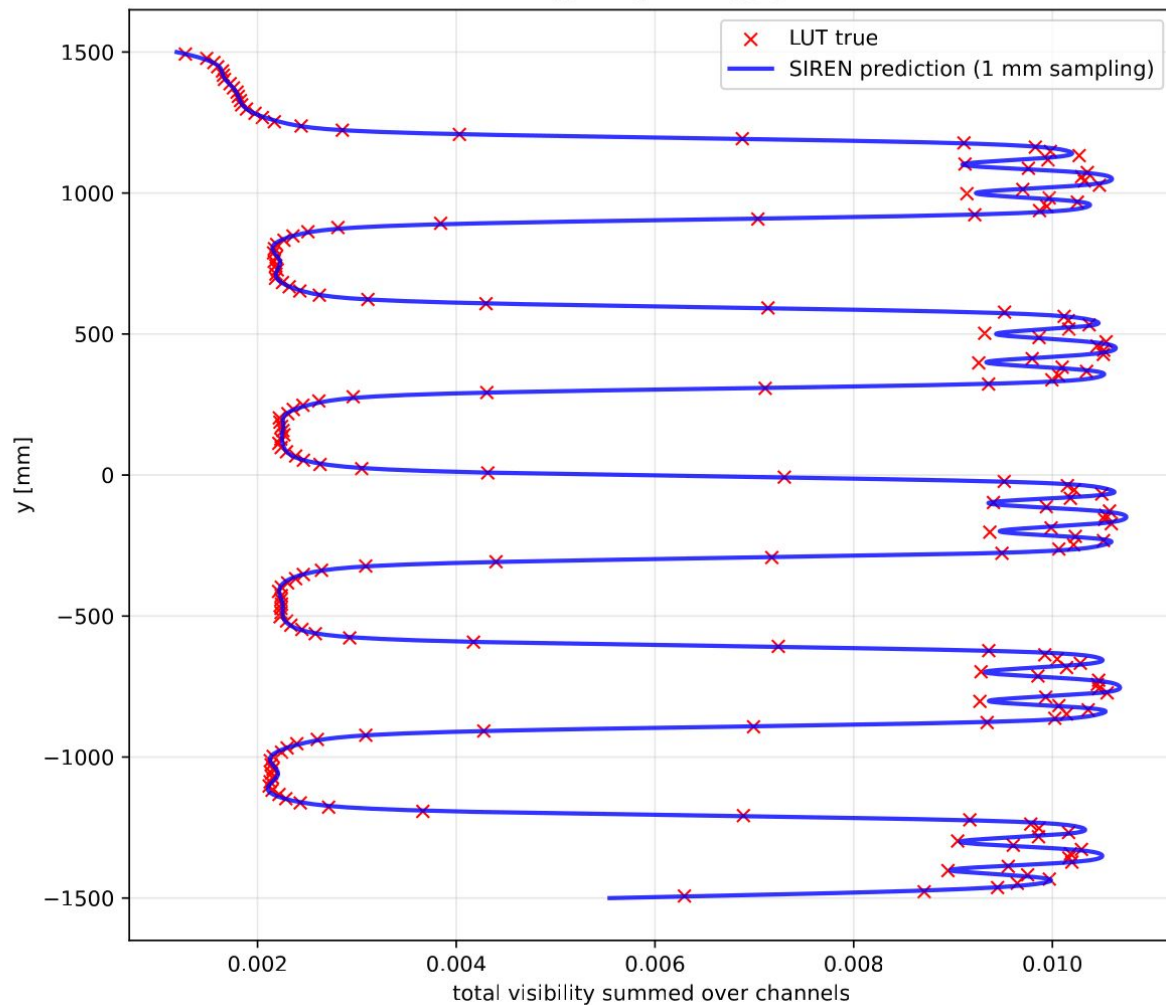
Selected voxel line
 $z = 471.09$ mm, $x = -168.48$ mm



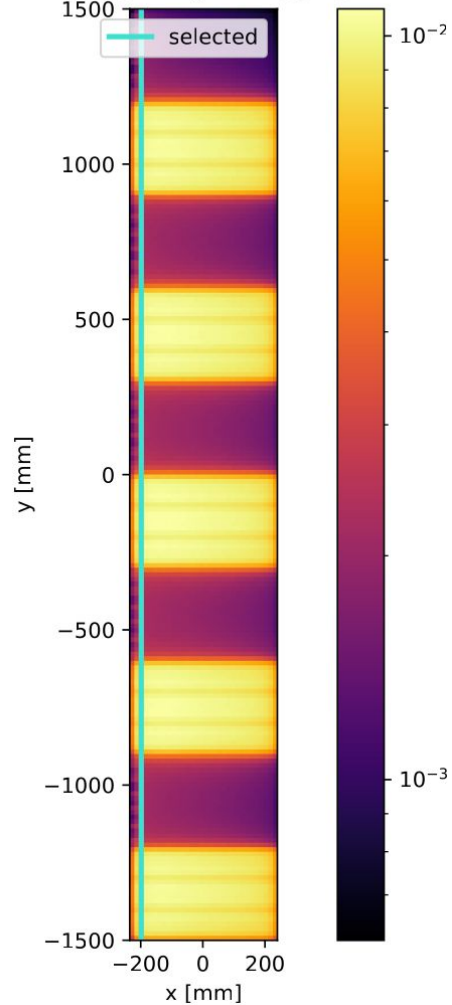
True total visibility at
 $z = 471.09$ mm, $x = -183.34$ mm



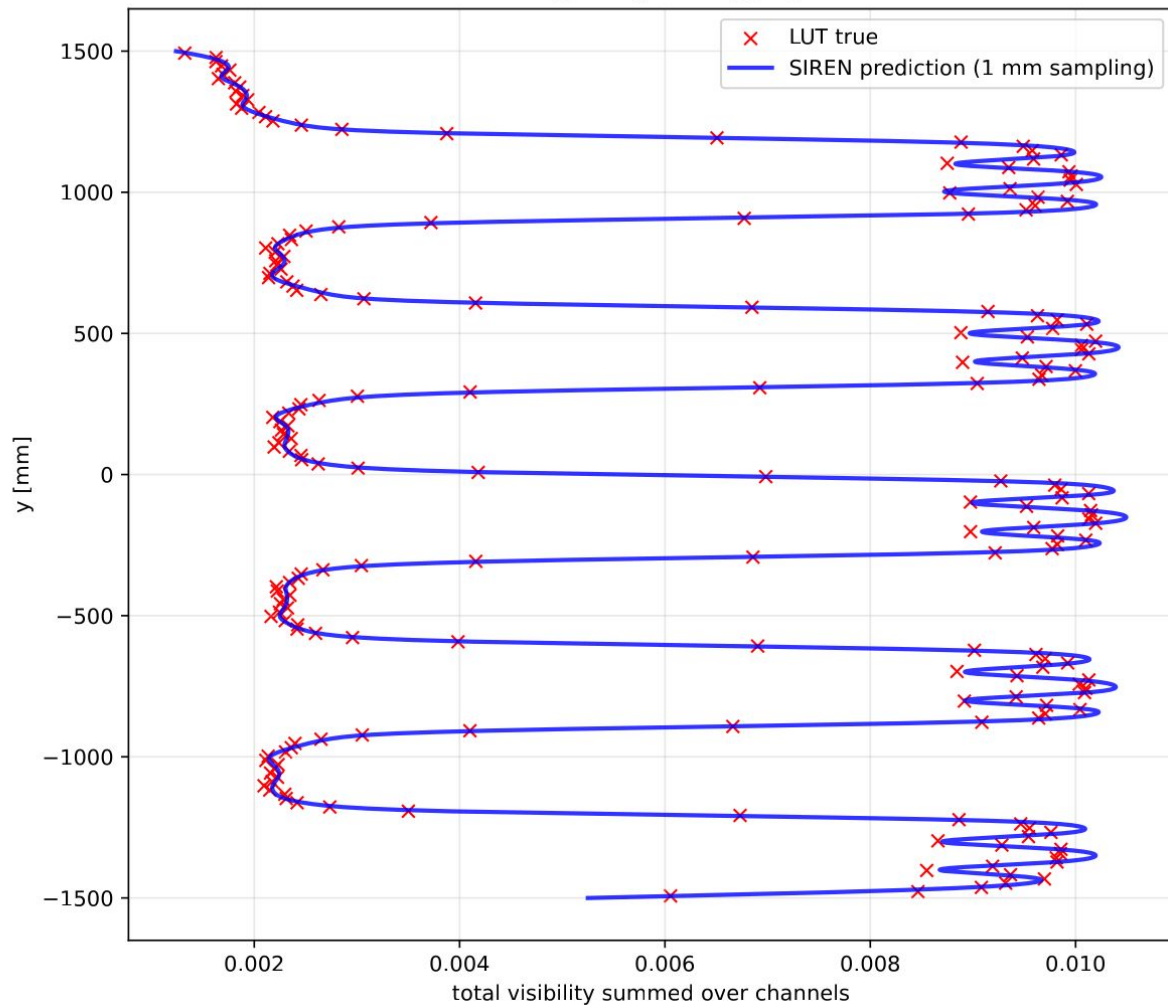
Selected voxel line
 $z = 471.09$ mm, $x = -183.34$ mm



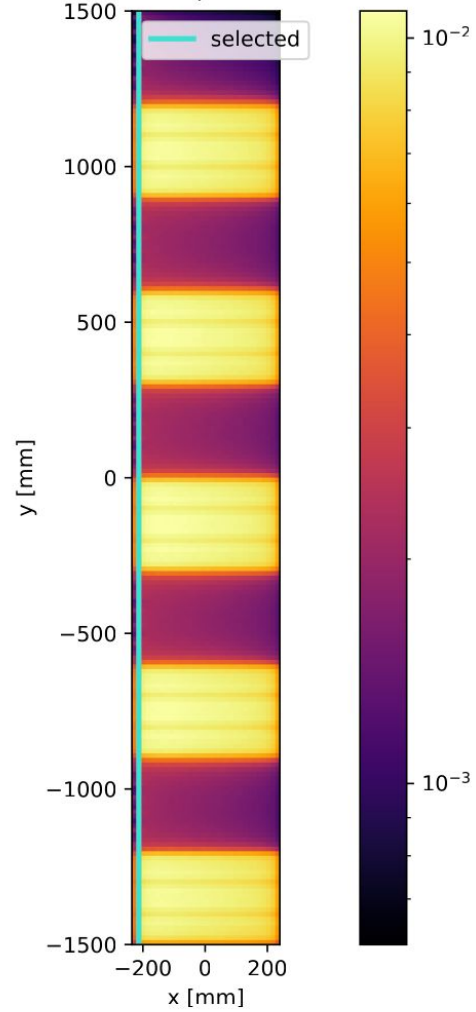
True total visibility at
 $z = 471.09$ mm, $x = -198.19$ mm



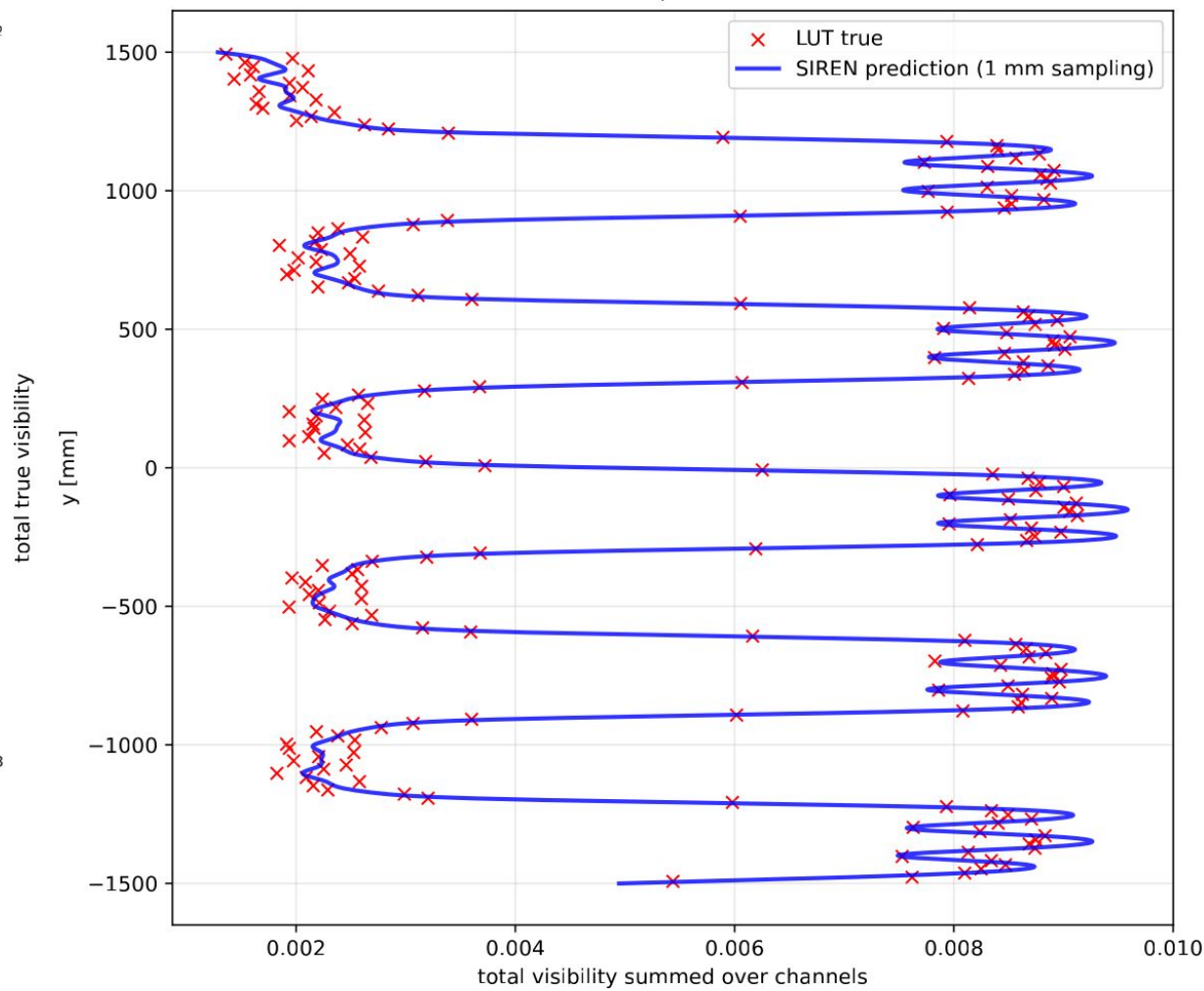
Selected voxel line
 $z = 471.09$ mm, $x = -198.19$ mm

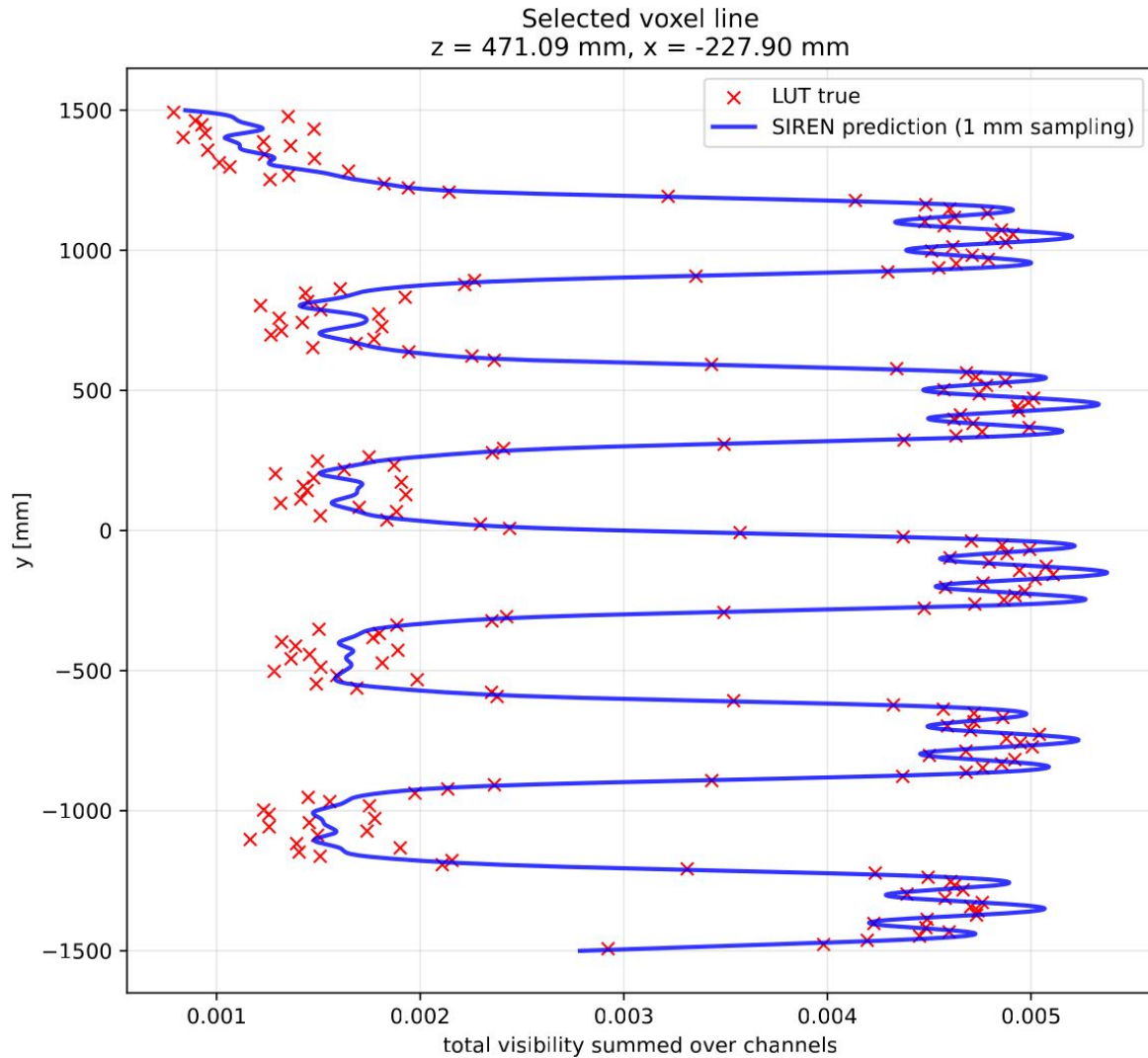
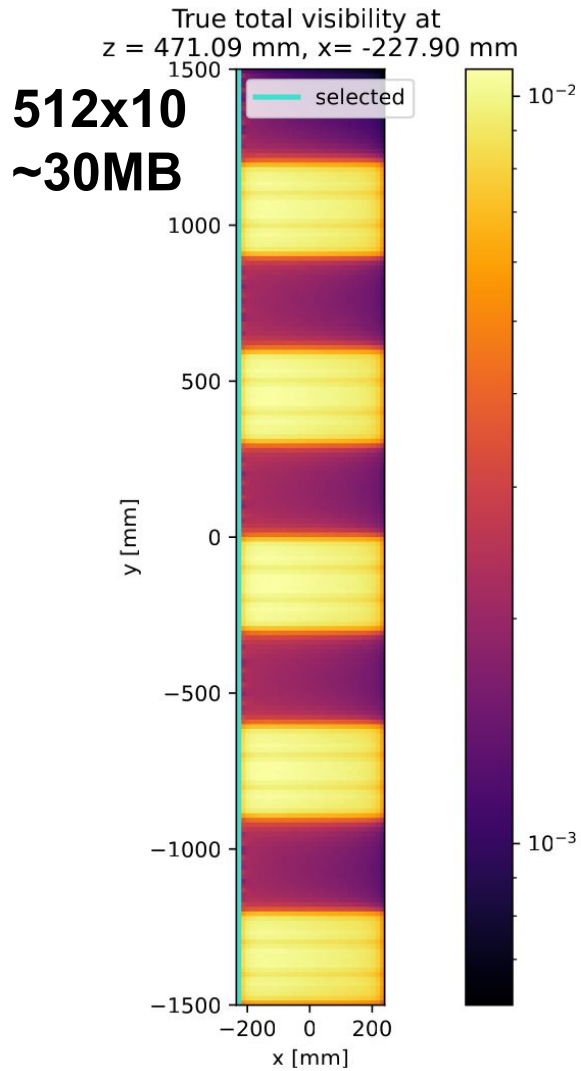


True total visibility at
 $z = 471.09$ mm, $x = -213.04$ mm



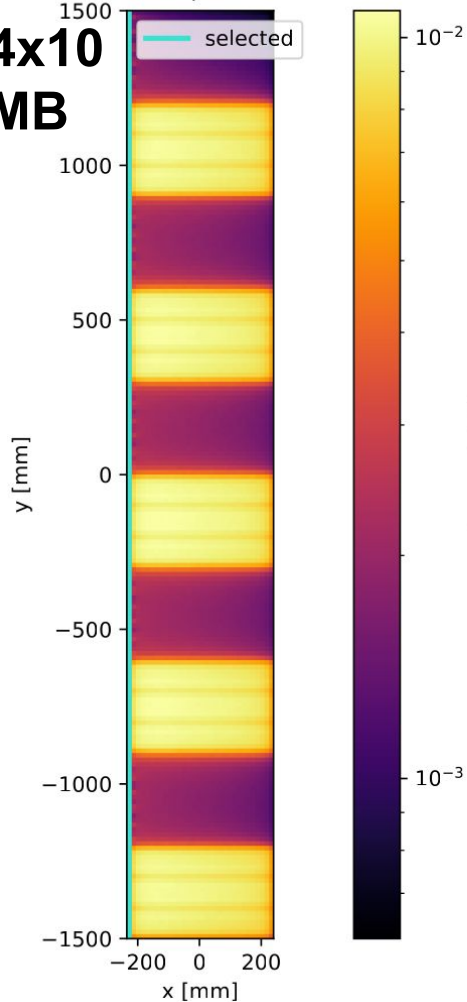
Selected voxel line
 $z = 471.09$ mm, $x = -213.04$ mm



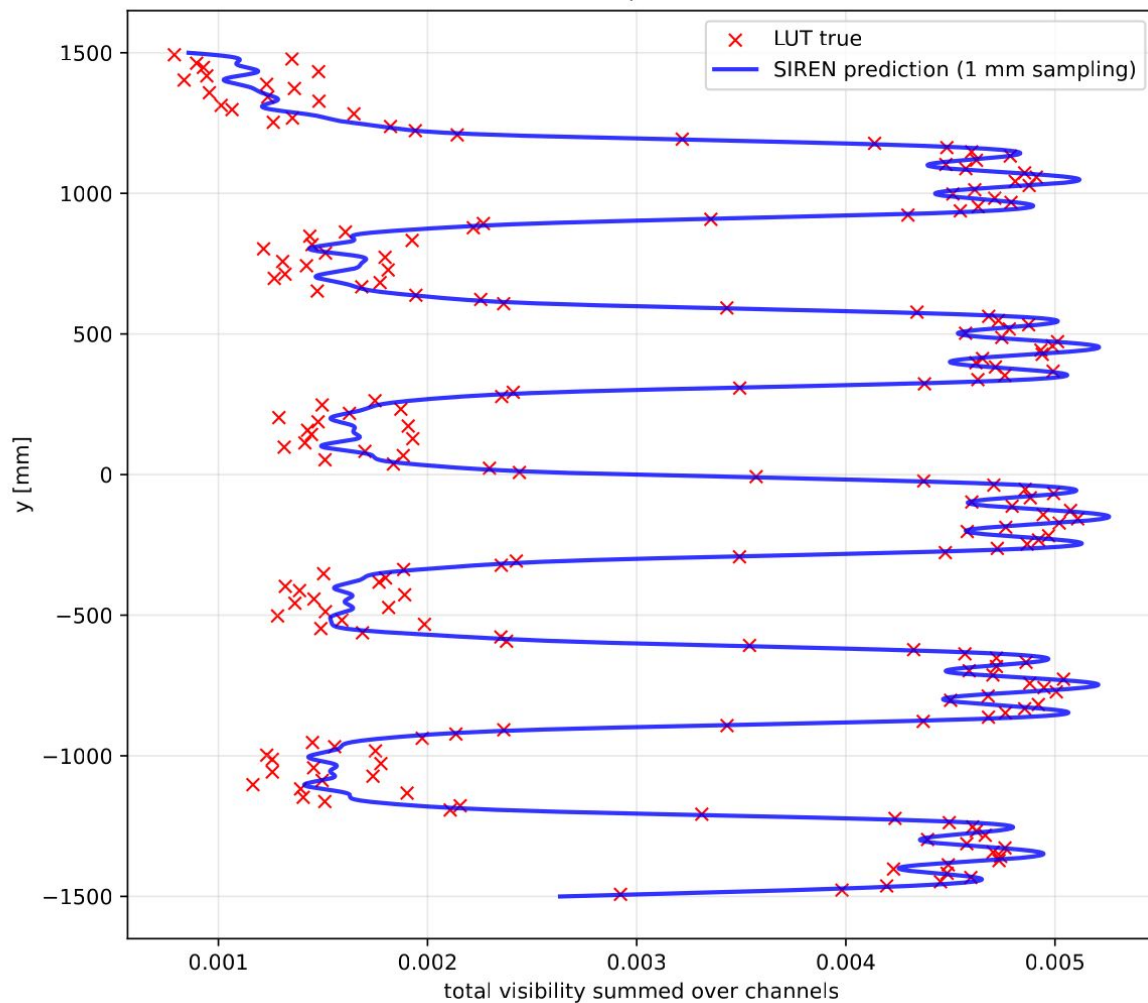


True total visibility at
 $z = 471.09$ mm, $x = -227.90$ mm

1024x10
300MB

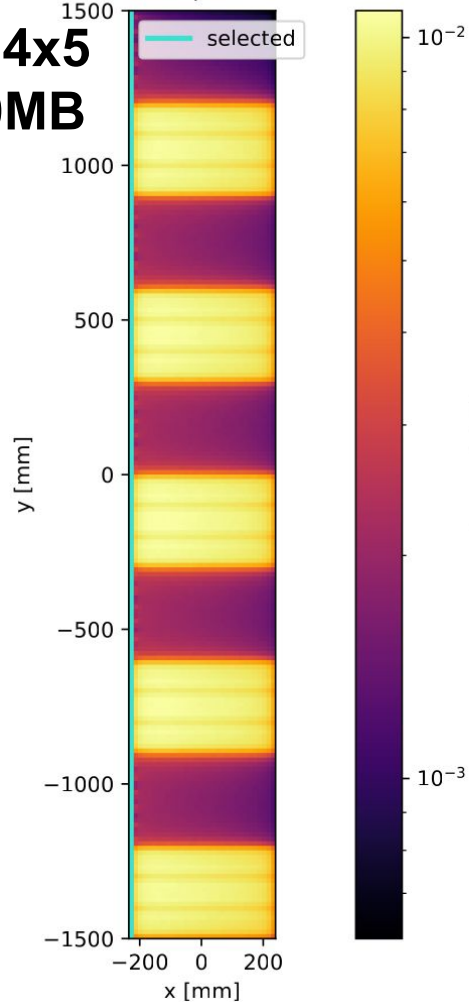


Selected voxel line
 $z = 471.09$ mm, $x = -227.90$ mm

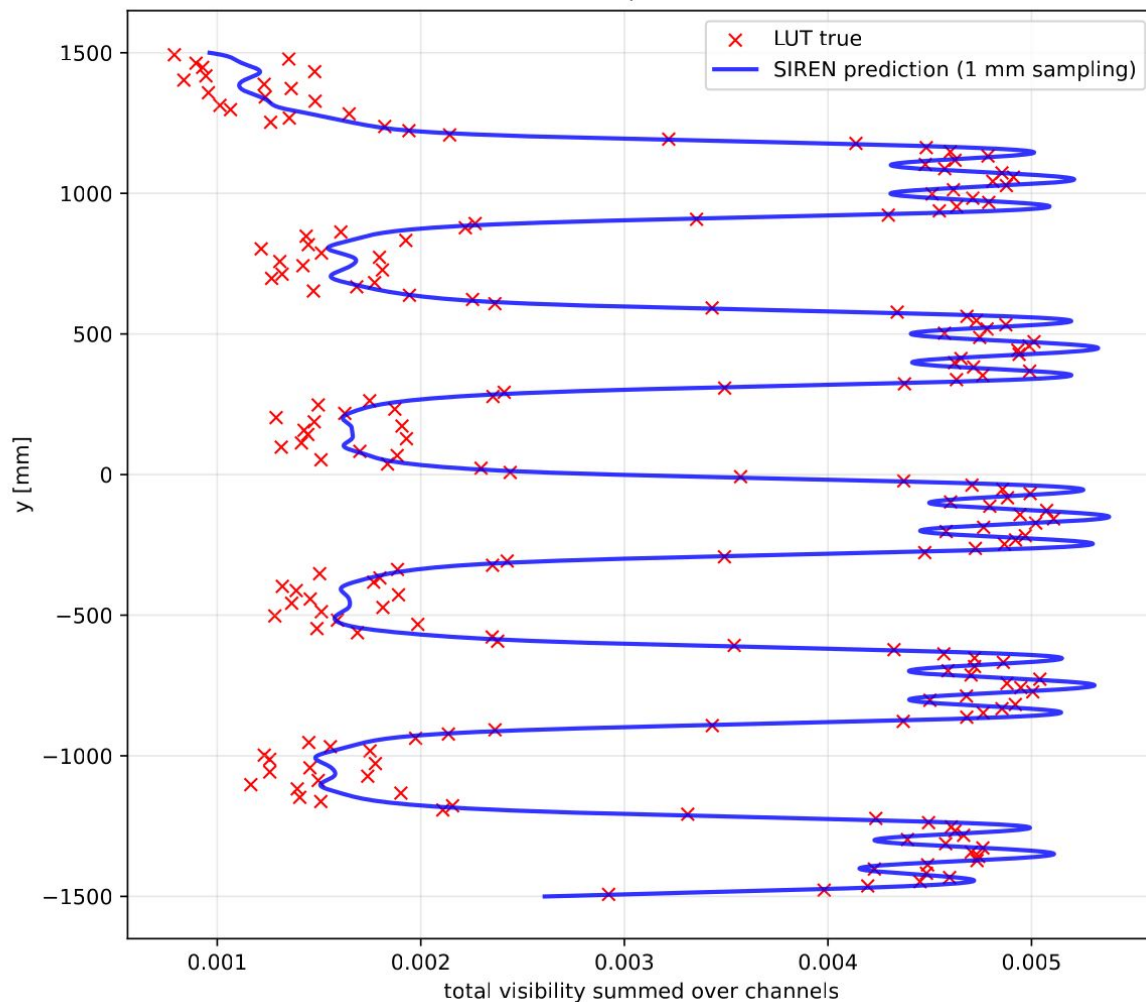


1024x5
~60MB

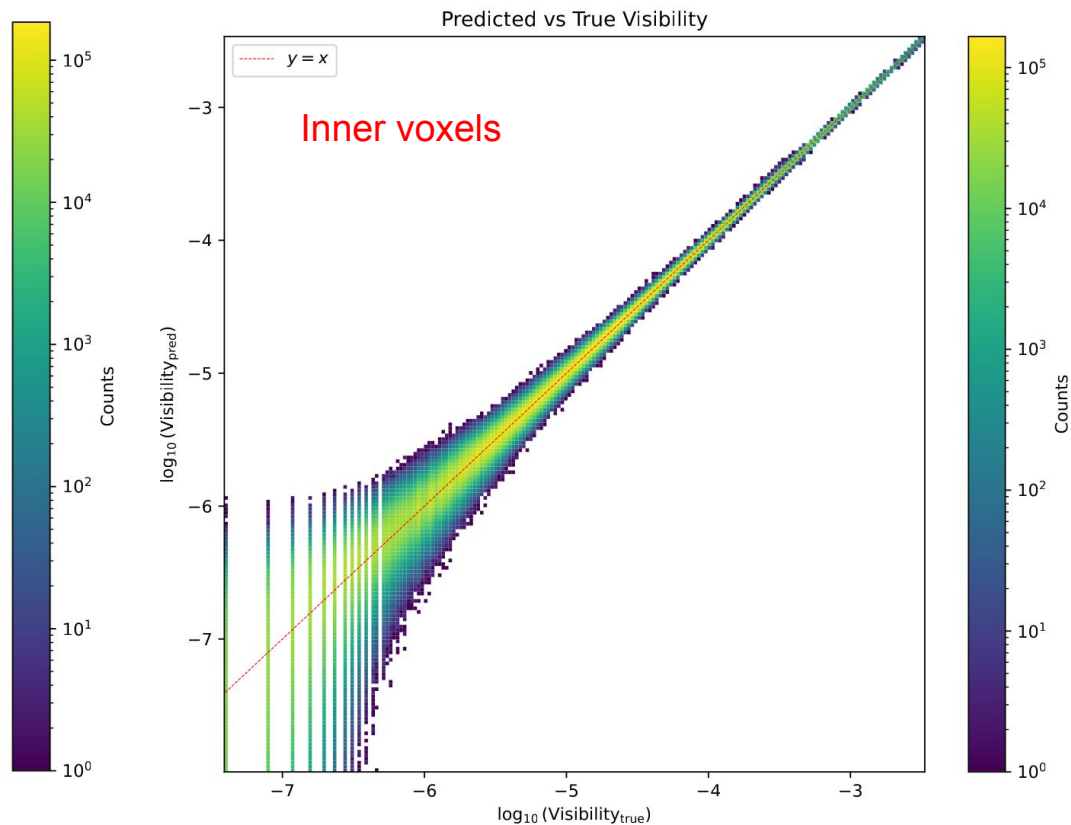
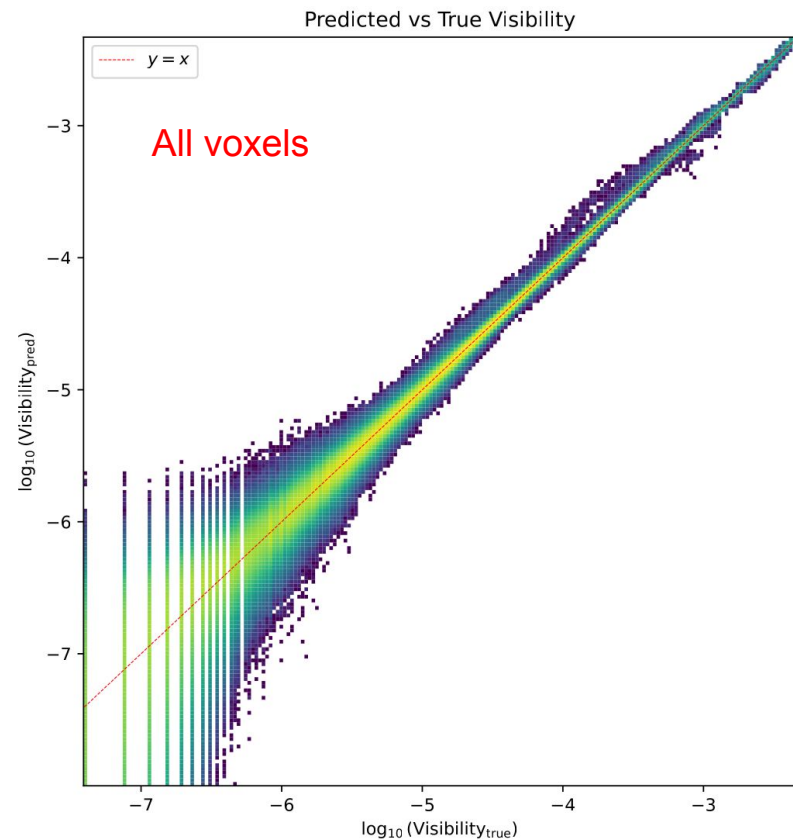
True total visibility at
z = 471.09 mm, x= -227.90 mm



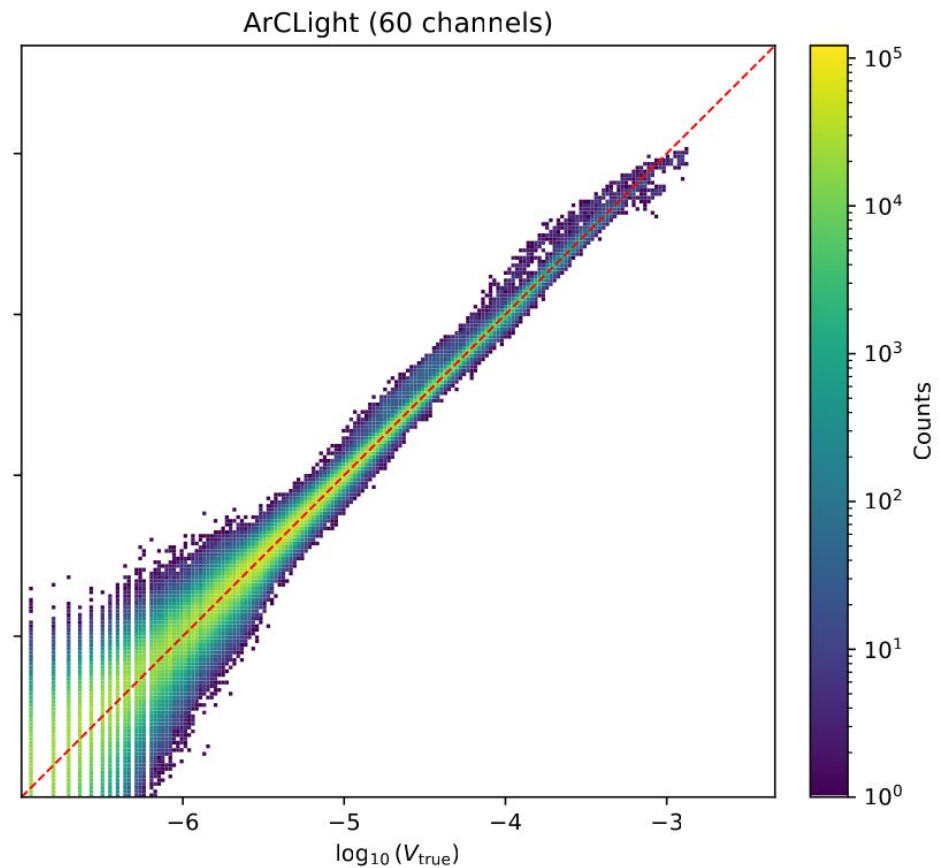
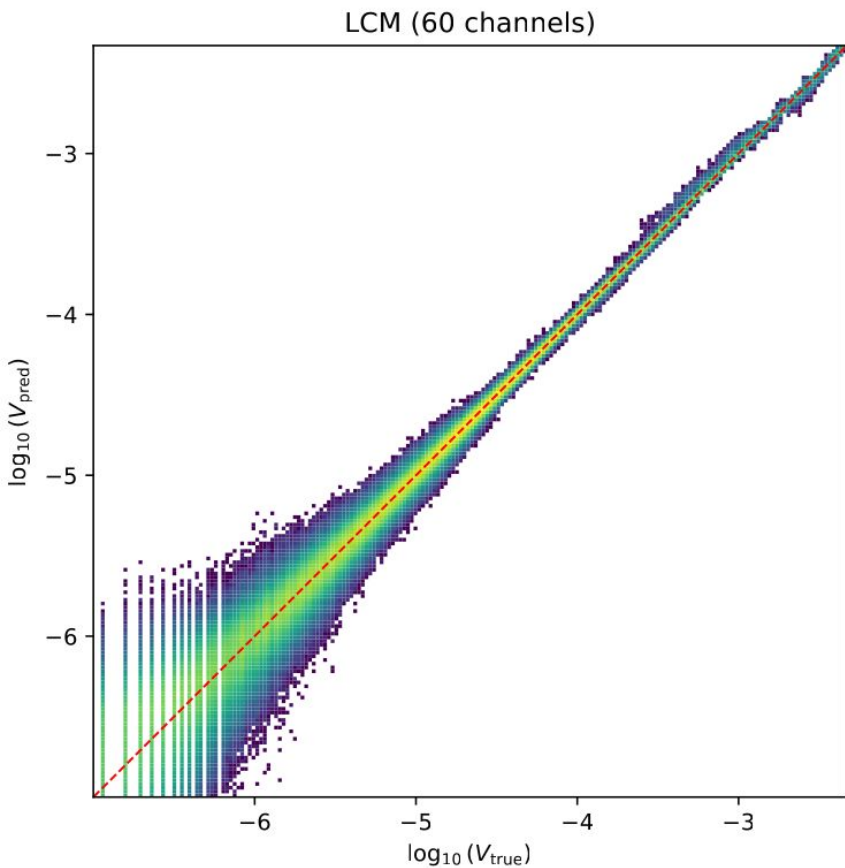
Selected voxel line
z = 471.09 mm, x = -227.90 mm



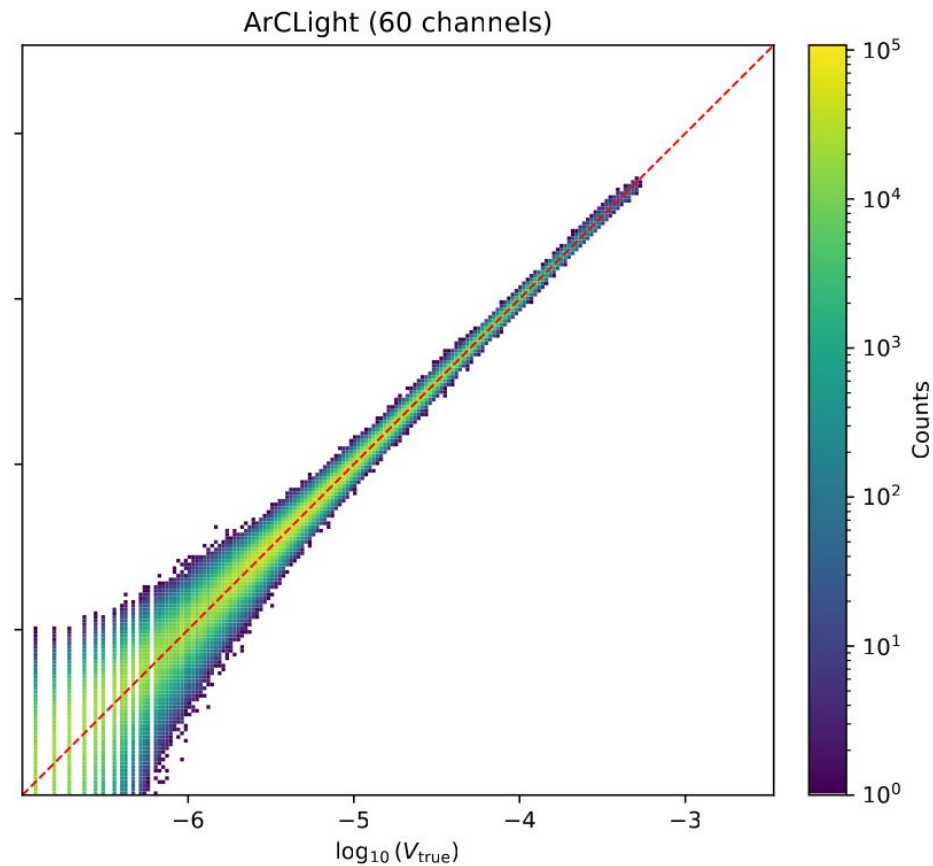
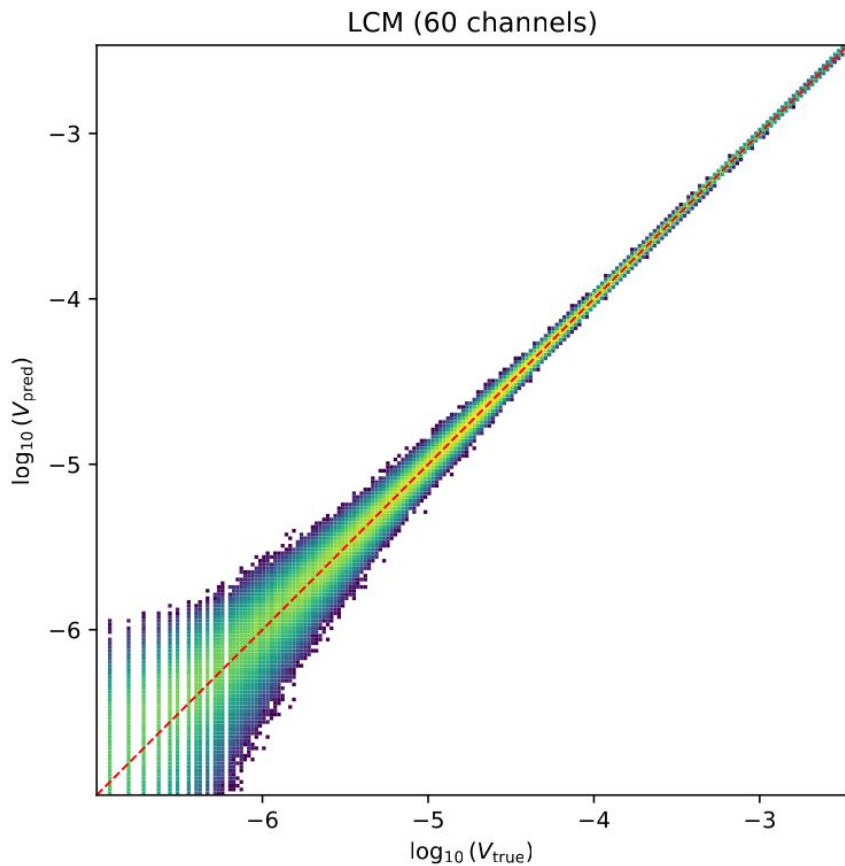
512x10 - All vs inner voxel comparison



512x10 - All voxels



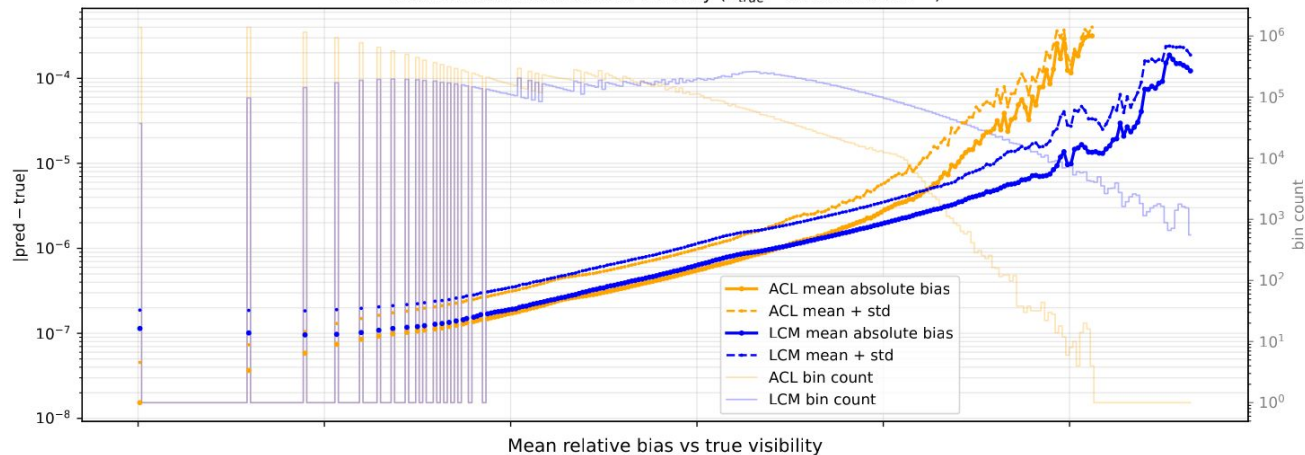
512x10 - Inner voxels



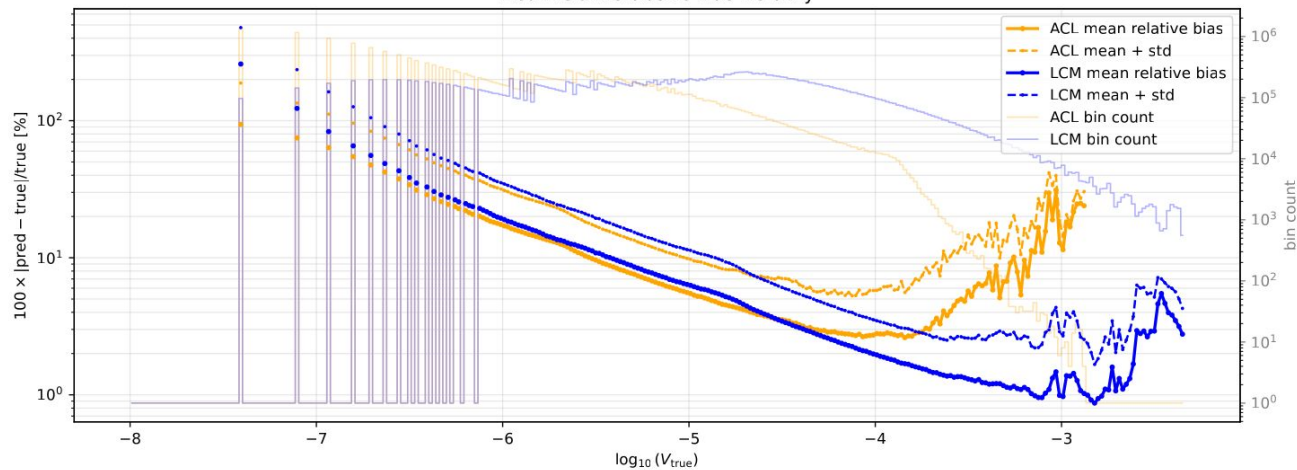
512x10 - All vs inner voxel comparison

Binned bias vs true visibility: ACL vs LCM (300 bins)

Mean absolute bias vs true visibility ($V_{true} = 0$ shown at 10^{-8})



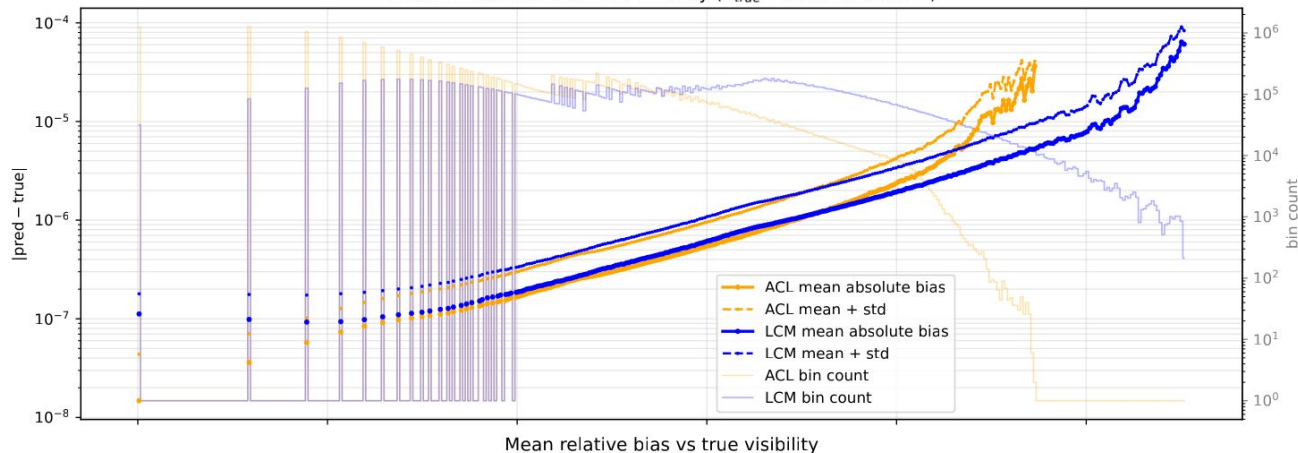
Mean relative bias vs true visibility



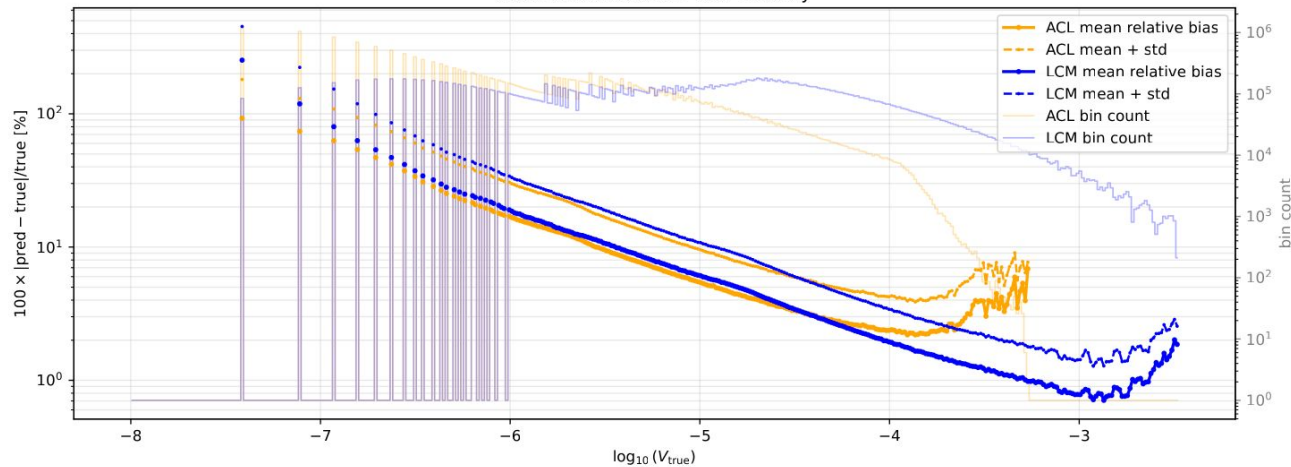
512x10 - All vs inner voxel comparison

Binned bias vs true visibility: ACL vs LCM (400 bins)

Mean absolute bias vs true visibility ($V_{true} = 0$ shown at 10^{-8})



Mean relative bias vs true visibility



Next steps

- Hyperparameter
- Higher training sample, finer LUT voxelization?
- Implement a log-likelihood loss function

Backup slides

Light simulation chain

Step 1 - Photon generation:

Charge Recombination model (E-Field)

$$\frac{dN_\gamma}{dx} = \left(W_{\text{ph}}^{-1} - \mathcal{R} W_{\text{ion}}^{-1} \right) \frac{dE}{dx}$$

Step 2 - Light propagation:

Use LUT to predict N_{pe} on SiPM channels (i) of track segment (j)

$$N_{\text{pe},ij} = \epsilon_i V_i [i_x, i_y, i_z] \left(\frac{dN_\gamma}{dx} \right)_j dx_j$$

LUT visibility map

edep segment length

edep-sim (Geant4 wrapper)
Runs on CPU

Passage of particles
through matter

larnd-sim
Runs on GPU

1.)

Recombination

2.)

Light
propagation

Photocurrent
simulation

Optical detector
electronics

Drifting
electrons

Induced current
on pixel

Pixelated readout
response

Ionization electrons simulation

Scintillation light simulation

HDF5 output file
in LArPix format

Light simulation chain

Step 3 - Time distr. per channel:

Use time distribution of LUT to predict time of photon hits per channel (i)

$$I_{pe,i}[i_t] = \frac{\Delta t_{LUT}}{\delta t} \sum_j f_i[i_x, i_y, i_z, i_t - t_0/\Delta t_{LUT}] N_{pe,i,j},$$

LUT time distr.

time of energy dep.

Step 4 - Apply scintillation model:

Convolve the scintillation Kernel with the time distribution I_{pe} of channel (i)

Kernel:

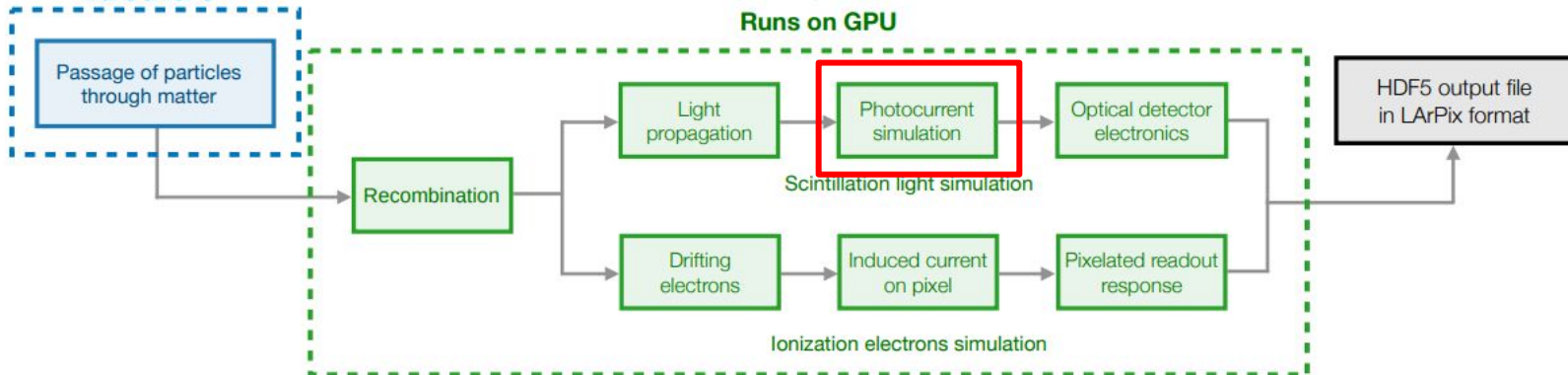
$$f_{scint}(t) = \frac{f_p}{\tau_p} e^{-t/\tau_p} + \frac{1-f_p}{\tau_s} e^{-t/\tau_s}$$

$$I_{scint,i}[i_t] = \sum_{j_t=j_{t,min}}^{i_t} I_{pe,i}[j_t] * f_{scint}((i_t - j_t)\delta t),$$

f_p prompt fraction, τ_p/τ_s : prompt & slow decay times

edep-sim (Geant4 wrapper)
Runs on CPU

larnd-sim
Runs on GPU



Light simulation chain

Step 5 - Bin-wise Poisson smearing:

Apply Poisson smearing on N_{pe} for each time bin

$$\lambda_i = I_{\text{scint}}[i_t] \delta t, \quad I_{\text{pe, disc}}[i_t] \sim \frac{\text{Poisson}(\lambda_i)}{\delta t}$$

edep-sim (Geant4 wrapper)
Runs on CPU

Passage of particles
through matter

larnd-sim
Runs on GPU

Recombination

Light
propagation

Photocurrent
simulation

Optical detector
electronics

Drifting
electrons

Induced current
on pixel

Pixelated readout
response

Scintillation light simulation

Ionization electrons simulation

HDF5 output file
in LArPix format

Log transformation for visibility

Suppose looking at two true visibilities:

$$v(\text{near-true}) = v_{n,t} = 10^{-3} ; v(\text{far-true}) = v_{f,t} = 10^{-3}$$

And two predicted visibilities:

$$v(\text{true-pred}) = v_{n,p} = 2 \times 10^{-6} ; v(\text{far-pred}) = v_{f,p} = 2 \times 10^{-6}$$

Using MSE directly:

- $\text{Loss}(\text{near}) = (v_{n,p} - v_{n,t})^2 = 10^{-12}$
- $\text{Loss}(\text{far}) = (v_{f,p} - v_{f,t})^2 = 10^{-6}$

Using log scale in MSE:

- $\text{Loss}(\text{near}) = (\log_{10}(v_{n,p}) - \log_{10}(v_{n,t}))^2 = (\log_{10}(2 \times 10^{-3}) - \log_{10}(10^{-3}))^2 = \log_{10}(2)^2$
- $\text{Loss}(\text{far}) = (\log_{10}(v_{f,p}) - \log_{10}(v_{f,t}))^2 = \log_{10}(2)^2$

SIREN - overview

First layer

$$h_i^{(0)} = \sin \left(\omega_0 \left(\sum_{j=1}^3 W_{ij}^{(0)} x_j + b_i^{(0)} \right) \right)$$

$$\mathbf{f}: \mathbf{R}^3 \rightarrow \mathbf{R}^{n_F}$$

n_F : Number of features

First hidden layer

$$h_i^{(1)} = \sin \left(\omega_0 \left(\sum_{j=1}^{n_F} W_{ij}^{(1)} h_j^{(0)} + b_i^{(1)} \right) \right)$$

$$\mathbf{f}: \mathbf{R}^{n_F} \rightarrow \mathbf{R}^{n_F}$$

n_L : Number of layers

w_0 : Constant for all layers

N : Number of voxels per batch

C : Number of channels =120

k-th layer

$$h_i^{(k)} = \sin \left(\omega_0 \left(\sum_{j=1}^{n_F} W_{ij}^{(k)} h_j^{(k-1)} + b_i^{(k)} \right) \right)$$

$$\mathbf{f}: \mathbf{R}^{n_F} \rightarrow \mathbf{R}^{n_F}$$

Output layer

$$h_i^{(n_L+1)} = \sum_{j=1}^{N_{ch}} W_{ij}^{(n_L+1)} h_j^{(n_L)} + b_i^{(n_L+1)}$$

$$\mathbf{f}: \mathbf{R}^{n_F} \rightarrow \mathbf{R}^{C=120}$$

Loss: weighted-MSE

$$L = \frac{1}{NC} \sum_{i=1}^N \sum_{j=1}^C w_{ij} (\hat{v}_{ij} - v_{ij})^2$$