Testing oversampling and DirectReco parameters

Kasper Pedersen, Msc. Student
IceCube, NBI
5/2/21

UNIVERSITY OF COPENHAGEN

Intro

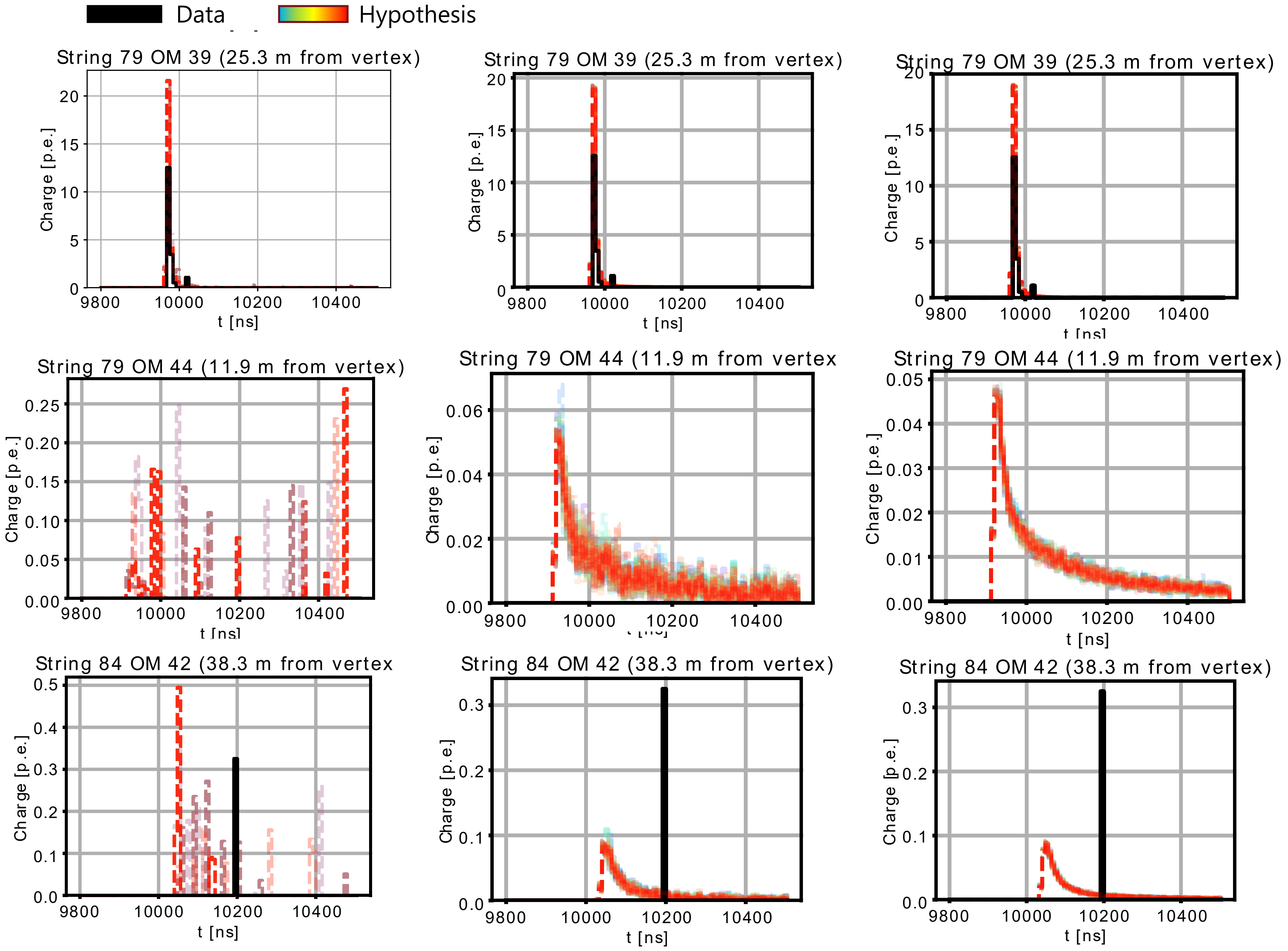
- We are now moving on to reconstruction in DirectReco.
- Check stability of DirectReco hypothesis at increasing oversampling.
- Run reconstruction of multiple events at different oversampling.
  - Using a gaussian smeared truth seed
- See if there is a ‘best’ oversampling.
- Expect to get better results (more precise and accurate) at higher oversampling.

Stability of hypothesis

- Extract hypothesis many times for same event (no reconstruction)
- See if hypothesis is stable at different oversampling

Select OMs

- Hit DOM example
- No-Hit DOM example
- Noise hit example



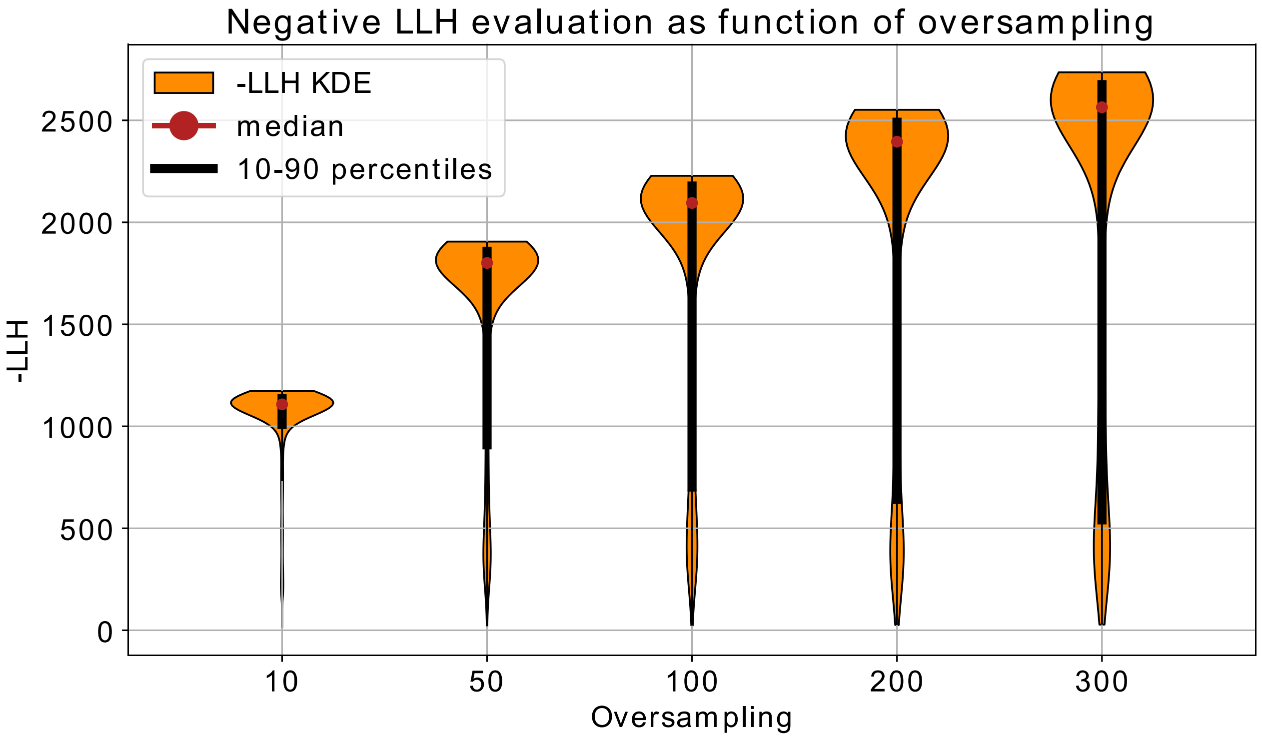
Oversampling

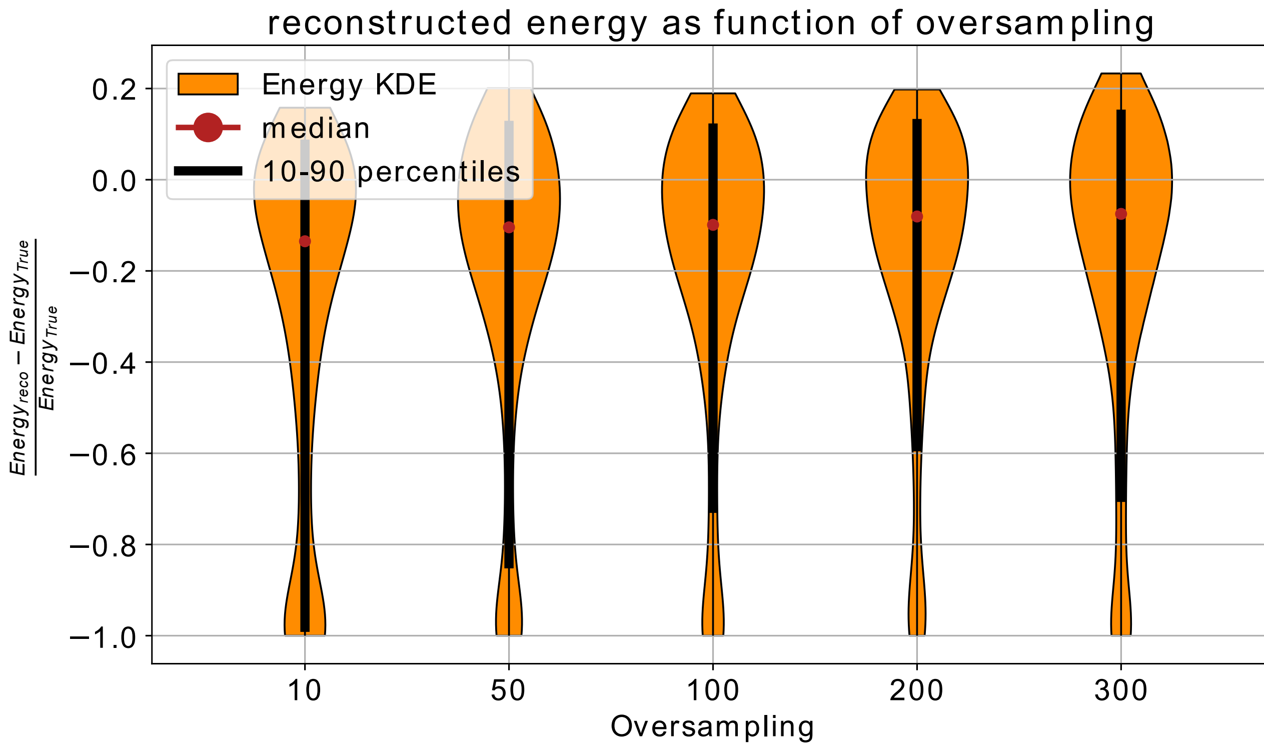
Reconstruction

- Run reconstruction of multiple events at different oversampling
- Check LLH and performance of reconstruction of physical parameters
- A Gaussian smearing has been applied to the truth seed

First Run

- LLH goes up and get a wider distribution as oversampling goes up
- Energy has small outlier distributions as well
- Need to look into the likelihood (next slide)





Dima LLH definition

Millipede implementation

from sec. 3 in 2018 dima paper

$$LH_{ratio} = \Pi_i \left(\frac{\mu_i}{\frac{s_i}{n_s}}\right)^{s_i} \cdot \Pi_i \left(\frac{\mu_i}{\frac{d_i}{n_d}}\right)^{d_i} \quad , \mu_i = \frac{s_i + d_i}{n_s + n_d}$$

$$, \mu_i = \frac{n_s \cdot s_i + d_i}{n_s + n_d}$$

$$\ln(LH_{ratio}) = \Sigma_i s_i \ln\left(\frac{n_s \mu_i}{s_i}\right) + d_i \ln\left(\frac{n_d \mu_i}{d_i}\right)$$

$\neq$

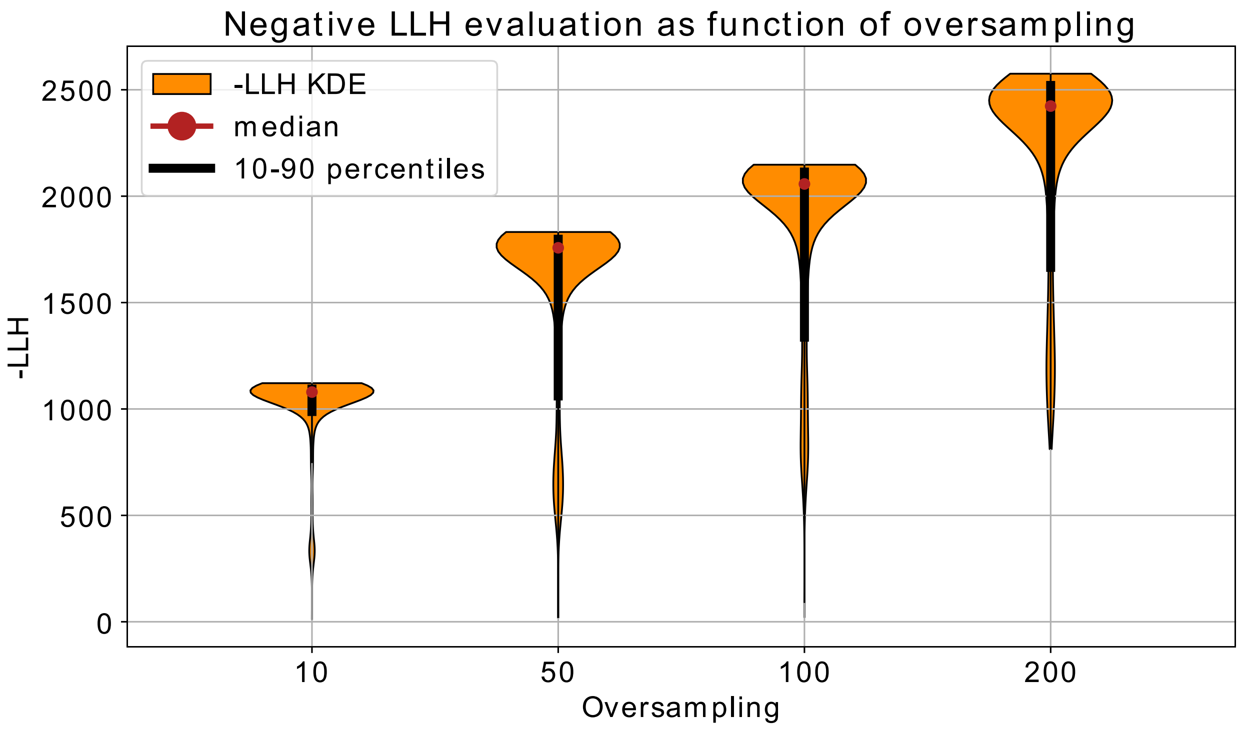
$$\ln(LH_{ratio}) = \Sigma_i n_s \cdot s_i \cdot \ln\left(\frac{\mu_i}{s_i}\right) + n_d \cdot d_i \cdot \ln\left(\frac{\mu_i}{d_i}\right)$$

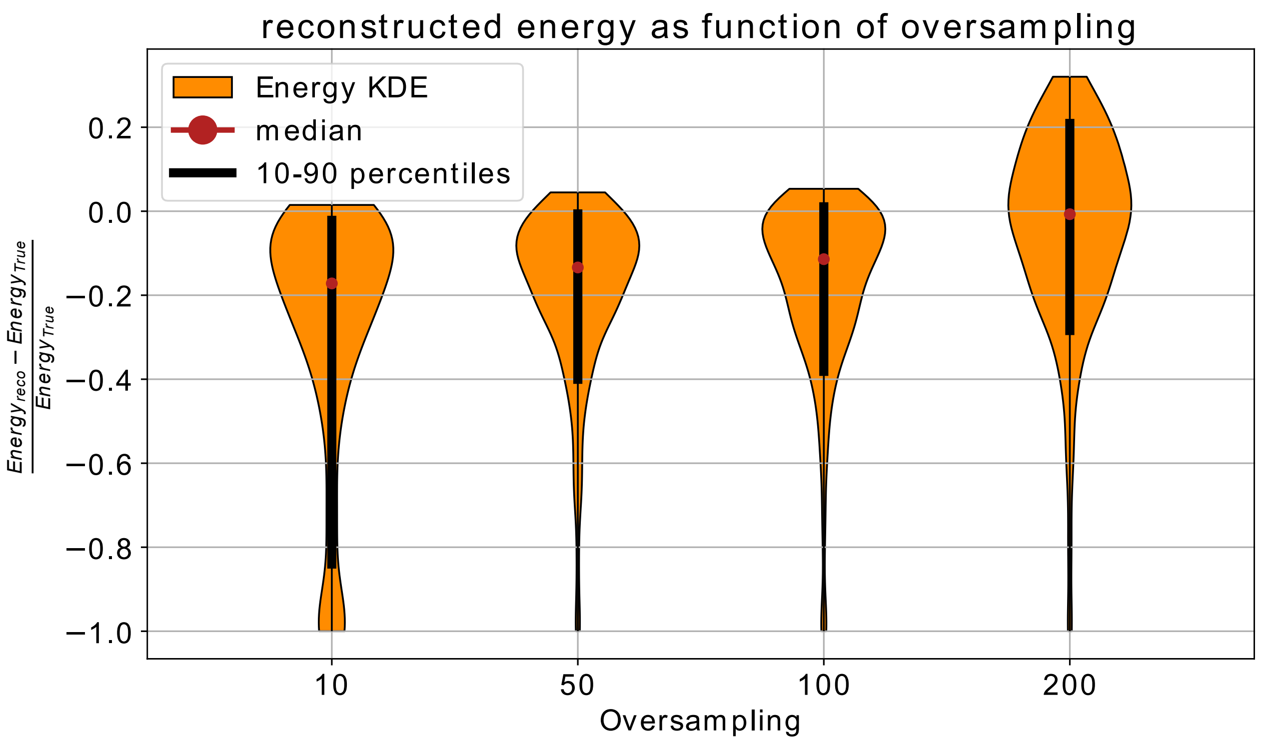
- It doesn't match up
- Try implementing own calculation (new DimaLLH)

n_d = number of 'data trials' i.e. 1 for reco
d_i = observed charge
n_s = number of 'sim trials' i.e. oversampling factor
s_i = expected charge

Second Run new DimaLLH

- LLH still goes up and gets wider but more stable
- Energy gets better with oversampling





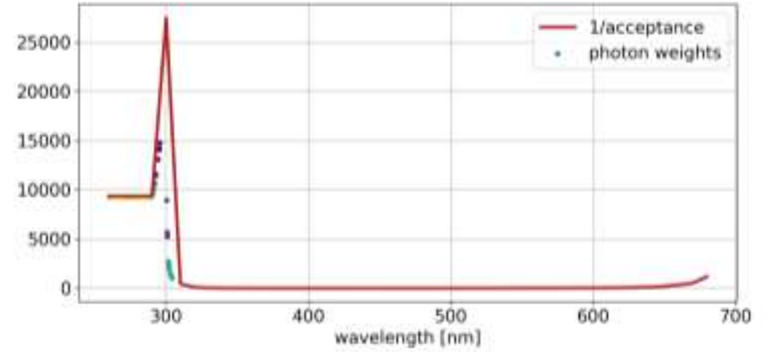
Clshim weights and new DimaLLH

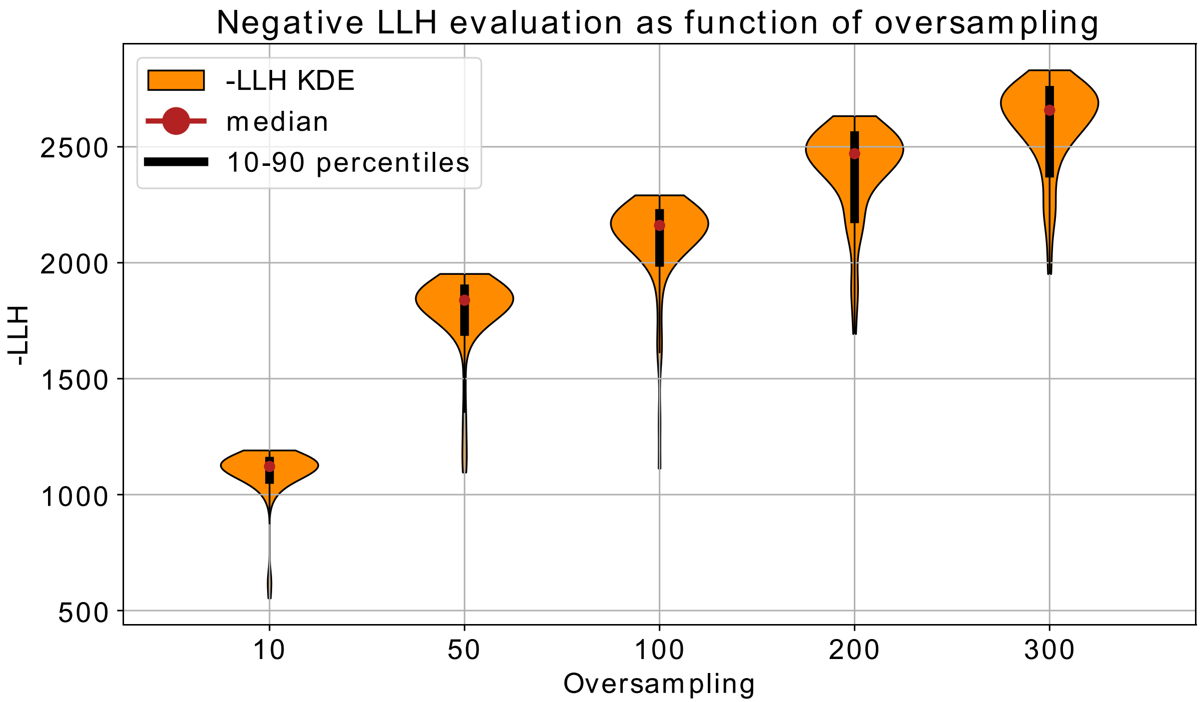
- In CLShim.cxx :

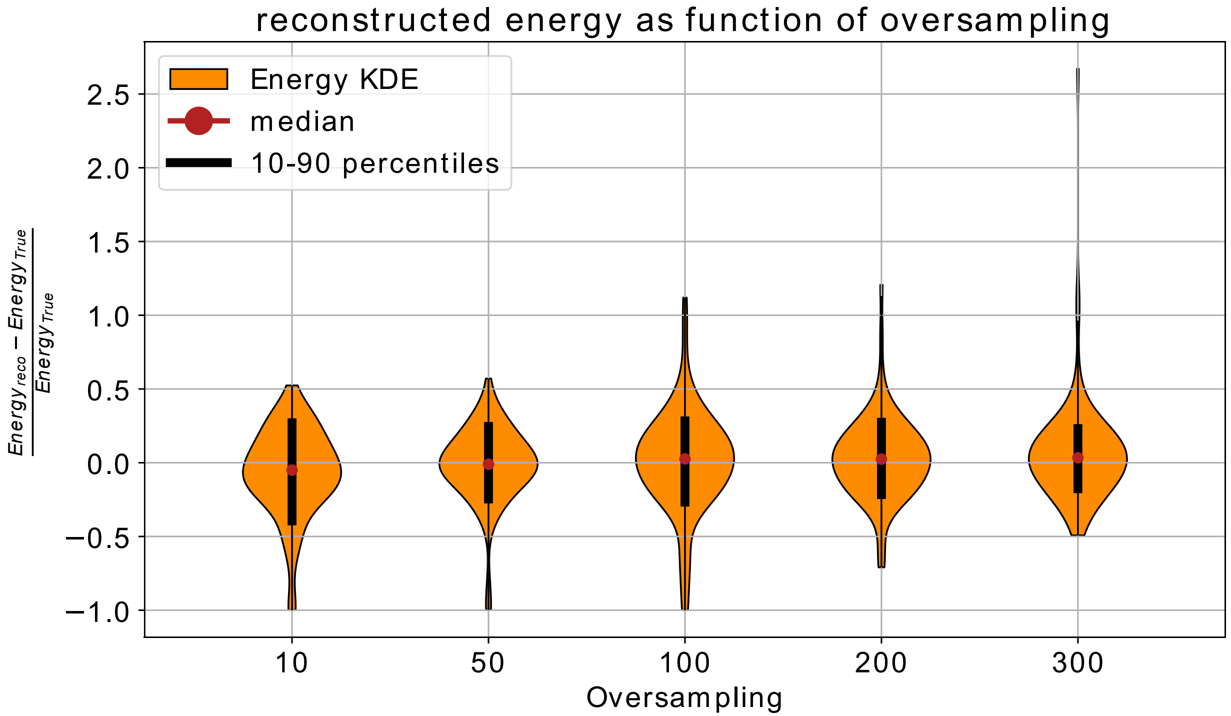
Weight = photon weight*wavelength acceptance* angular acceptance / oversample factor

- Removed photon weight from equation
- LLH even better
- Energy more stable

Remember in simulation photon weights = 1/(wavelength acceptance)

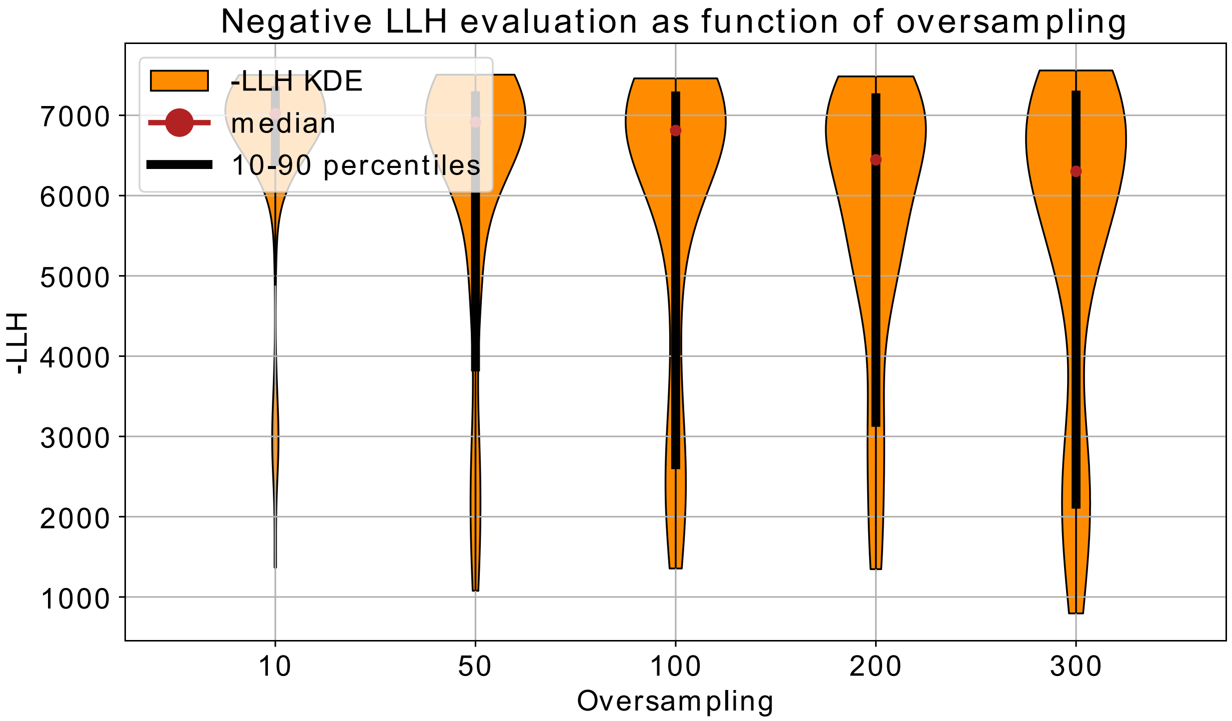


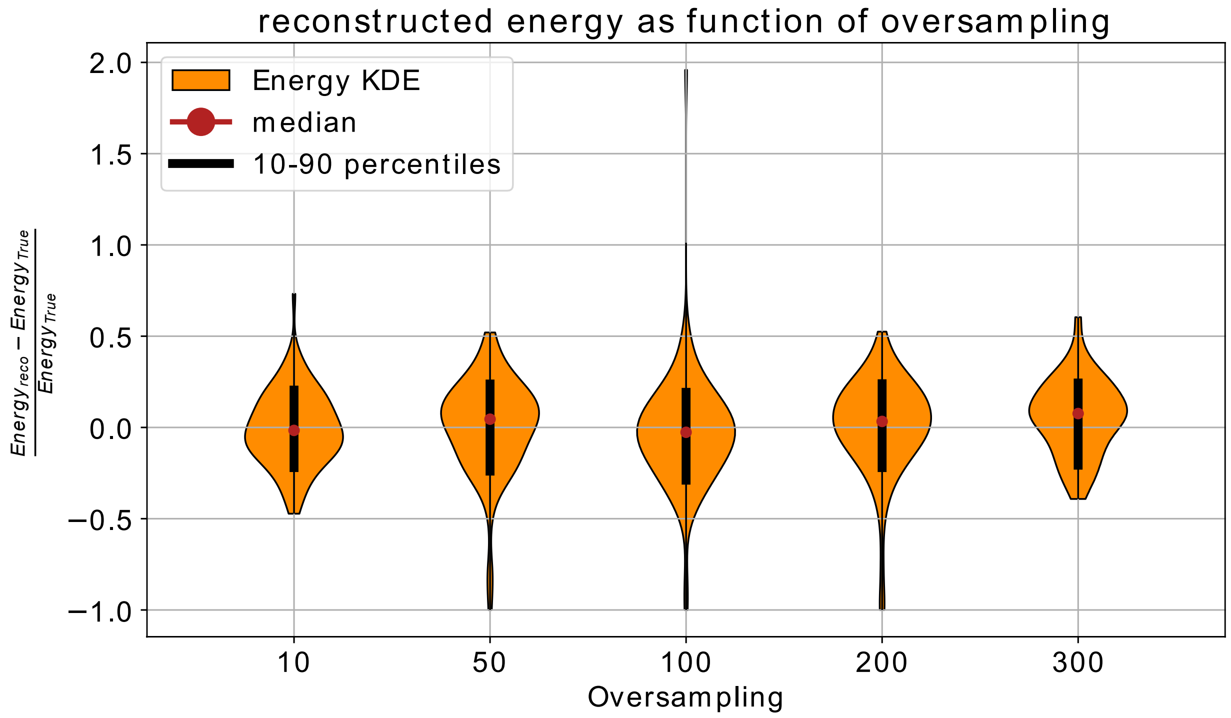




How is the normal Poisson LLH doing?

- LLH goes down with higher oversampling
- Energy is stable and comparable to new Dima LLH

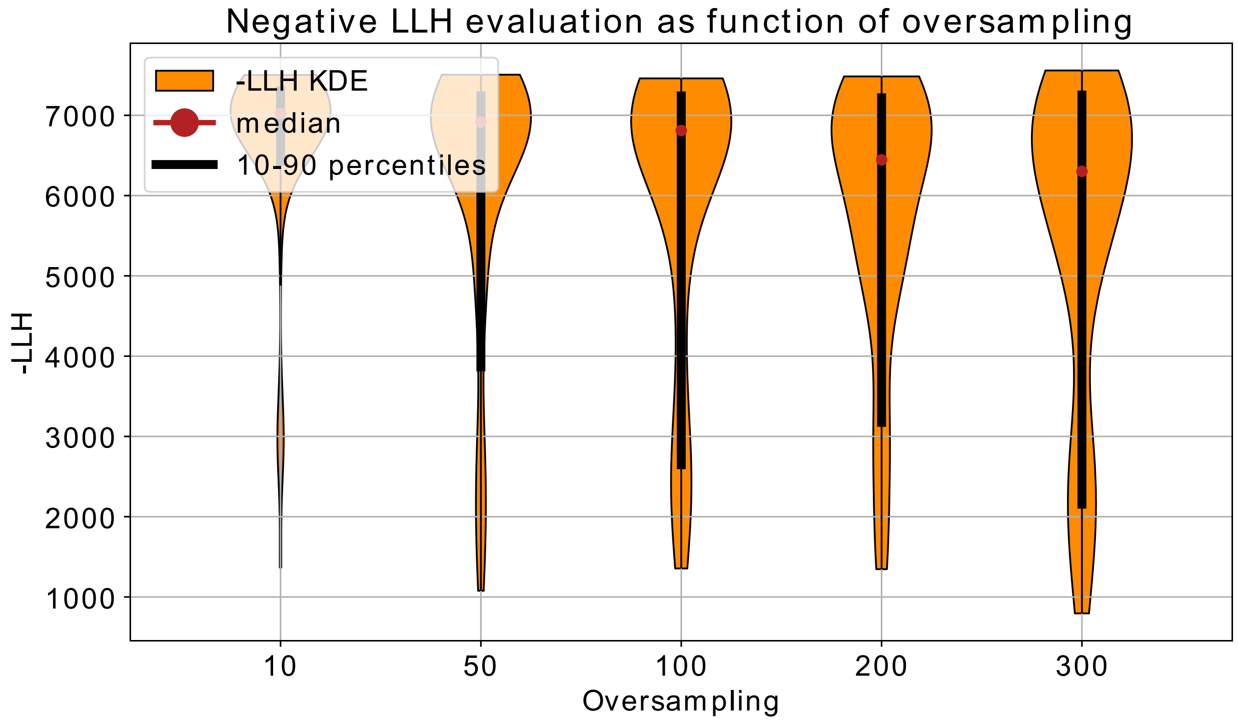


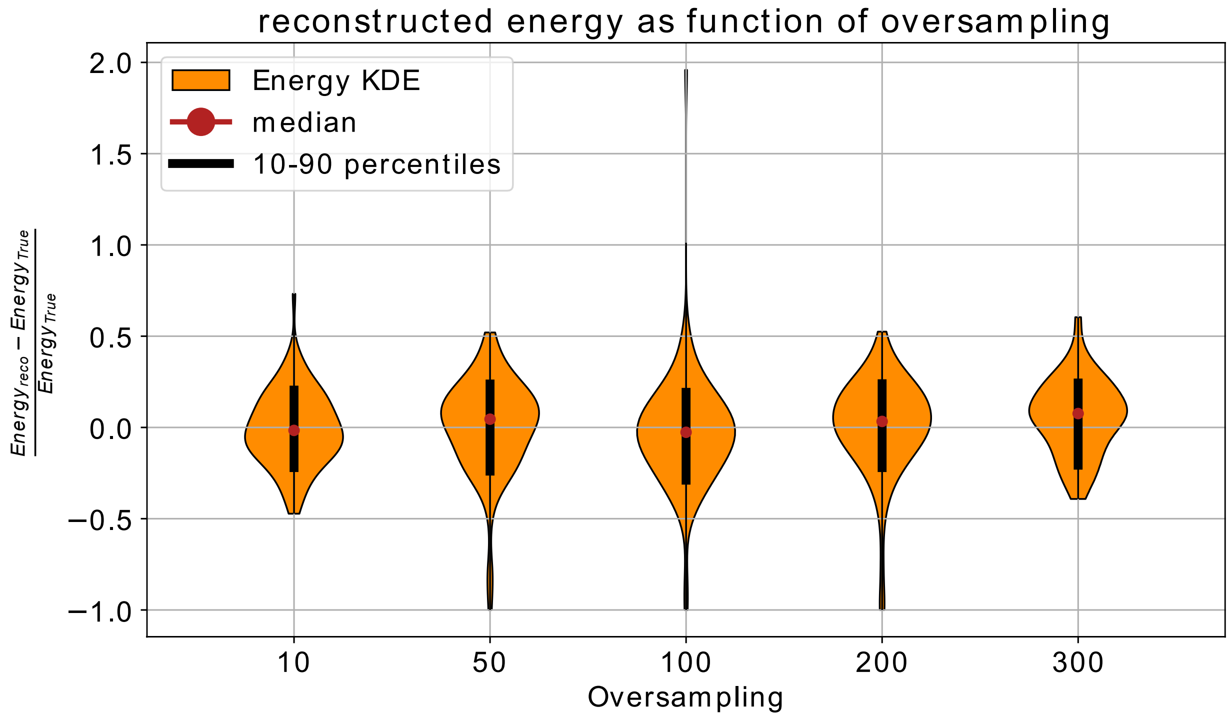


How is the normal Poisson LLH doing?

- LLH goes down with higher oversampling
- Energy is stable and comparable to new Dima LLH

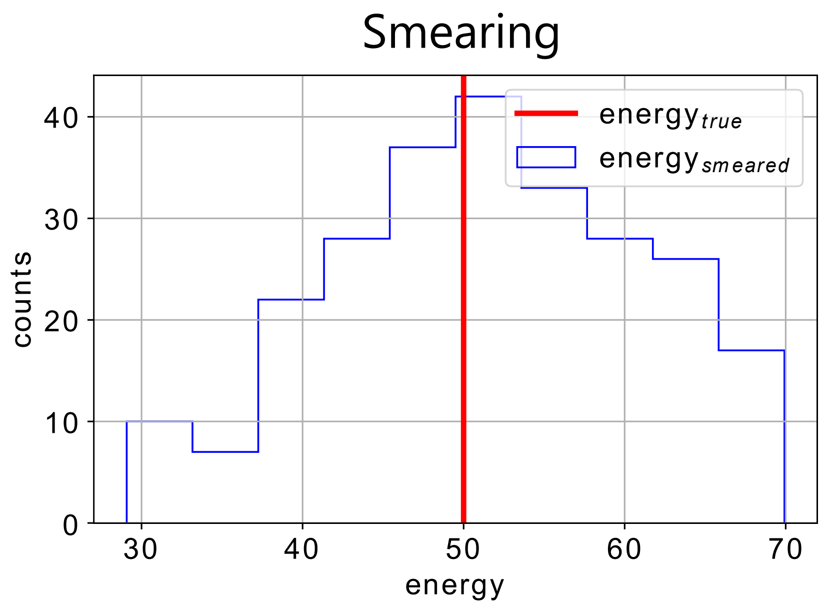
Is the minimizer moving away from seed?

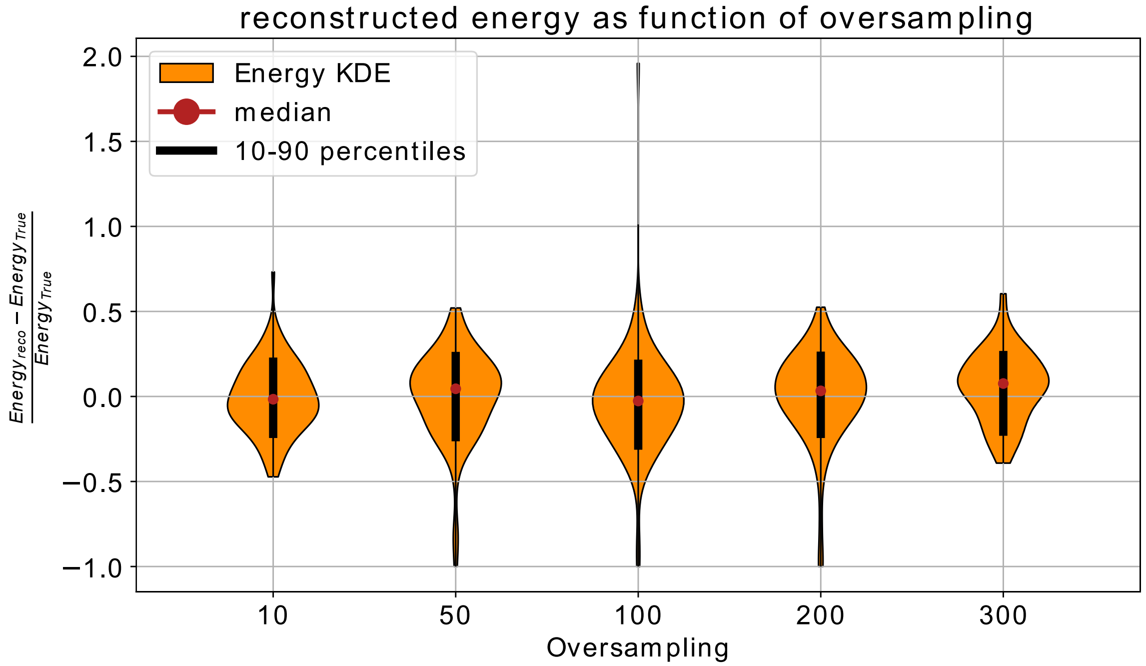




Smearing Issue?

- Is the shape of the energy reco basically just the minimizer not moving away from input seed?
- Now checking with a constant added to seed instead of drawn from a gaussian, same as Jonathan is doing
- Make energy-error plot using Energy seed instead of Energy reco and compare ?
- Save and plot minimizer movement at each iteration ?





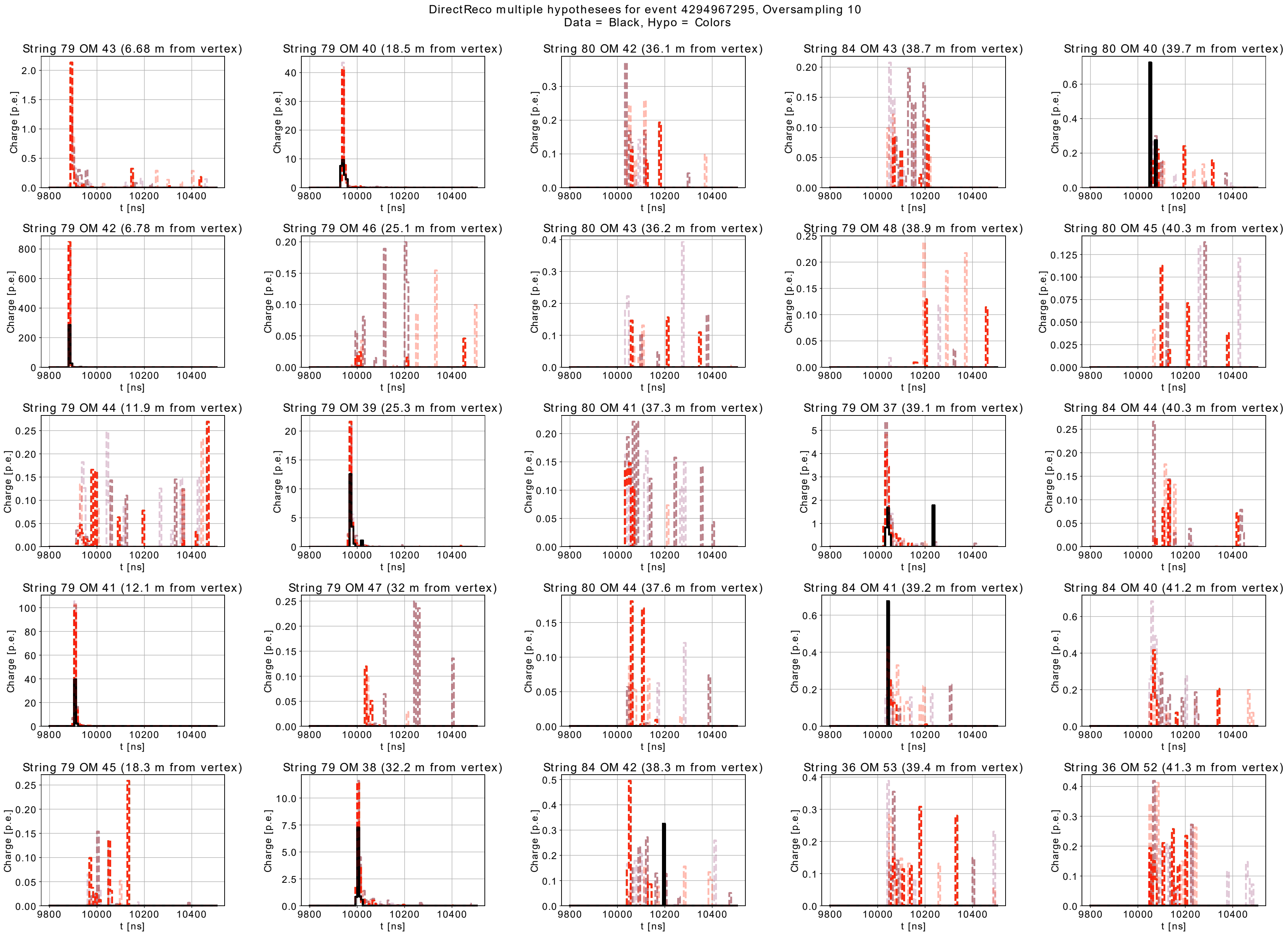
Future Work

- Why is Dima LLH getting worse with oversampling?

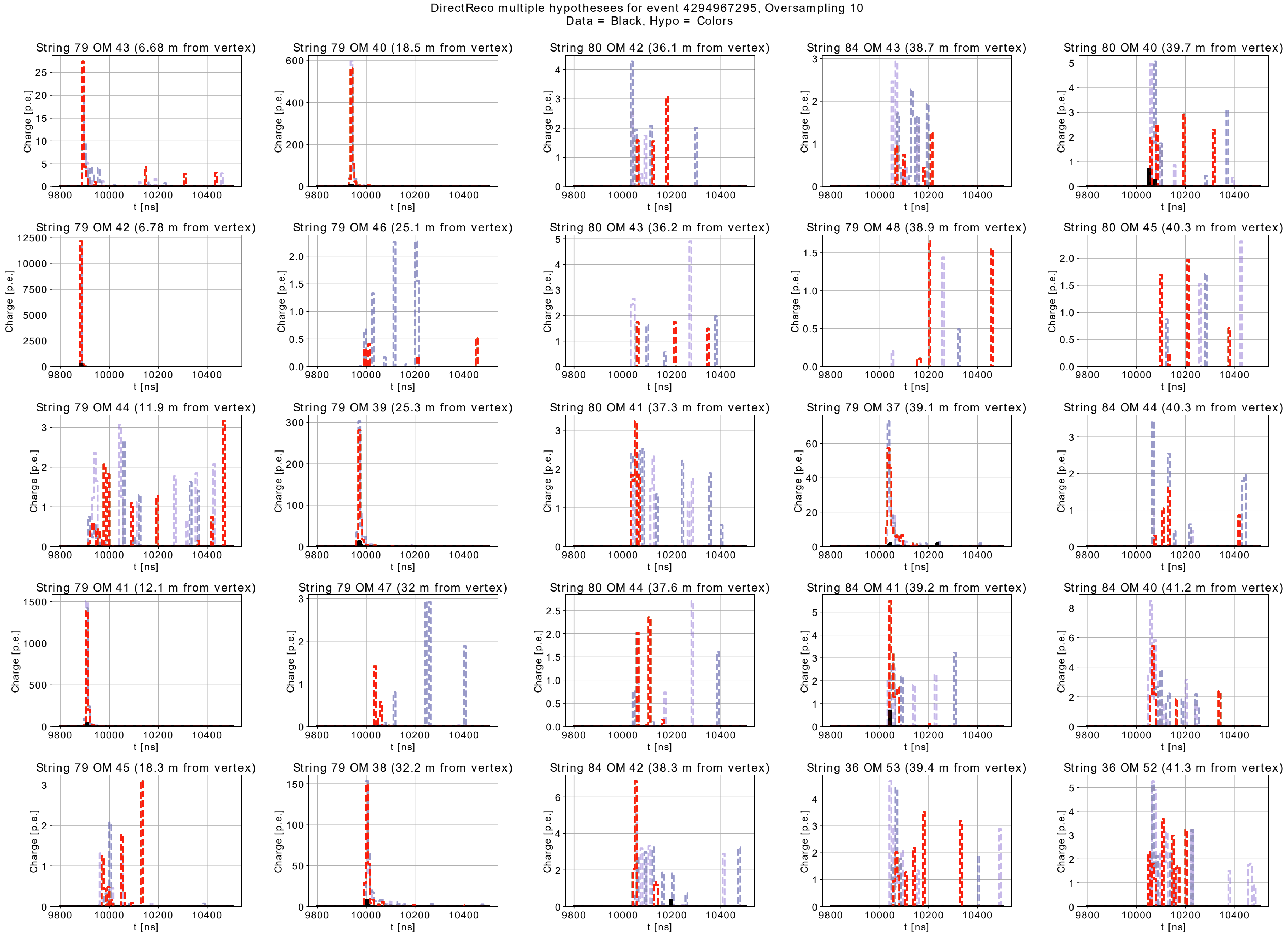
- Is minimizer moving?

Bonus Slides

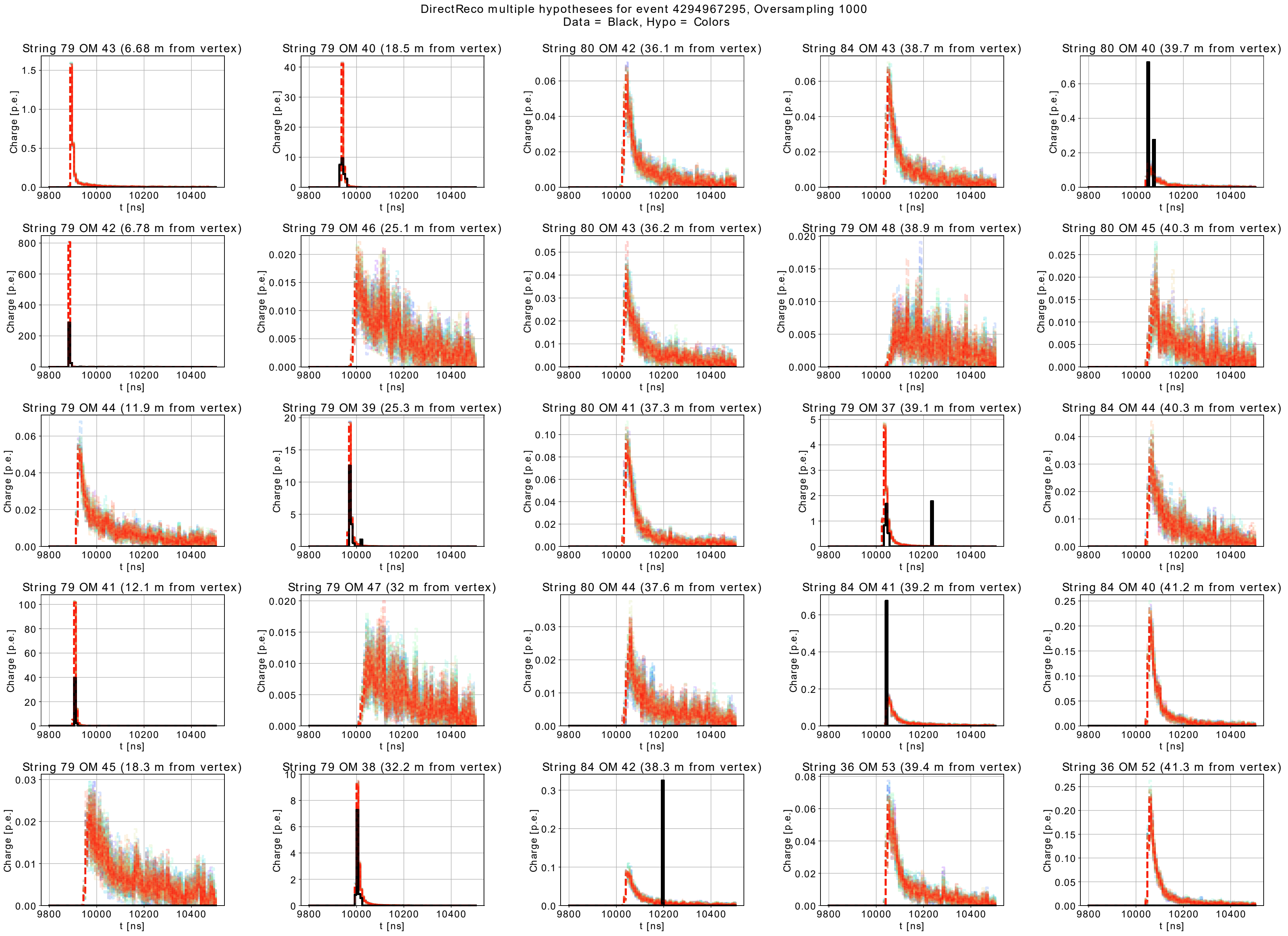
New CLShim Weight



Old CLShim Weight



New CLShim Weight



Old CLShim Weight