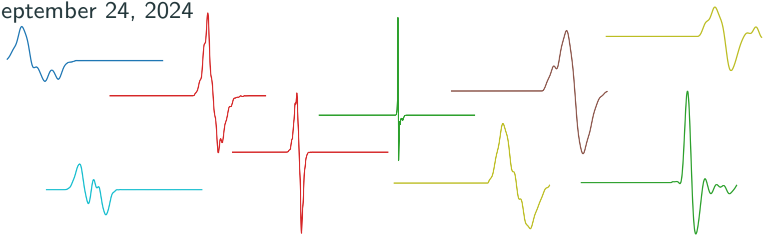
UPPSALA
UNIVERSITET

A surrogate model for the generation of radio pulses from neutrinos for IceCube-Gen2

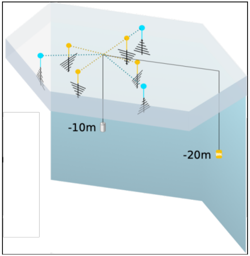
Philipp Pilar, Christian Glaser, Niklas Wahlström

September 24, 2024



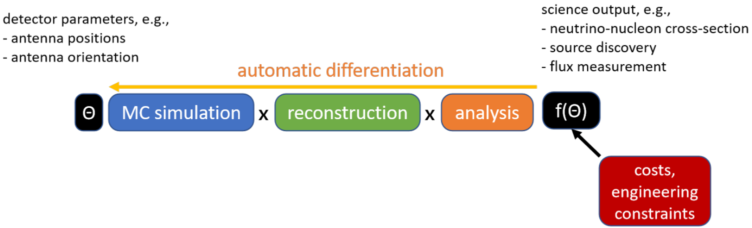
Overview – IceCube-Gen2 detector

- The IceCube detector at the South Pole can detect cosmological high-energy neutrinos.
- IceCube-Gen2 is being developed.
- We consider the radio emission of particle showers.
- Radio detectors can still be optimized.

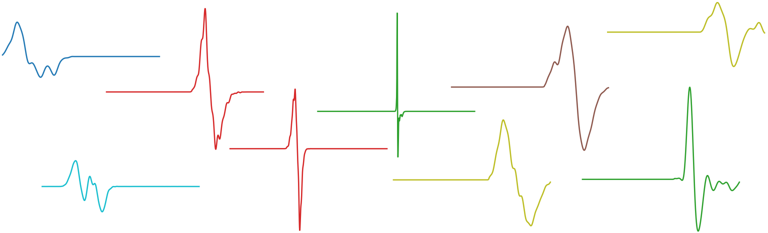


Overview – Aim of the project

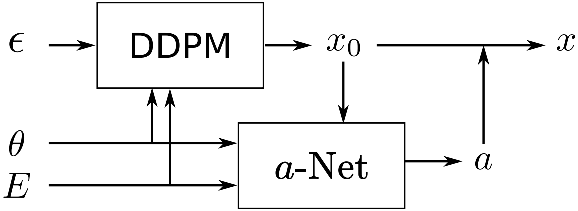
- Part of a bigger aim to create an end-to-end optimization pipeline.
  $\Rightarrow$ NuRadioOpt, see Christian's talk



- We focus on the MC simulation part.
- Train a differentiable surrogate model to generate radio signals.



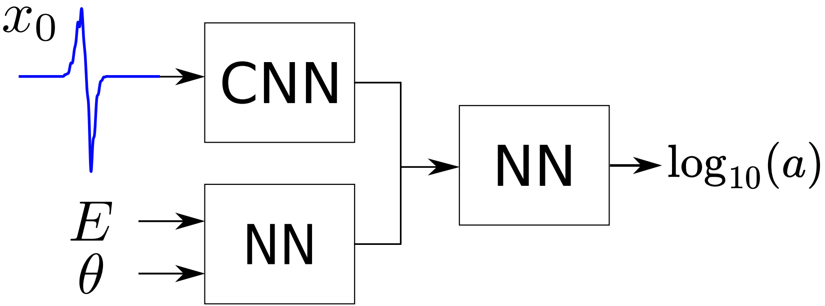
Model architecture



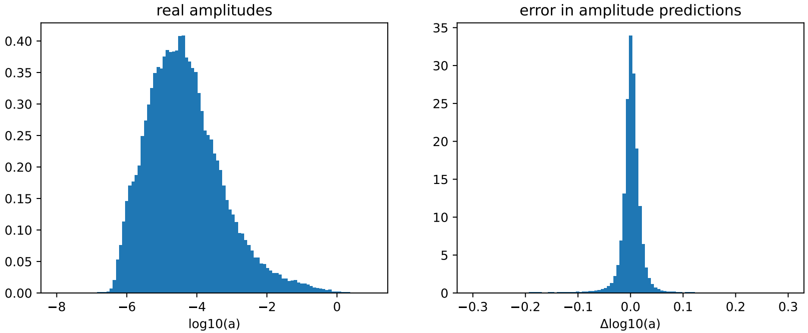
- A denoising diffusion probabilistic model (DDPM) is used to generate normalized samples x_0 from random noise ϵ, conditional on the viewing angle θ and the energy E.
- A neural network is employed to predict the amplitude a of the generated signal (a-Net).
- Subsequently, x_0 and a are combined to form the final signal x.

Amplitude prediction network (a-Net)

- Normalized waveforms x_0 are fed into a neural network to predict the amplitude.
- Combination of convolutional and fully-connected layers.
- Conditional on E and θ.



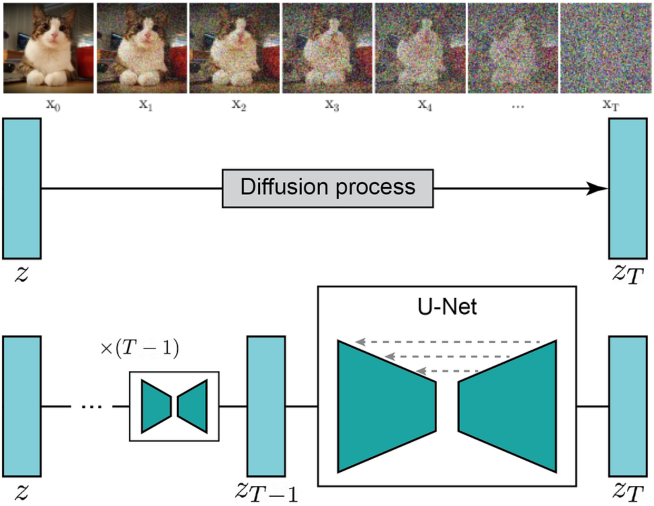
Results: amplitude prediction



Left: the main reason for using the a-Net is the large spread in signal amplitudes.

Right: the network manages to predict the amplitudes with high accuracy.

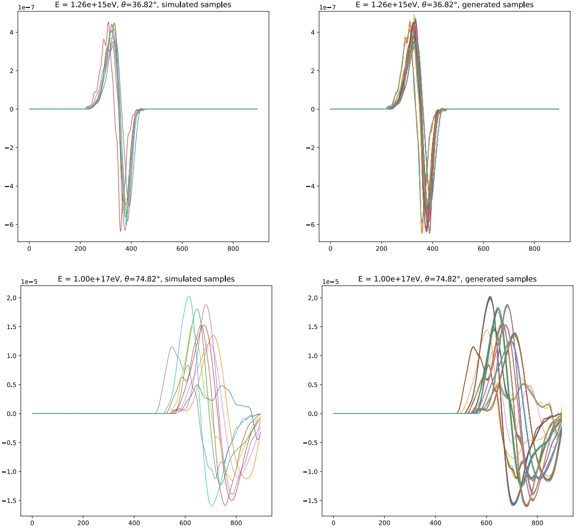
Denoising diffusion probabilistic models (DDPMs)



We use the implementation from https://github.com/lucidrains/denoising-diffusion-pytorch.

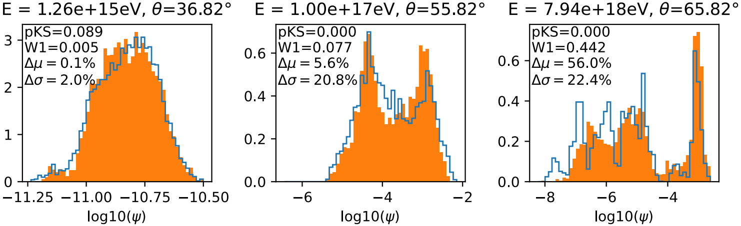
Results: generated samples

real samples generated samples



Results: distribution of summary statistics

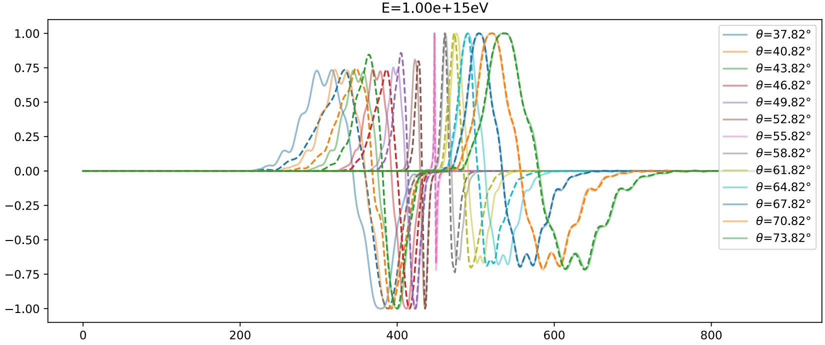
- Use summary statistics to evaluate the accuracy of the generated data distribution.
- E.g., the energy fluence ψ:



- There generally is a good match for low energies.
- At higher energies, we still often have $\Delta\mu > 10\%$.

Results: θ-dependence

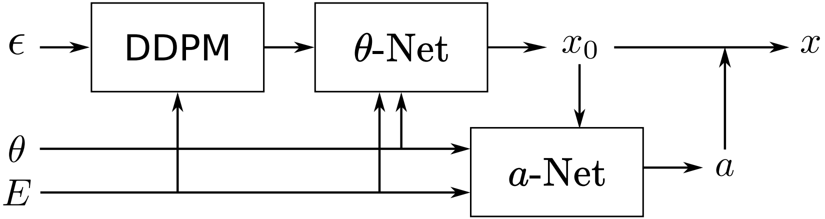
- The model needs to be capable of generating the resulting signals for the same event at different viewing angles θ.



- The vanilla DDPM does not automatically learn the correct θ-dependence.

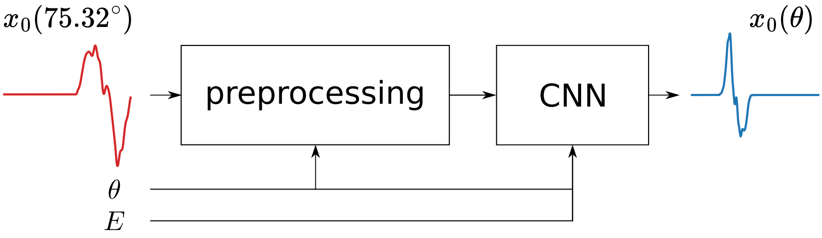
Model architecture to ensure correct θ-dependence

We modify the model architecture to enable the correct θ-dependence:



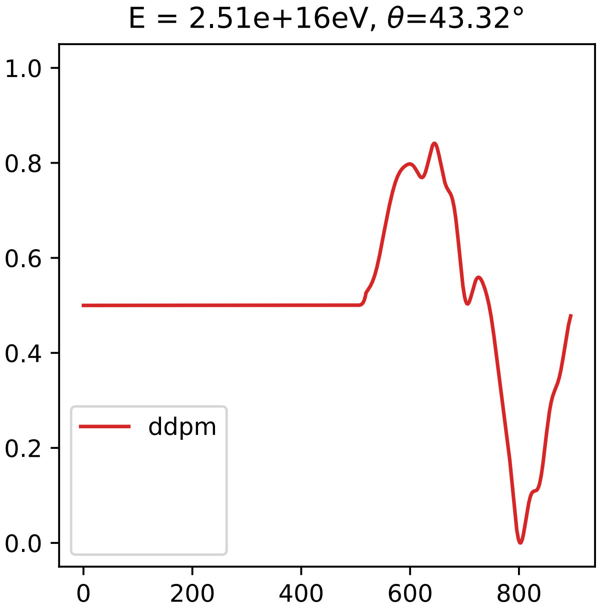
- The DDPM generates samples at a fixed angle θ_0.
- A separate network (the θ-Net) transforms these samples into the corresponding signals at arbitrary angles θ.

Angle modification network (θ-Net)

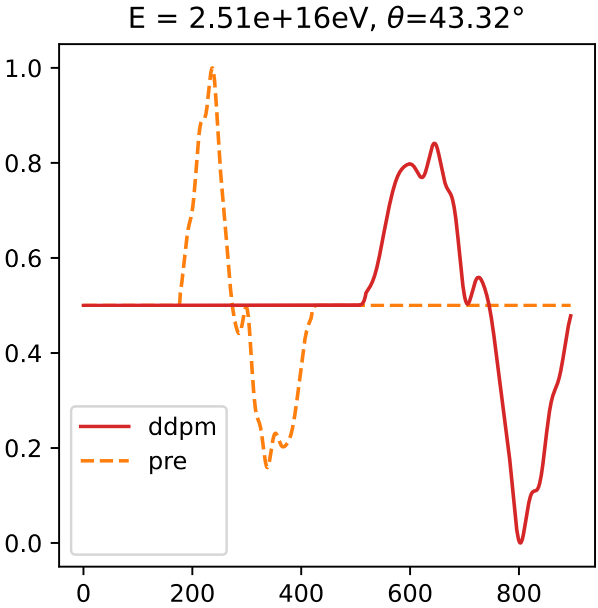


- Samples at the fixed angle $\theta_0 = 75.32^\circ$ serve as input.
- The samples are preprocessed via a loose geometric relationship.
- A convolutional network finetunes the transformed signals.

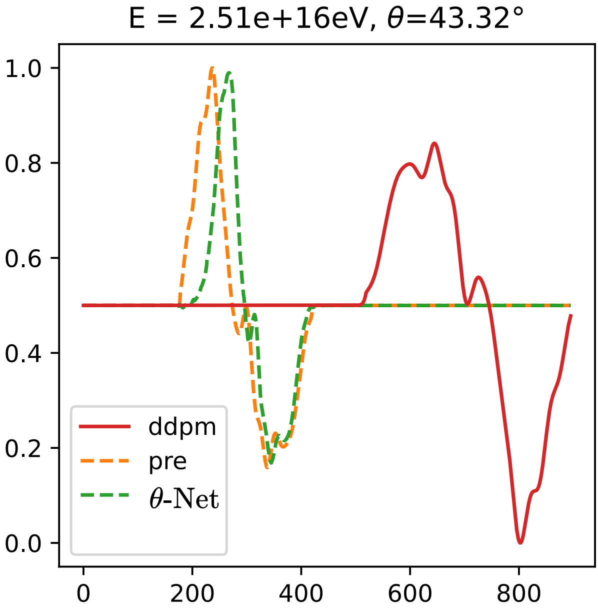
Generating samples using the θ-Net



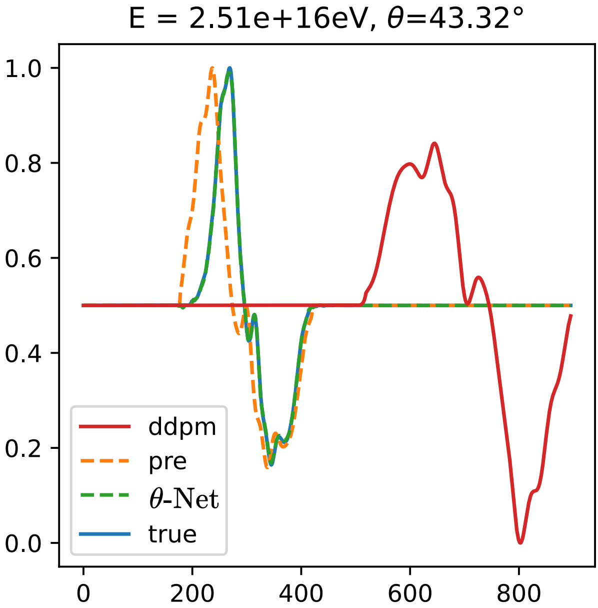
Generating samples using the θ-Net



Generating samples using the θ-Net

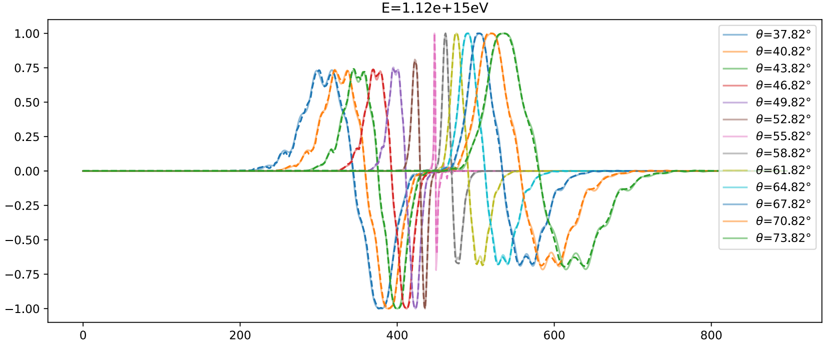


Generating samples using the θ-Net



Results: θ-dependence with θ-Net

- Using the θ-Net, the correct θ-dependence is recovered with high accuracy.



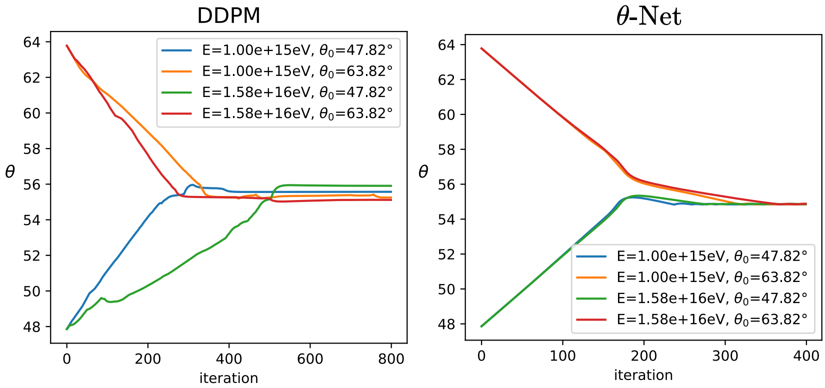
Simple optimization experiment

- Test whether the backpropagation through the networks yields useful gradients.
- We optimize the viewing angle θ to obtain (normalized) signals that are as 'squeezed' as possible.
- That is, we minimize

$$A(x_0(\theta)) = \int_{-\infty}^{\infty} |x_0(t,\theta)| dt \tag{1}$$

with respect to θ.

Results: simple optimization experiment



- Reasonable results with both architectures.
- The model with θ-net converges faster and the results are more accurate.
- The VRAM requirements are significantly lower with the θ-Net architecture (1GB vs 20GB).

Summary and outlook

Summary:

- DDPMs can generate realistic radio signals
- Modular model design improves the results
    - deal with wide range of amplitudes
    - ensure the correct angle-dependence
- Gradients are suitable for optimization

Summary and outlook

Summary:

- DDPMs can generate realistic radio signals
- Modular model design improves the results
  - deal with wide range of amplitudes
  - ensure the correct angle-dependence
- Gradients are suitable for optimization

Outlook:

- Further model tuning to obtain good results at all energies
- Reduce overfit to existing data
- Combine with other elements of optimization pipeline

Summary and outlook

Summary:

- DDPMs can generate realistic radio signals
- Modular model design improves the results
  - deal with wide range of amplitudes
  - ensure the correct angle-dependence
- Gradients are suitable for optimization

Outlook:

- Further model tuning to obtain good results at all energies
- Reduce overfit to existing data
- Combine with other elements of optimization pipeline

Thanks for your attention!