Increasing sparse detector data using generative adversarial networks

R. Taylor, A. Vacheret

Imperial College London

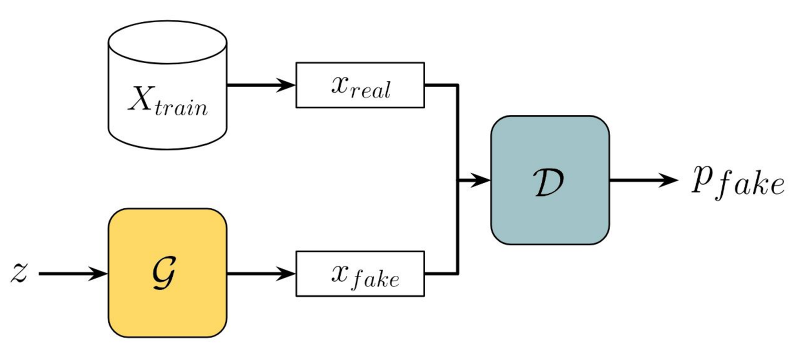
NuSec Nuclear Security Science Network

Introduction

- Large datasets are routinely required to accurately quantify the performance of radiation detector systems and subsequent analysis algorithms
- Low signal to noise or limited access to radiation sources can limit the useful data that can be acquired, in which case detailed Monte Carlo simulations may be necessary
- Generative adversarial networks (GANs) have the potential to be a fast alternative to Monte Carlo simulation, with the ability to generate a realistic detector response

Generative Adversarial Networks

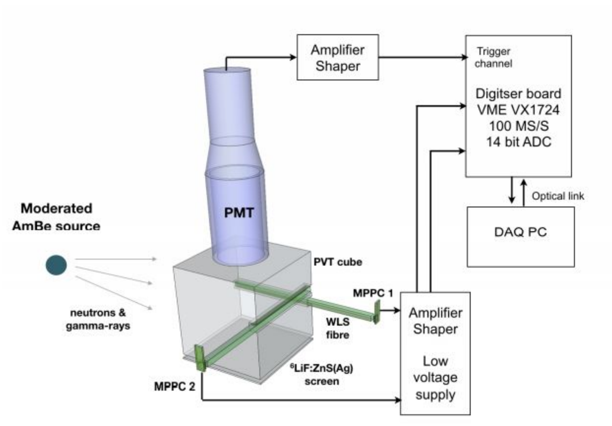
- Generative adversarial networks are a class of machine learning algorithm that learn to generate realistic high-dimensional data
- GANs consist of two neural networks:
  - A ***generator*** that creates fake data from random noise
  - A ***discriminator*** that learns to correctly classify real and fake data

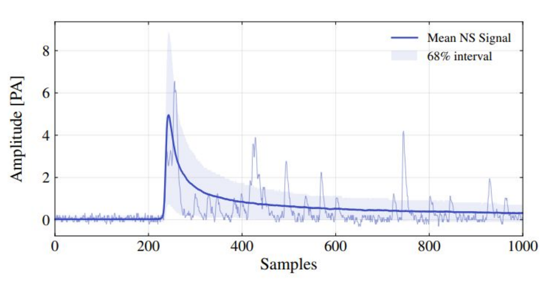


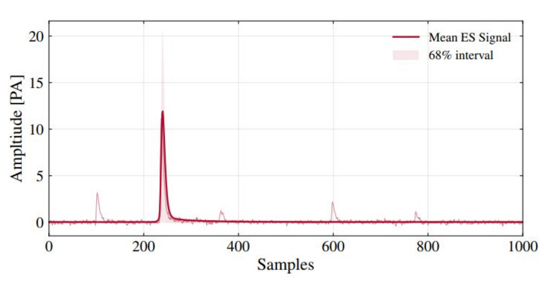
- During training, the generator learns to create realistic data by fooling the discriminator
- The generator is optimised when the discriminator is unable to tell the difference between real and fake samples

Training Data

- Training data was collected by exposing a PVT scintillator with a ^{6}LiF:ZnS(Ag) neutron absorbing layer to an AmBe source [1]







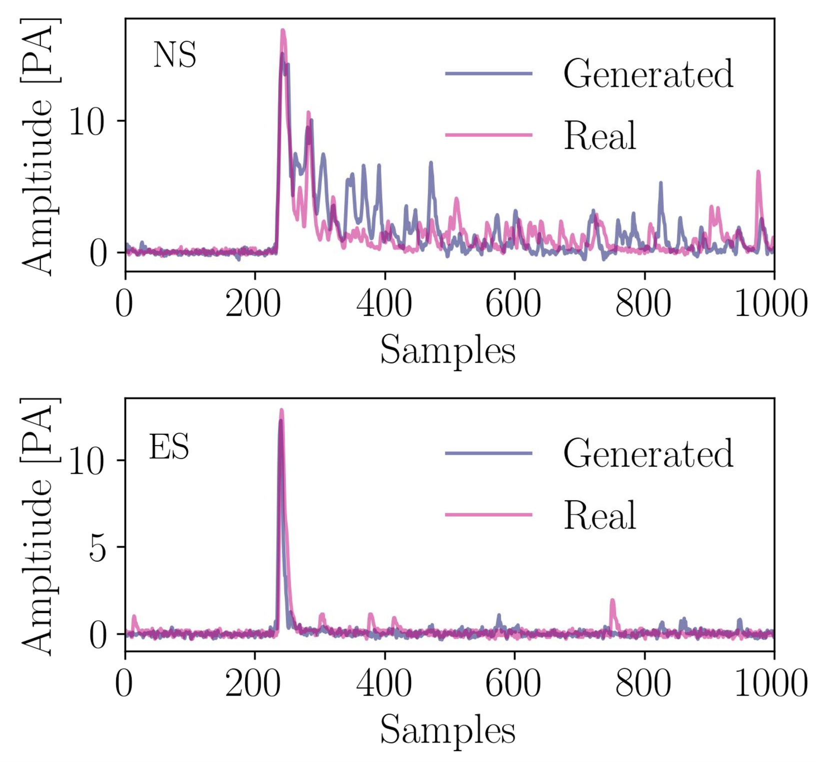
- Two distinct signals:
  a. ***Nuclear Scintillation (NS)***: Long-tailed pulse with multiple subsequent peaks
  b. ***Electron Scintillation (ES):*** prompt pulse with single peak
- Approximately 6000 digitised waveforms of each signal type were recorded

References & Acknowledgements

[1] J. Griffiths, S. Kleinegesse, D. Saunders, R. Taylor, A. Vacheret arXiv:1807.06853
[2] I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, A. Courville arXiv:1704.00028

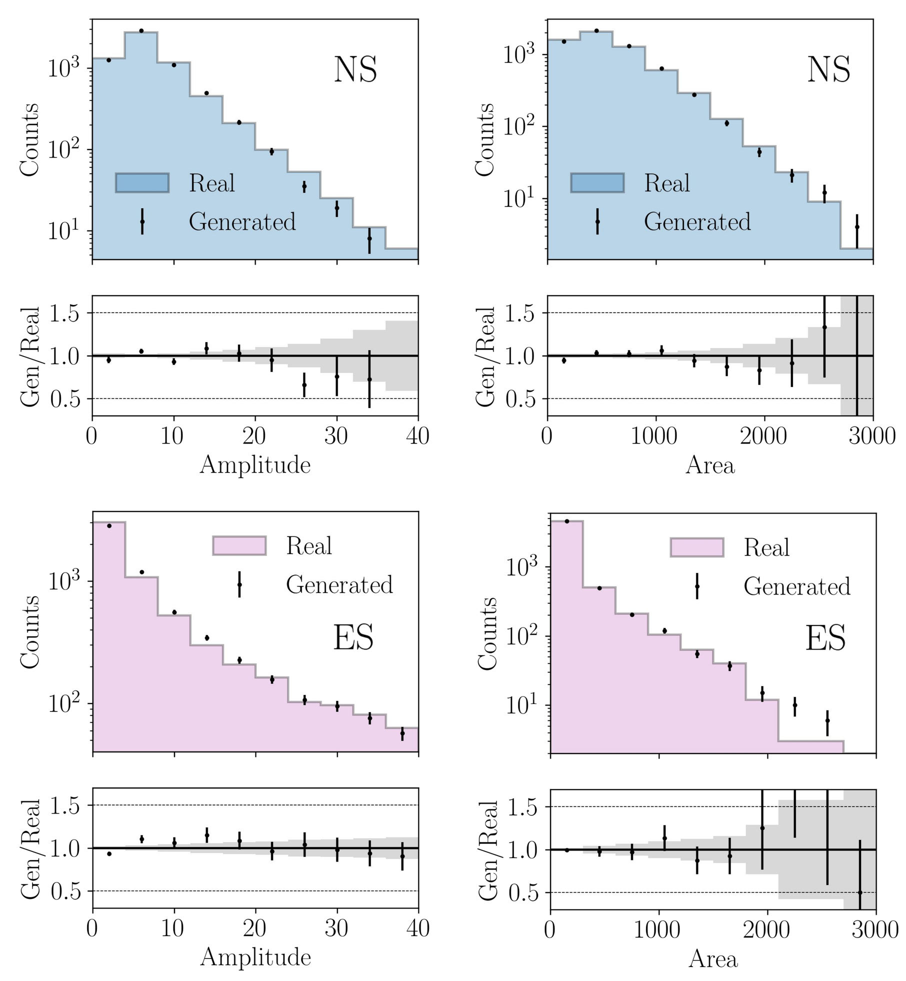
- Thanks to T. Delahaye and A. Ghataura for the initial development of the GAN architecture

Generated Pulses

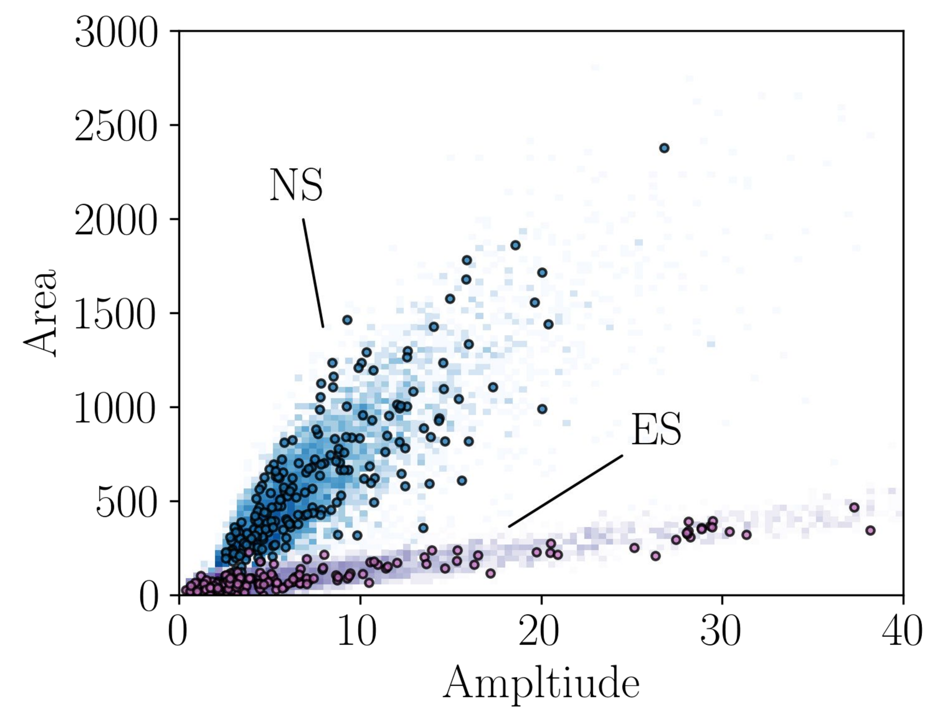


- ***Wasserstein GAN*** [2] architecture with gradient-penalty loss trained on NVIDIA K40 Tesla GPUs
- Separate models trained to generate ES and NS signals
- Individual generated pulses look realistic and the pulse shape distributions are accurately reproduced

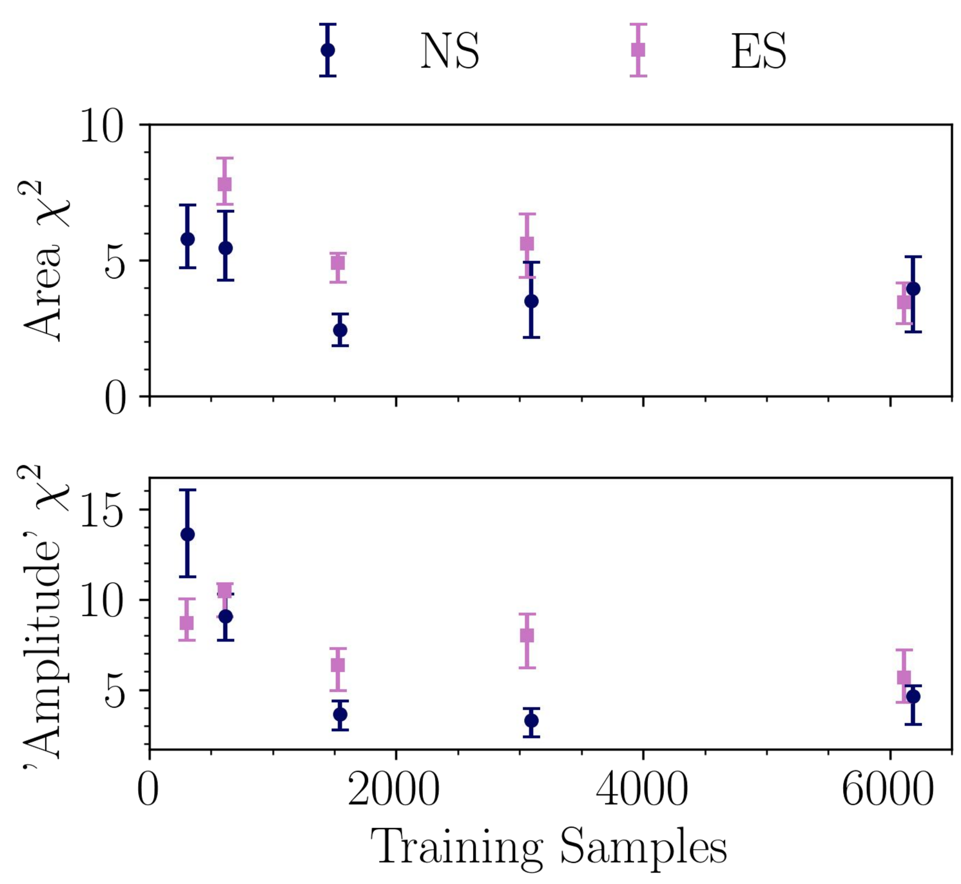
Pulse Parameter Distributions



- Variance and correlation of distributions is accurately modelled
- GAN quality measured with χ^2 between real and generated pulse distributions
- Reasonable agreement achieved for wide range of pulses areas and amplitudes



Impact of Dataset Size



- Improvement in quality of GAN seen with increasing dataset size
- However, significantly smaller datasets can still produce reasonable results.
- Performance becomes limited by ability to model the tails of training distributions

Conclusions & Future Work

- GAN architecture developed that can produce artificial pulse level data for two distinct and varied scintillation signals
- Demonstrated that the GAN can approximate pulse shape distributions, even with a limited dataset
- Further work needed to investigate different GAN architectures:
  - ***Conditional GAN*** - generator can be conditioned to produce a specific output, e.g interaction type, energy
- Study the generation of a more complex detector response, e.g with multiple channels