LUDWIG-
MAXIMILIANS-
UNIVERSITY
MUNICH

Generating PXD Background Hitmaps with Generative Adversarial Networks at Belle II

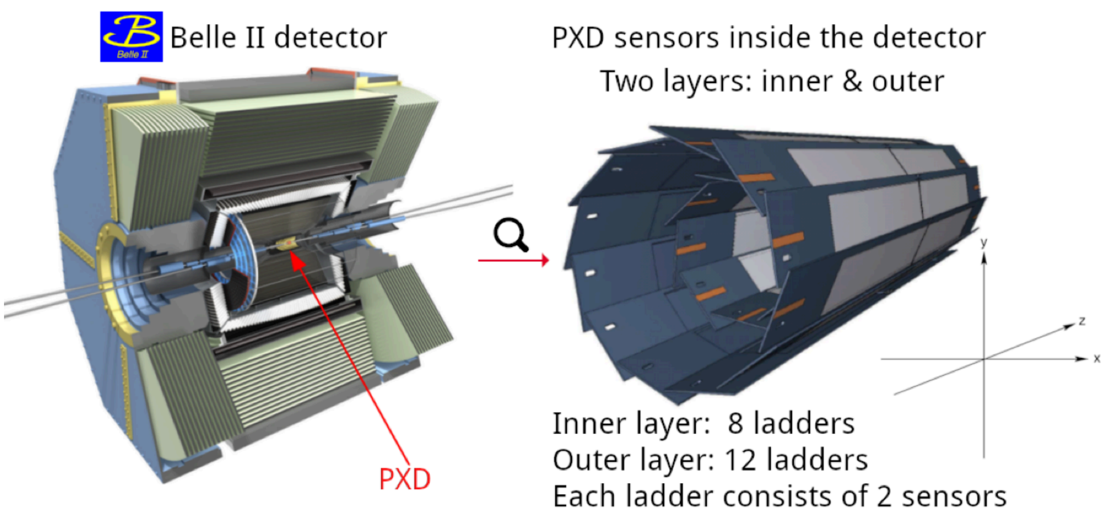
Hosein Hashemi, Nikolai Hartmann, Thomas Kuhr, Martin Ritter, Matej Srebre

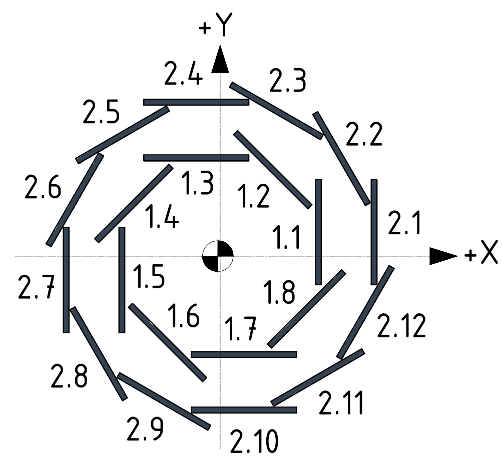
Ludwig-Maximilians-Universität München
The ORIGINS Excellence Cluster

Bundesministerium für Bildung und Forschung

Introduction

- ***The Pixel Vertex Detector (PXD)*** is the innermost semi-conductor sub-detector at Belle II.
- The sensitive area of the PXD is assembled from ***40 modules***, where each module consists of a ***250*** $\times$ ***768*** pixel matrix of the pixel sensors.
- **The inner layer:** 16 modules implemented into 8 ladders
- **The outer layer:** 24 modules implemented into 12 ladders

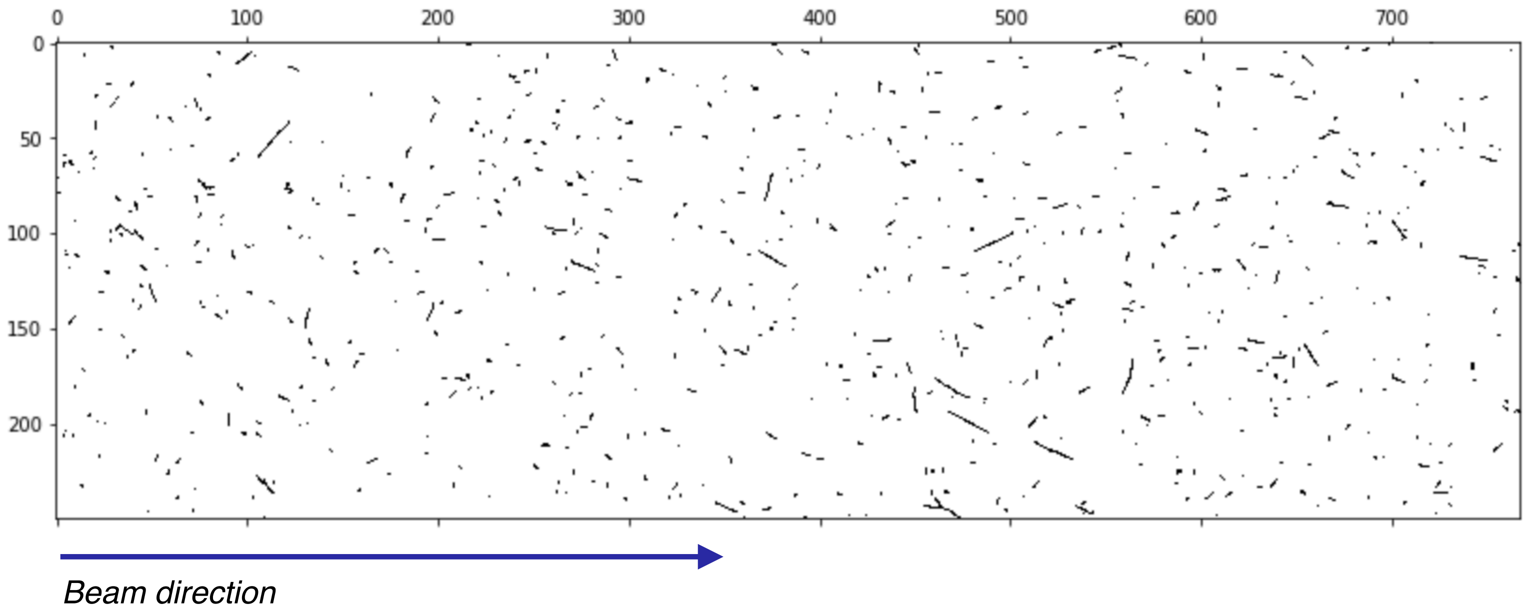




Backgrounds

- ***The PXD hits are coming from two sources:***
  - **Signal Decays:**
    - A. *particles which originate from the physics processes of interest.*
  - **Backgrounds:**
    - A. ***Beam-induced****: intra-beam scattering, Beam-Gas scattering, synchrotron radiation*
    - B. ***Luminosity dependent:*** *Radiative Bhabha scattering, two-photon process*

Colour-reversed real (simulated) image



Backgrounds

- **Realistic detector simulation has to take into account effects from background processes**
  - *Simulation requires many PXD hitmaps with statistically independent background.*
  - *Overlay hits from simulated background or random trigger data to hits from signal MC.*
  - *PXD hits have the highest storage consumption, almost 200 kB per event cost.*
  - *Requires distributing over all sites where MC is produced.*

Backgrounds

- **Realistic detector simulation has to take into account effects from background processes**
  - *Simulation requires many PXD hitmaps with statistically independent background.*
  - *Overlay hits from simulated background or random trigger data to hits from signal MC.*
  - *PXD hits h... ...ge consumption, almost 200 kB per event cost.*
  - *Re... ...re MC is produced.*

Solution:
Let's generate ze bkg on ze way of analysis with Generative Adversarial Networks, instead of storing them.

■ ***What Is Adversarial?***

Evolve with competition

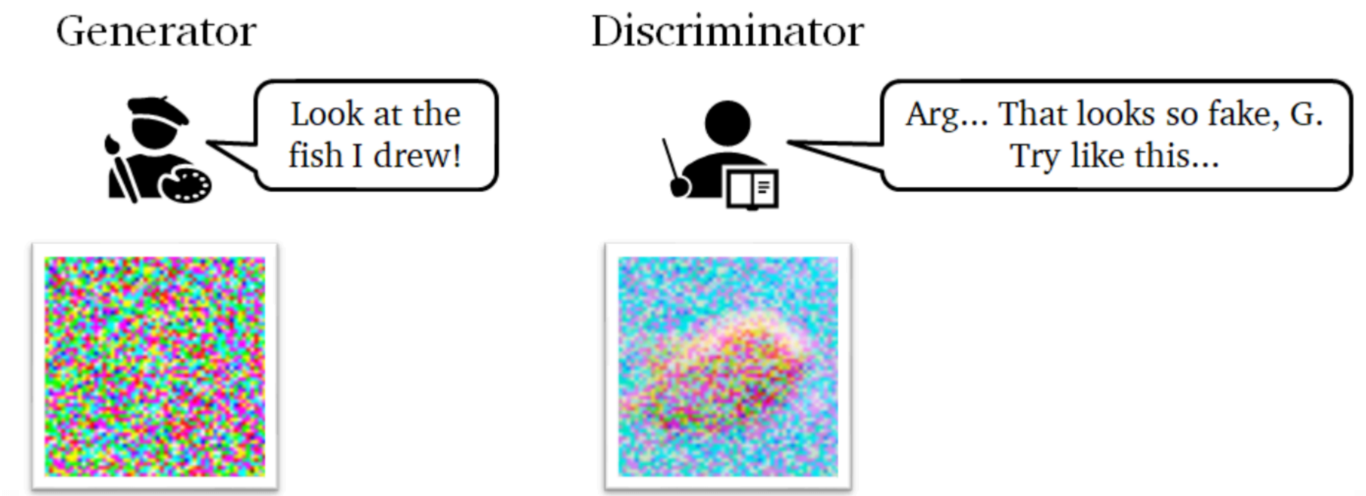
Deer-leopard minimax game

$$\min_{\text{leopard}} \max_{\text{deer}} V(\text{deer}, \text{leopard}) = \text{distance between deer and leopard}$$

What Doesn't Kill You Makes You Stronger!

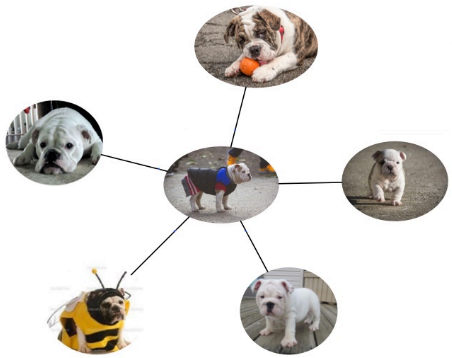
GAN is all you need!

- Generate PXD background events with Generative Adversarial Network **(GAN)**
- **Whats is GAN?**



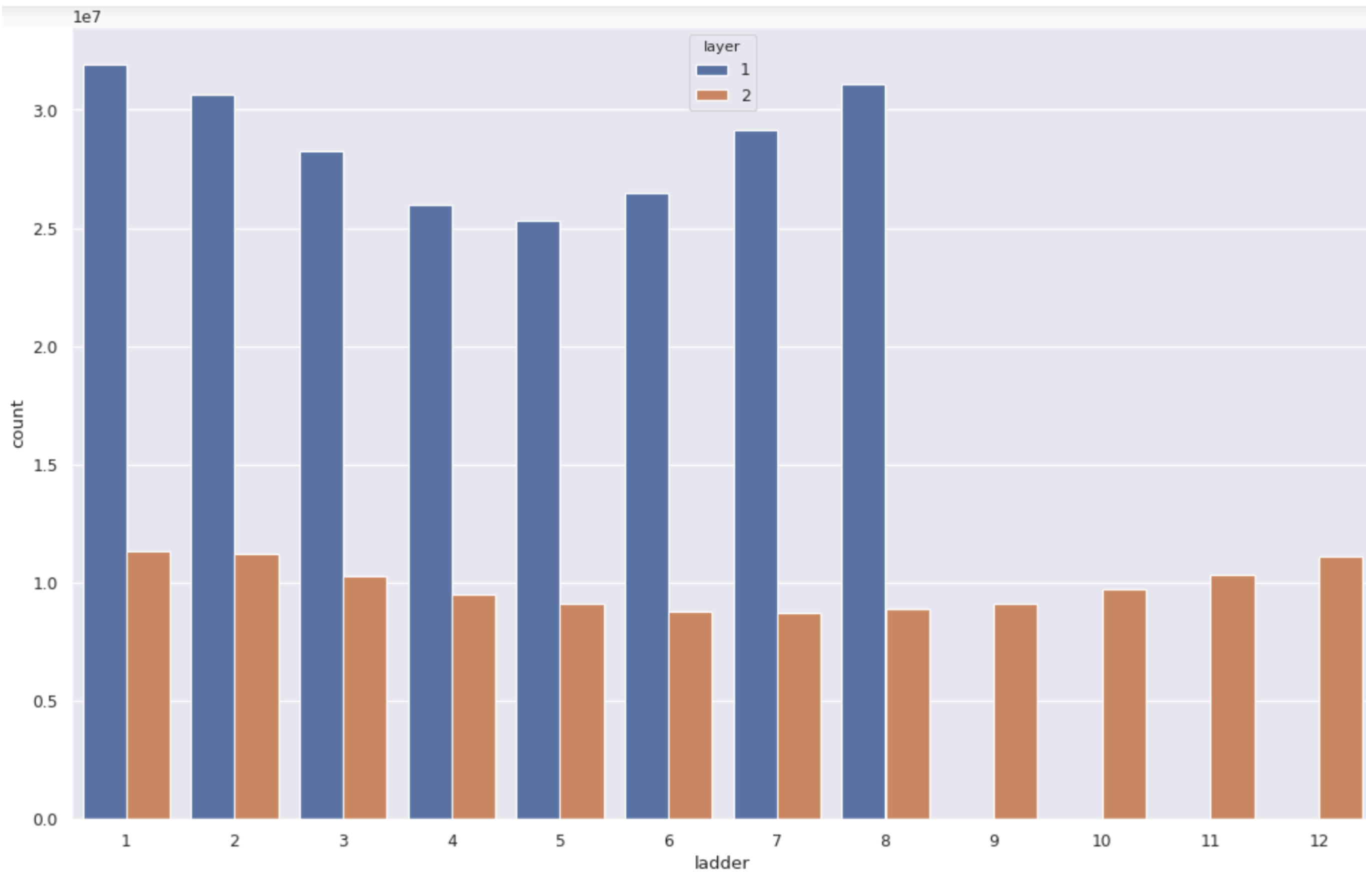
The GAN game

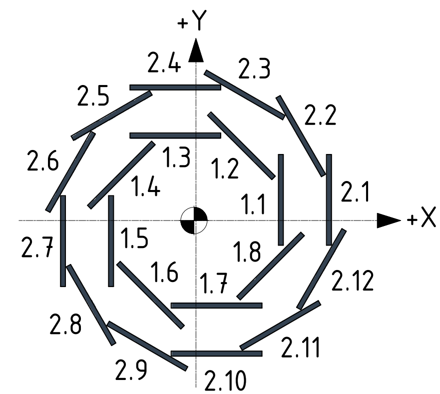
- **Conditional GAN :** *The type of animal is the condition*
- ***Close-Conditional (relational)***



Conditional GAN

- **Using spatial class-conditions based on the sensor number 1-40:**
  - ☑ *Increase the image fidelity*
  - ☑ *Generation of sensor-dependent images*

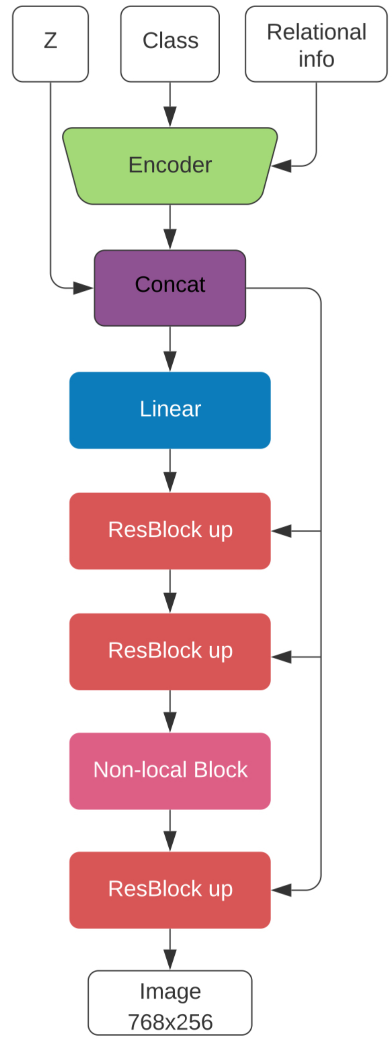




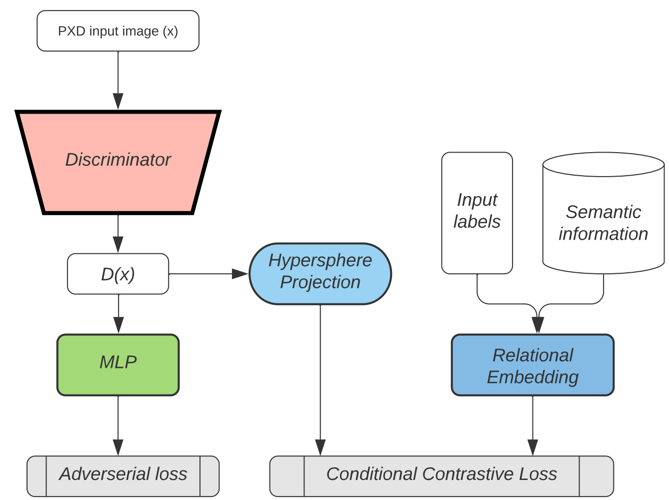
Model

Modified BigGAN-deep
with Contrastive
(metric) Learning

- **Technologies:**
  - Self-Attention Block
  - Residual blocks
  - Spectral Normalisation
  - Orthogonal Weight init.
  - Orthogonal regularisation
  - Contrastive Learning
  - Hinge Loss
  - Consistency Regularisation



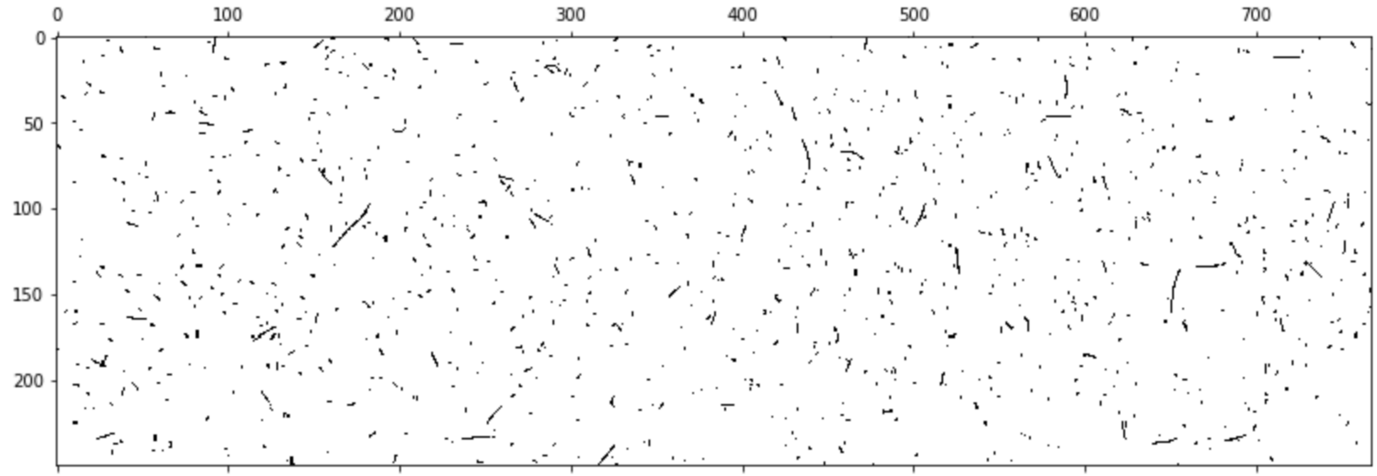
Generator



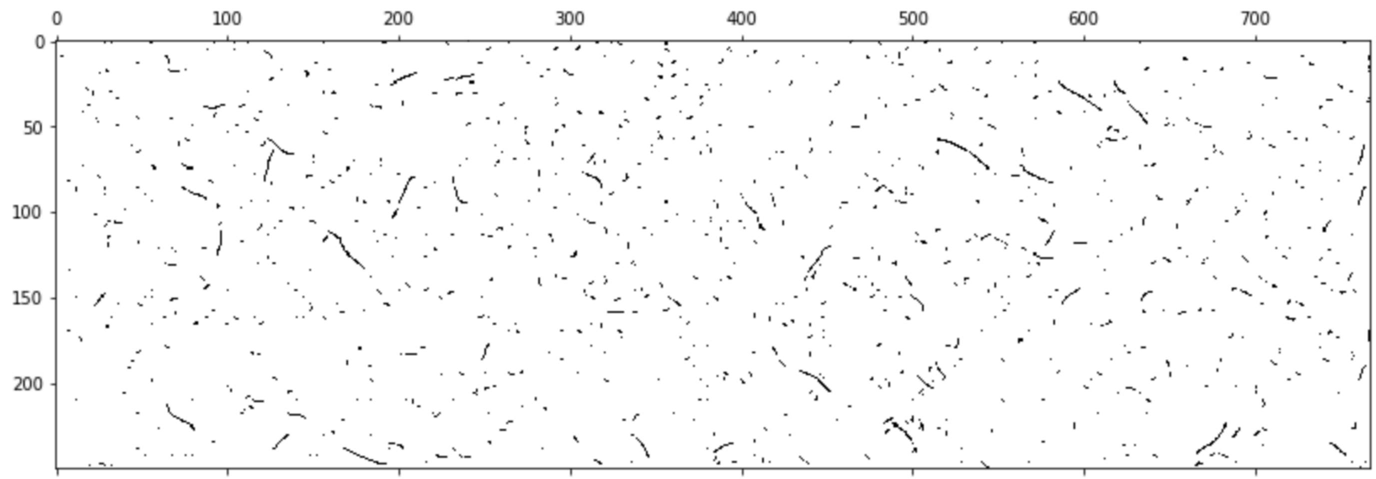
Discriminator

Generated vs Real PXD Images

Colour-reversed real (simulated) image



Colour-reversed generated image



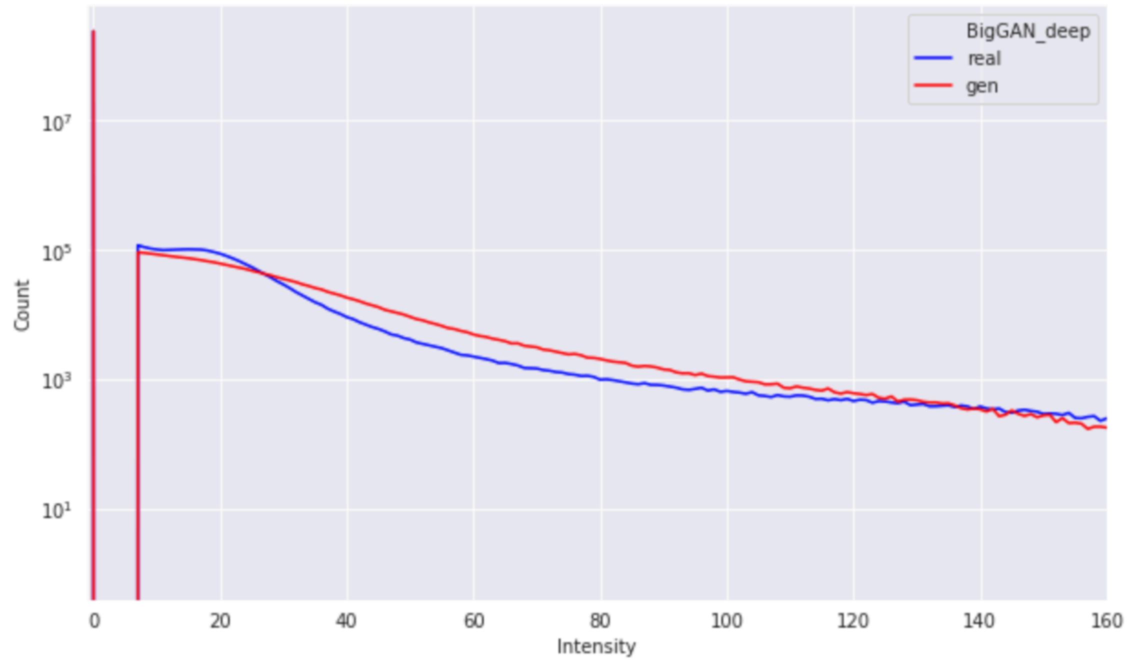
Validation of generated PXD images

Problem:

Finding a metric to say how good the generated images are.

Solutions:

- Image pixel intensity analysis:
  - Pixel value 0 means complete blackness
  - The fired pixels are only read out if their value exceeds a threshold, 7 ADU.
  - How to capture this prior information about the image?



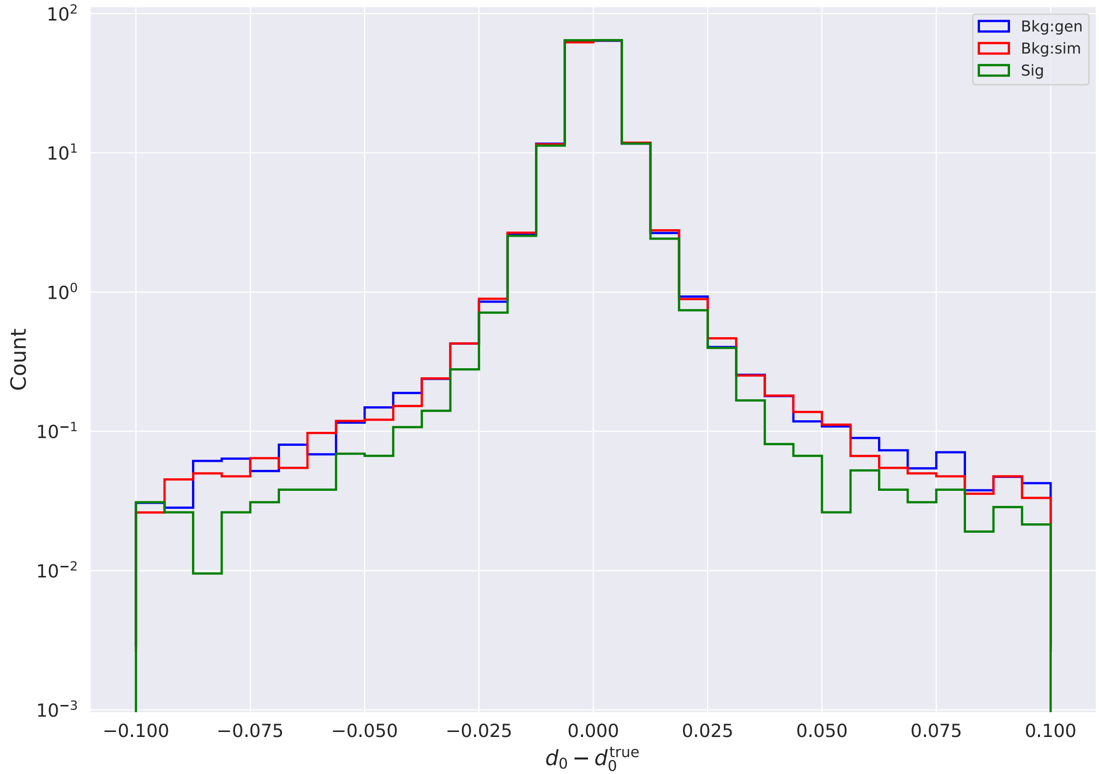
Validation of generated PXD images

Problem:

Finding a metric to say how good the generated images are.

Solutions:

- ☑ Evaluate tracking performance impact parameter residuals:
  - Signal + no bkg.
  - Signal + simulated bkg.
  - Signal + generated bkg.



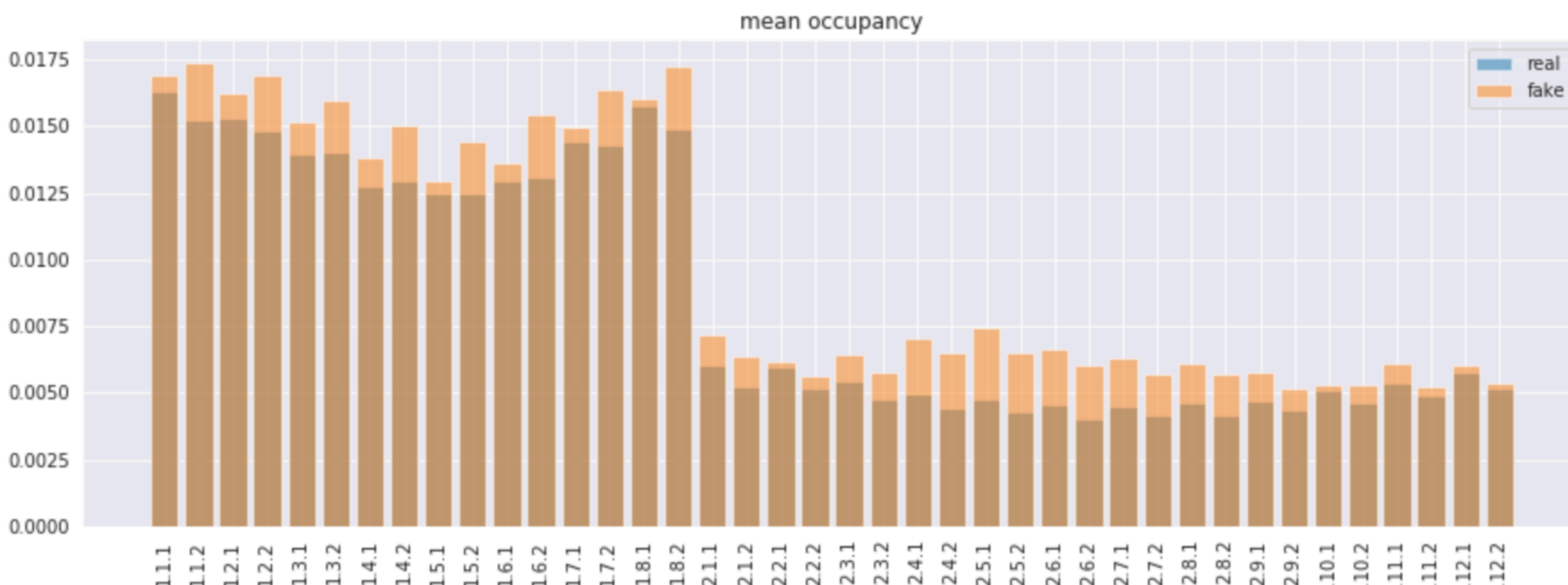
Validation of generated PXD images

Problem:

Finding a metric to say how good the generated images are.

Solutions:

☑ Compare the occupancy information for PXD



Summary and Outlook

- Successful proof of principle that conditional GANs can be used to generate sensor-dependent PXD background
- Refine the GAN setup in order to capture **correlation** between two layers or each sensor relative to each other for PXD detector.
- Adding bkg types as colour channels to the images.
- Create a custom **Inception Score (IS)**, based on simulated events in order to have a fully automated evaluation metric.
- Doing a comprehensive validation of generated hitmaps by estimating the systematic uncertainty on the tracking efficiency, fake rate and resolution.
- Simulation Software implementation.

Thank You

Back up Slides

- The fired pixels are only read out if their value exceeds a threshold, 7 ADU.
- How to capture this prior information about the image?
- Solution: To add this information to the training

Algorithm 1 Pixel-Aware regularization.

Input: generator and discriminator parameters θ_G, θ_D, pixel-aware regularization coefcient λ, Adam hyperparameters α, β_1 , β_2, batch size M, number of discriminator iterations per generator iteration N_D

```
for number of training iterations do
    for t = 1, ..., N_D do
        for i = 1, ..., M do
            sample z ~ p(z),  x = p_data(x)
            L_D^(i) ← D[G(z)] − D(x)
        end for
        θ_D ← Adam((1/M) Σ_{i=1}^{M} (L_D^(i)), α, β_1, β_2)
    end for
    sample {z^(i)}_{i=1}^{M} ~ p(z)
    x_fake = G(z)
    F[G(z)] : x_fake ↦ x_fake^cutoff          ▷ Threshold wrt. the pixel constraints.
    L_pr^(i) ← ||G(z) − F[G(z)]||^2
    L_G^(i) ← −D[G(z)]
    θ_G ← Adam((1/M) Σ_{i=1}^{M} (L_G^(i) + L_pr), α, β_1, β_2)
end for
```

Where:

$L_D^{(i)} \leftarrow D[G(z)] - D(x)$

$\theta_D \leftarrow Adam(\frac{1}{M}\sum_{i=1}^{M}(L_D^{(i)}), \alpha, \beta_1, \beta_2)$

$F[G(z)] : x_{fake} \mapsto x_{fake}^{cutoff}$

$L_{pr}^{(i)} \leftarrow ||G(z) - F[G(z)]||^2$

$\theta_G \leftarrow Adam(\frac{1}{M}\sum_{i=1}^{M}(L_G^{(i)} + L_{pr}), \alpha, \beta_1, \beta_2)$

Back up Slides

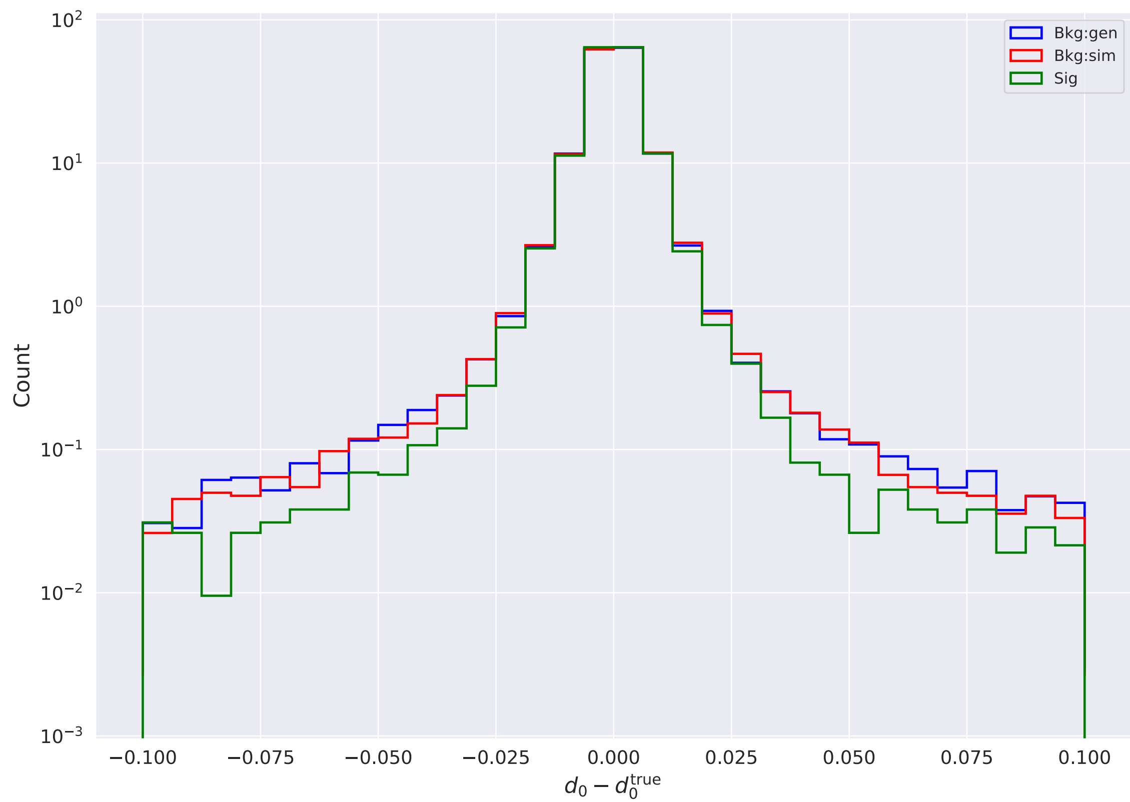
- Evaluate tracking performance for
  - Signal + no bkg.
  - Signal + nominal bkg.
  - Signal + generated bkg.

- **Scoring**: *Using Frechet Distance (2-Wasserstein distance):*

$$W_2(\mu_1, \mu_2)^2 = \|m_1 - m_2\|_2^2 + \mathrm{trace}\left(C_1 + C_2 - 2\left(C_2^{1/2} C_1 C_2^{1/2}\right)^{1/2}\right).$$

for $\mu_1 = N(m_1, C_1)$ and $\mu_2 = N(m_2, C_2)$.

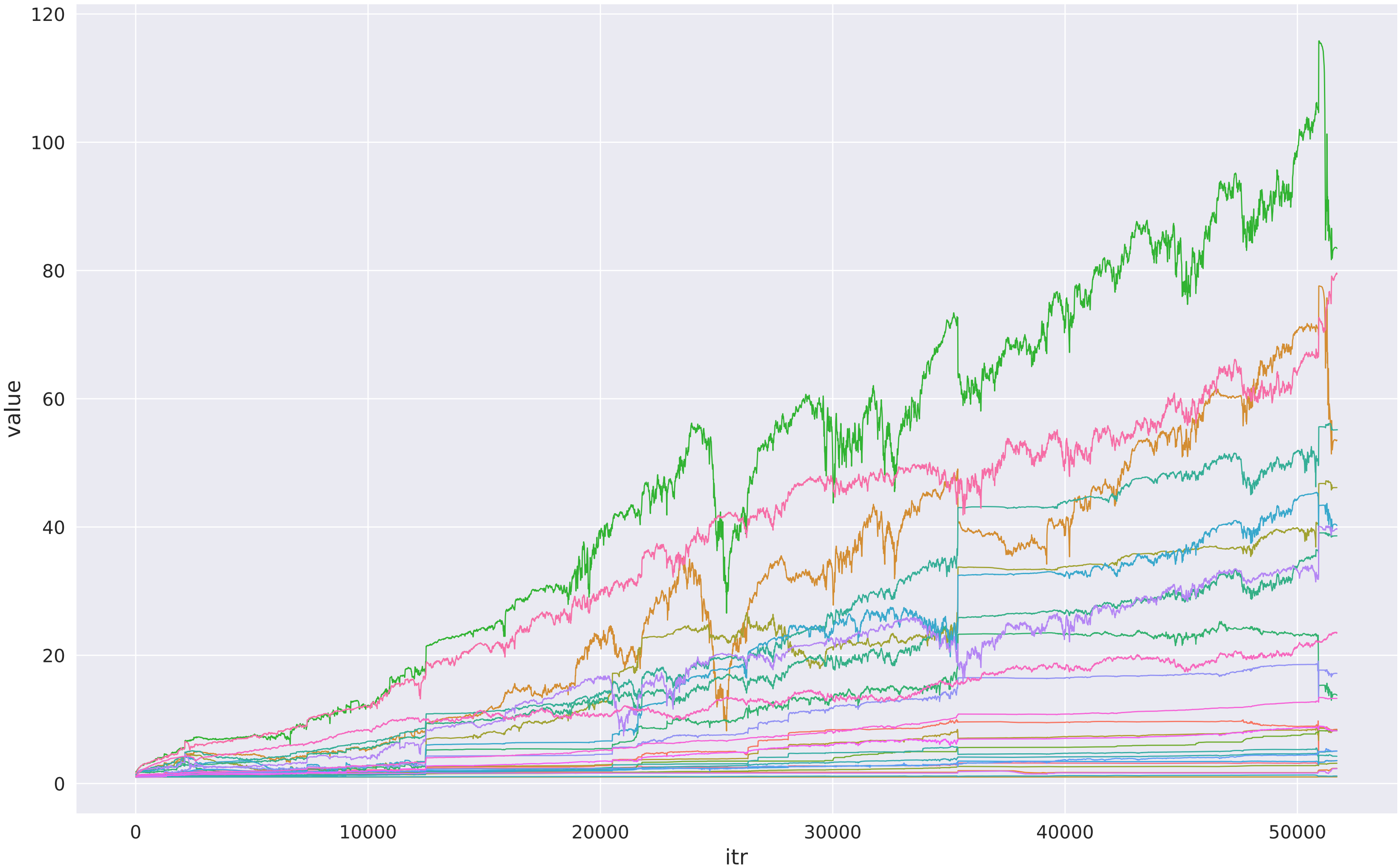
The Lower the FD score, the better the image quality and diversity from the physics point of view.



FD scores of between:

A. sim-sig : 5.42e-4
B. gen-sig : 7.41e-4
C. gen-sig : 1.64e-5

An Example of Collapse



A Healthy Model