

Introduction
OO

GAN
OOOO

Results
OOOOOOO

Outlook
O

How to GAN LHC Events

Anja Butter

ITP, Universität Heidelberg

arXiv:1907.03764

with Tilman Plehn and Ramon Winterhalder



Building a ML Toolbox

Making use of the ML hype!

Some examples from Hammer and Nails 2019:

- Jet classification/ top tagging - CNNs, point clouds, customized layers
- Track reconstruction - Graph networks
- Anomaly detection - Autoencoder
- Shower generator - GANs
- Pileup removal - Graph networks
- Systematic uncertainties - Bayesian networks
- ...



Phase-Space Sampling

Monte Carlo simulations at the heart of any LHC analysis

Phase-Space Sampling

Monte Carlo simulations at the heart of any LHC analysis

Problem: High-dimensionality and rich phase-space structures

Task: Finding an optimal phase-space mapping

→ Computationally time consuming

Phase-Space Sampling

Monte Carlo simulations at the heart of any LHC analysis

Problem: High-dimensionality and rich phase-space structures

Task: Finding an optimal phase-space mapping

→ Computationally time consuming

How to generate events more efficiently?

→ Neural networks!

Neural Network Approach

Task: Generate events with a neural network (generator)

Direct comparison to data → **produce unweighted events**

Standard MC: unweighting algorithm needed → **inefficient**

Neural Network Approach

Task: Generate events with a neural network (generator)

Direct comparison to data → **produce unweighted events**

Standard MC: unweighting algorithm needed → **inefficient**

NN approach

- Input: random numbers
- Output: unweighted events
- Training data:
 - unweighted MC events or real data
 - can include parton showers, hadronization and detector effects

Neural Network Approach

Task: Generate events with a neural network (generator)

Direct comparison to data → **produce unweighted events**

Standard MC: unweighting algorithm needed → **inefficient**

NN approach

- Input: random numbers
- Output: unweighted events
- Training data:
 - unweighted MC events or real data
 - can include parton showers, hadronization and detector effects

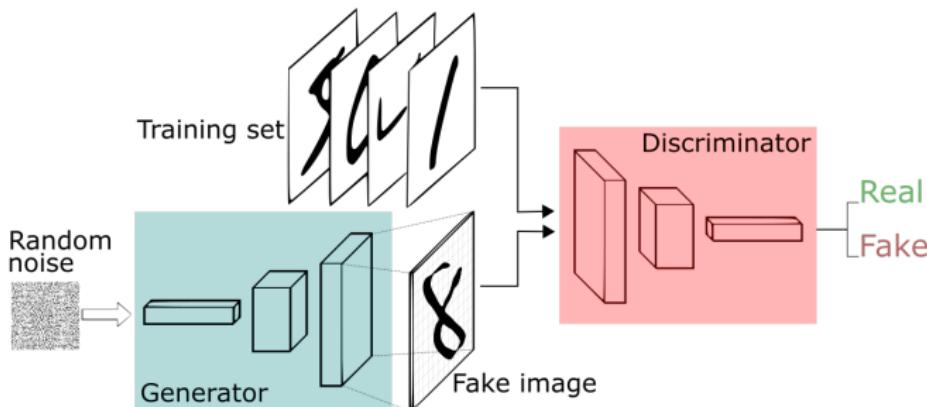
Network architecture? → **GAN**

Generative Adversarial Networks

GAN: two competing networks → generator and discriminator

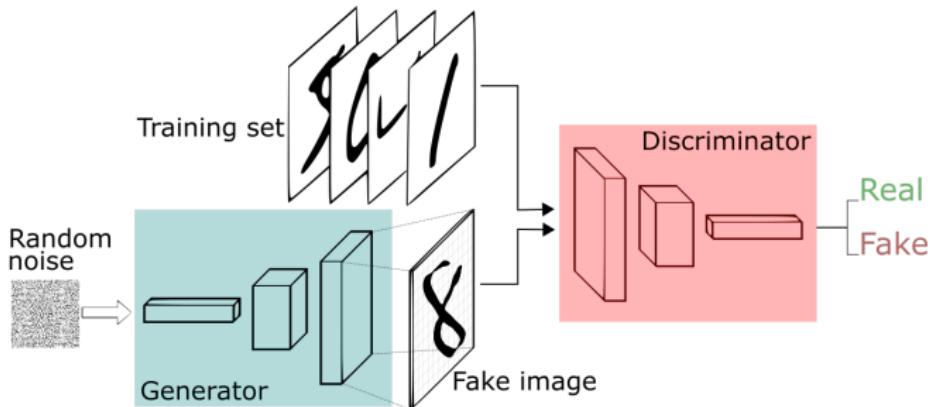
Generative Adversarial Networks

GAN: two competing networks → generator and discriminator



Generative Adversarial Networks

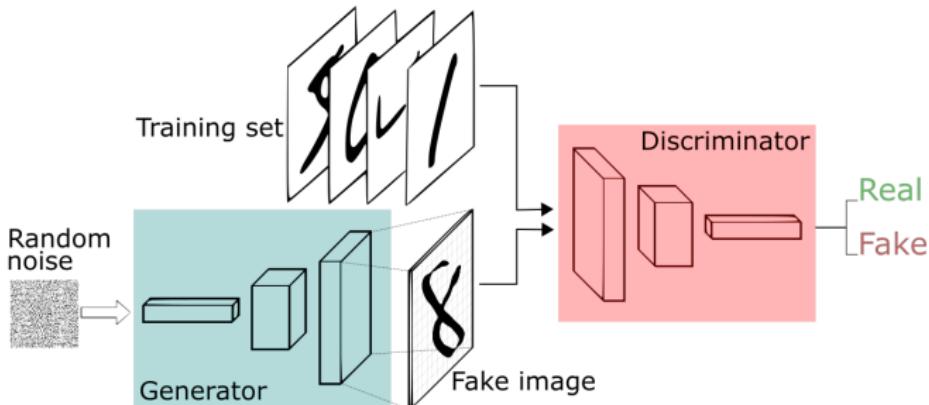
GAN: two competing networks → generator and discriminator



GANs used in many applications like video and image generation and physics.

Generative Adversarial Networks

GAN: two competing networks → generator and discriminator

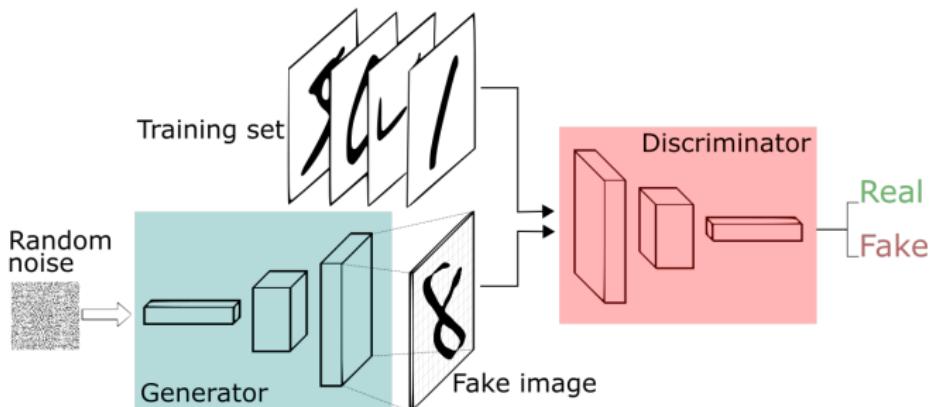


GANs used in many applications like video and image generation and physics.

- Particle Detectors - Paganini et al. [arXiv:1705.02355], Paganini et al. [arXiv:1712.10321]
Musella et al. [arXiv:1805.00850], Erdmann et al. [arXiv:1807.01954]

Generative Adversarial Networks

GAN: two competing networks → generator and discriminator

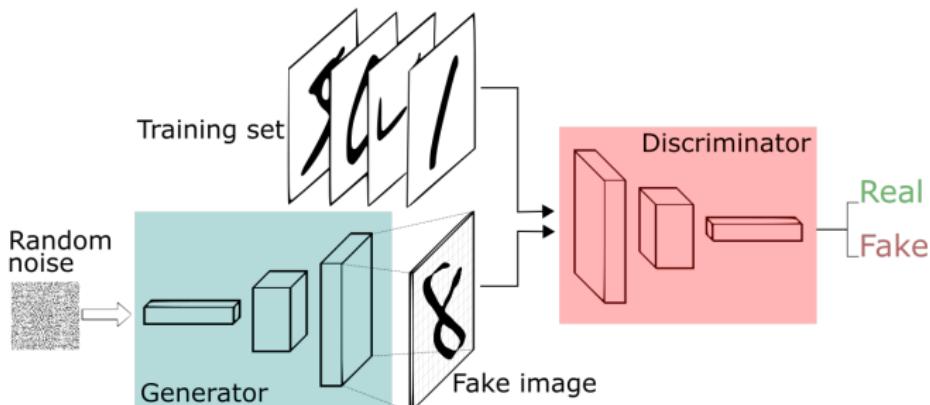


GANs used in many applications like video and image generation and physics.

- Particle Detectors - Paganini et al. [arXiv:1705.02355], Paganini et al. [arXiv:1712.10321]
Musella et al. [arXiv:1805.00850], Erdmann et al. [arXiv:1807.01954]
- Monte Carlo integration - Bendavid [arXiv:1707.00028], Klimek et al. [arXiv:1810.11509]

Generative Adversarial Networks

GAN: two competing networks → generator and discriminator

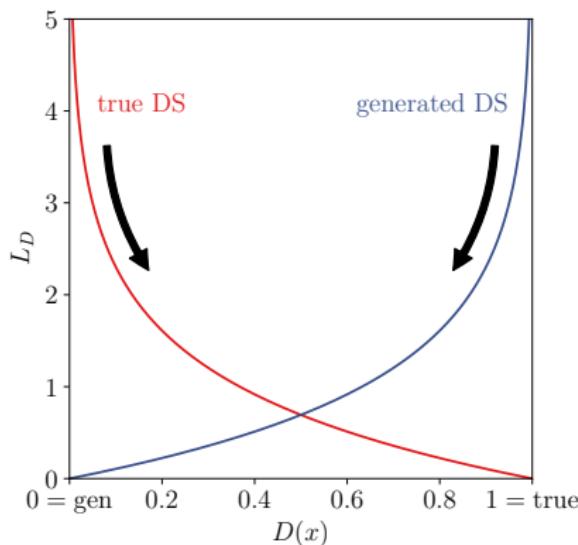


GANs used in many applications like video and image generation and physics.

- Particle Detectors - Paganini et al. [arXiv:1705.02355], Paganini et al. [arXiv:1712.10321]
Musella et al. [arXiv:1805.00850], Erdmann et al. [arXiv:1807.01954]
- Monte Carlo integration - Bendavid [arXiv:1707.00028], Klimek et al. [arXiv:1810.11509]
- Event generation - Otten et al. [arXiv:1901.00875], Hashemi et al. [arXiv:1901.05282]

Training the Discriminator

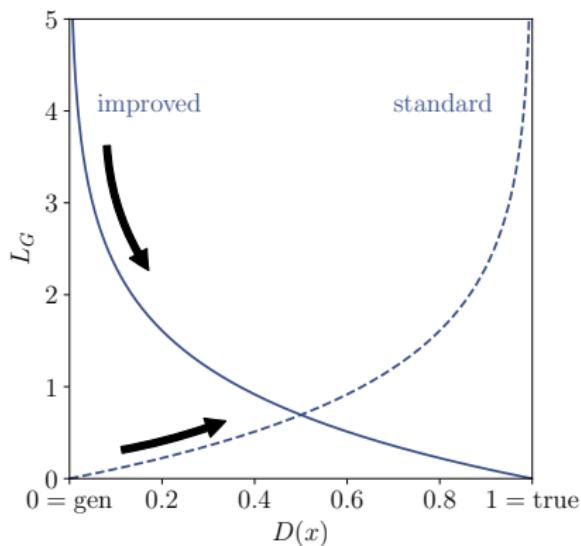
Discriminator loss



$$L_D = \langle -\log D(x) \rangle_{x \sim P_T} + \langle -\log(1 - D(x)) \rangle_{x \sim P_G}$$

Training the Generator

Generator loss

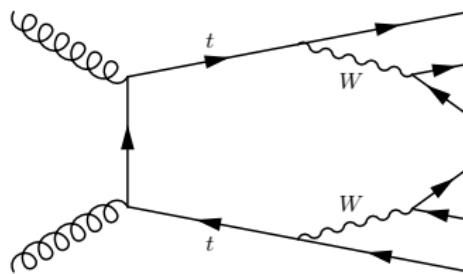


$$L_G = \langle -\log D(x) \rangle_{x \sim P_G}$$

Top-Pair Production

GAN events for the $2 \rightarrow 6$ particle production process

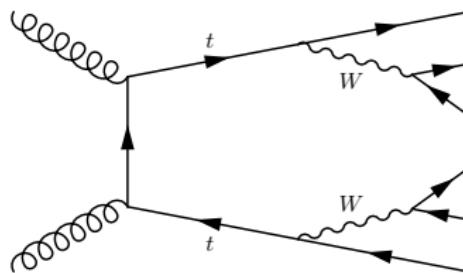
$$pp \rightarrow t\bar{t} \rightarrow (bW^-)(\bar{b}W^+) \rightarrow (bq_1\bar{q}'_1)(\bar{b}q_2\bar{q}'_2).$$



Top-Pair Production

GAN events for the $2 \rightarrow 6$ particle production process

$$pp \rightarrow t\bar{t} \rightarrow (bW^-)(\bar{b}W^+) \rightarrow (bq_1\bar{q}'_1)(\bar{b}q_2\bar{q}'_2).$$

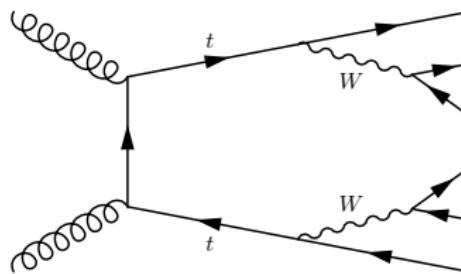


Challenges: 16-dimensional phase-space, 4 resonances,
phase-space boundaries, tails

Top-Pair Production

GAN events for the $2 \rightarrow 6$ particle production process

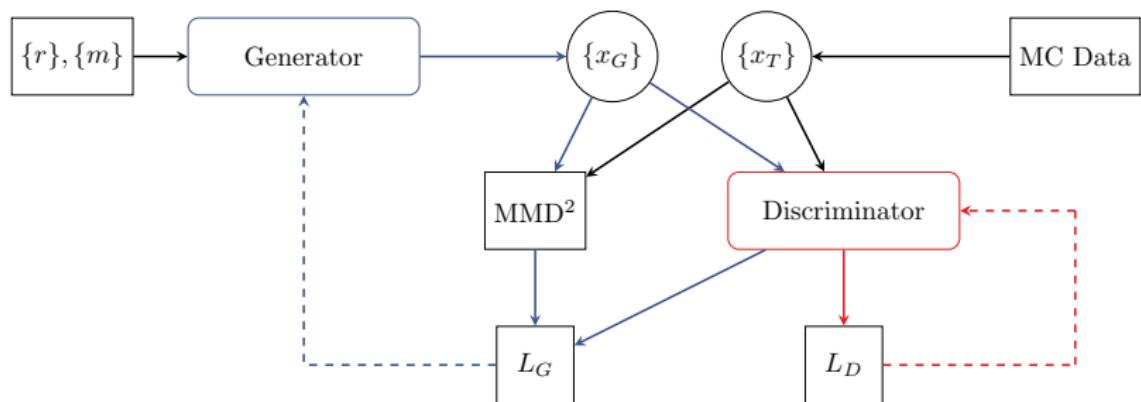
$$pp \rightarrow t\bar{t} \rightarrow (bW^-)(\bar{b}W^+) \rightarrow (bq_1\bar{q}'_1)(\bar{b}q_2\bar{q}'_2).$$



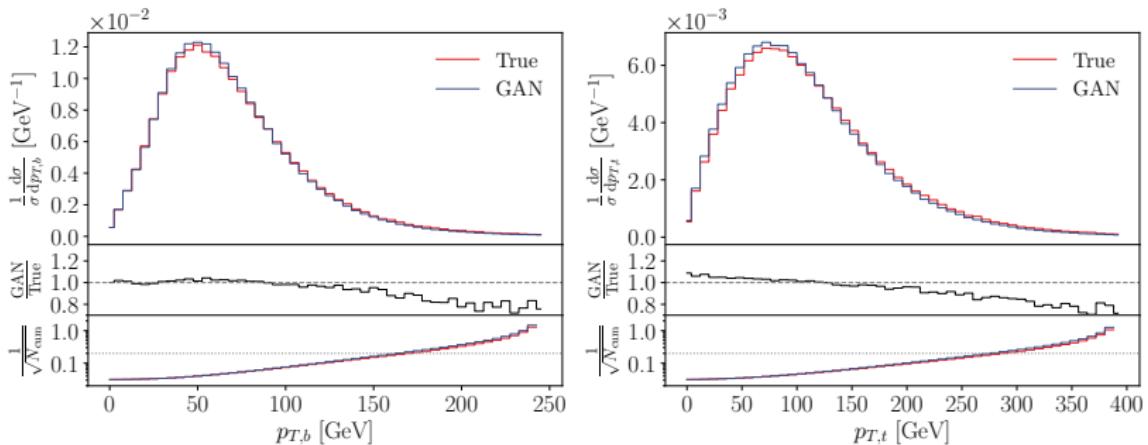
Challenges: 16-dimensional phase-space, 4 resonances,
phase-space boundaries, tails

Remarks: fix masses of final state particles
→ generate 18 dim output
additional loss focusing on phase-space structures
→ MMD Loss

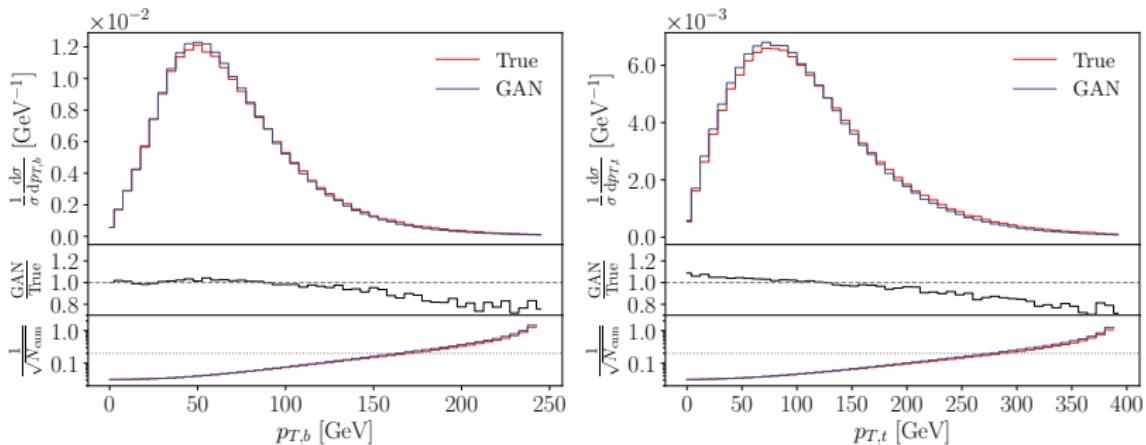
GAN Workflow



Momentum Distributions

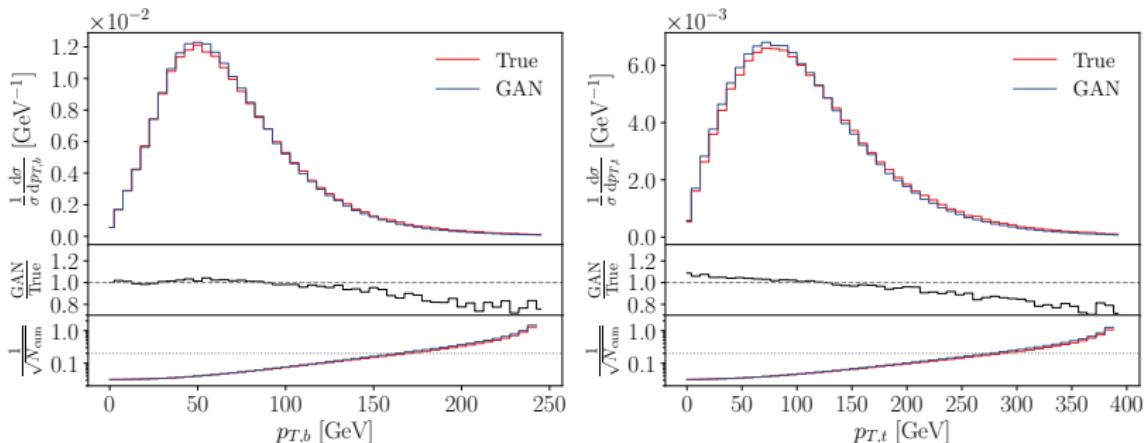


Momentum Distributions



$$N_{\text{tail}}(p_T) = 1 - N_{\text{cum}}(p_T) = b \left(1 - \frac{1}{\sigma} \int_0^{p_T} dp'_T \frac{d\sigma}{dp'_T} \right)$$

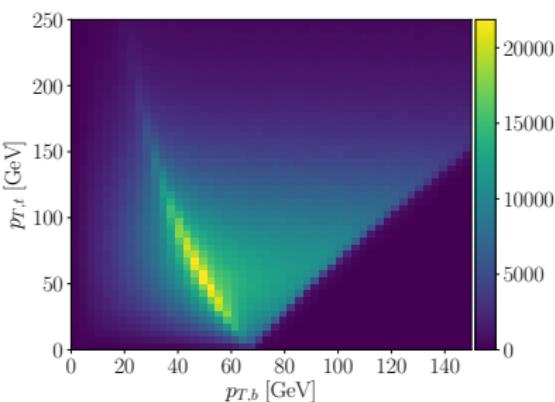
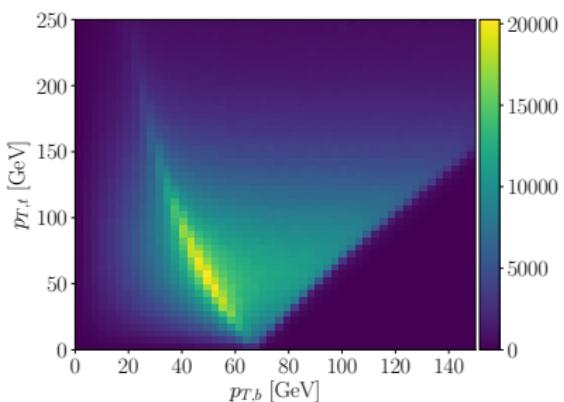
Momentum Distributions



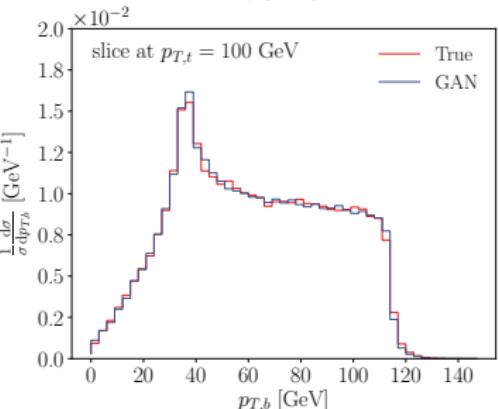
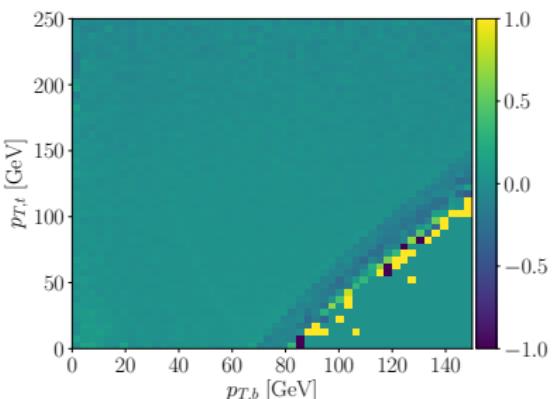
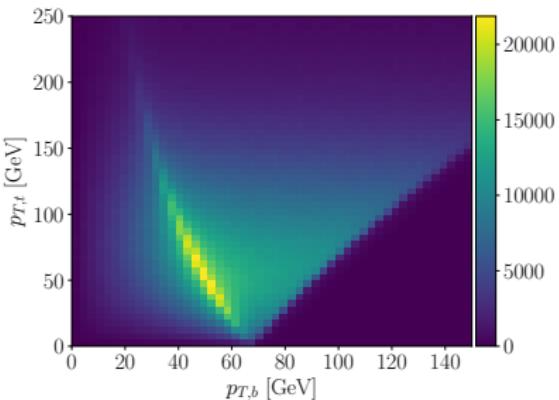
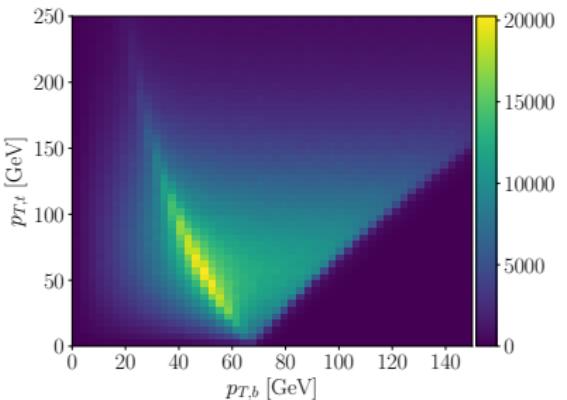
$$N_{\text{tail}}(p_T) = 1 - N_{\text{cum}}(p_T) = b \left(1 - \frac{1}{\sigma} \int_0^{p_T} dp'_T \frac{d\sigma}{dp'_T} \right)$$

→ flat distributions easy to learn!

2-dimensional Correlations



2-dimensional Correlations

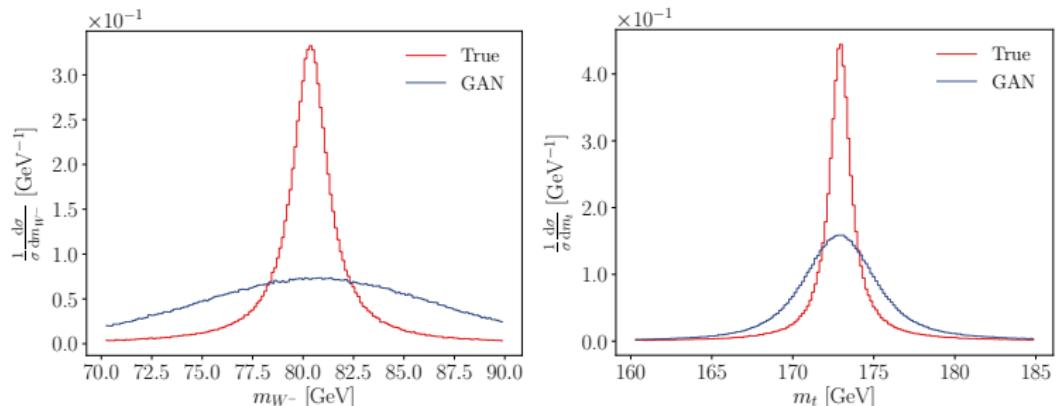


Invariant Mass Peaks

What about the resonances?

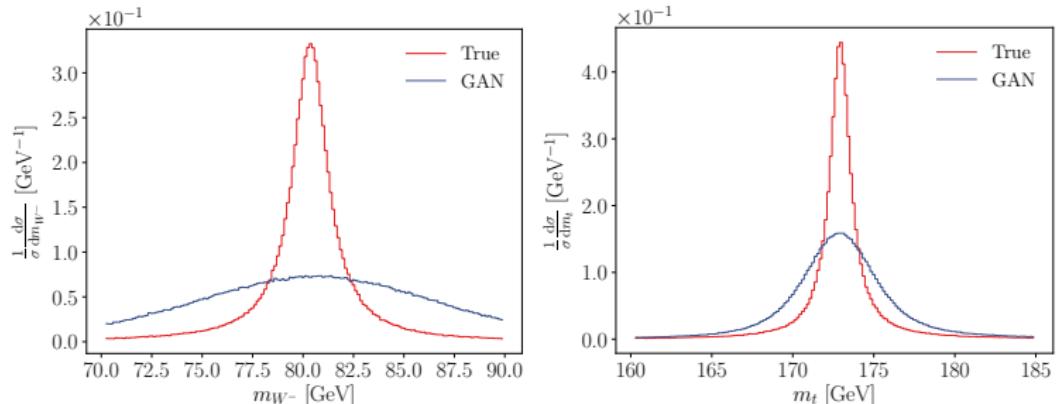
Invariant Mass Peaks

Without the additional loss:



Invariant Mass Peaks

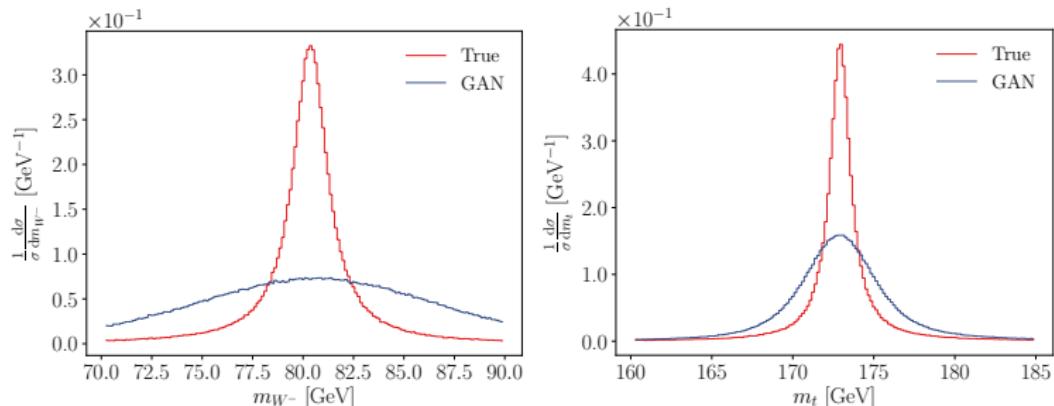
Without the additional loss:



Challenge: resolve the mass peaks

Invariant Mass Peaks

Without the additional loss:



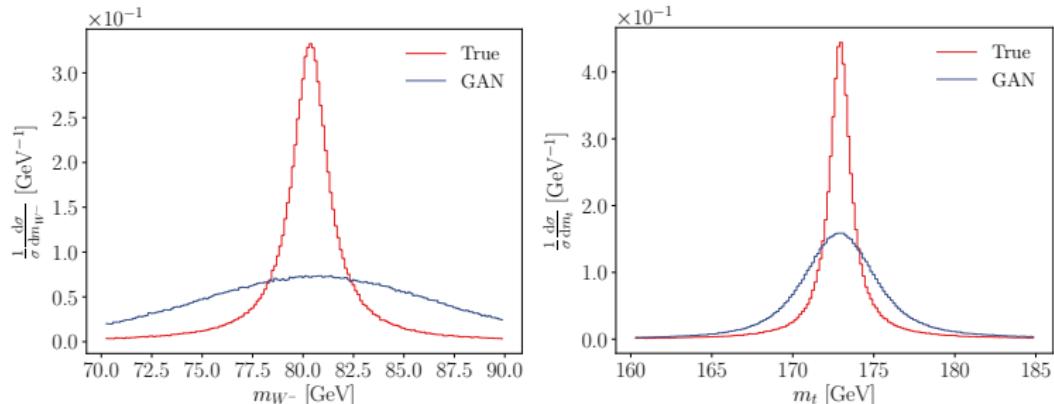
Challenge: resolve the mass peaks

Standard solution: phase-space remapping

$$\int ds \frac{F(s)}{(s - m^2)^2 + m^2\Gamma^2} = \frac{1}{m\Gamma} \int dz F(z) \quad \text{with} \quad z = \arctan \frac{s - m^2}{m\Gamma} .$$

Invariant Mass Peaks

Without the additional loss:



Challenge: resolve the mass peaks

Standard solution: phase-space remapping

$$\int ds \frac{F(s)}{(s - m^2)^2 + m^2\Gamma^2} = \frac{1}{m\Gamma} \int dz F(s) \quad \text{with} \quad z = \arctan \frac{s - m^2}{m\Gamma} .$$

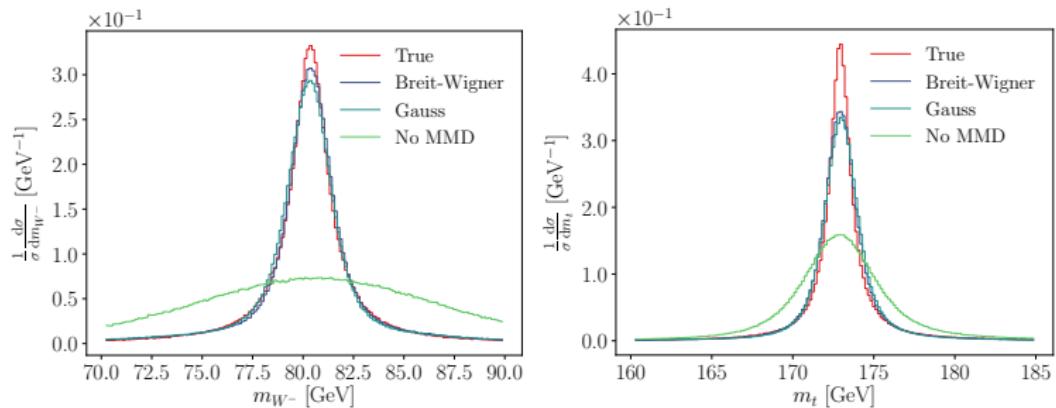
However: knowledge of m and Γ needed

Invariant Mass Peaks

Can we do it better?

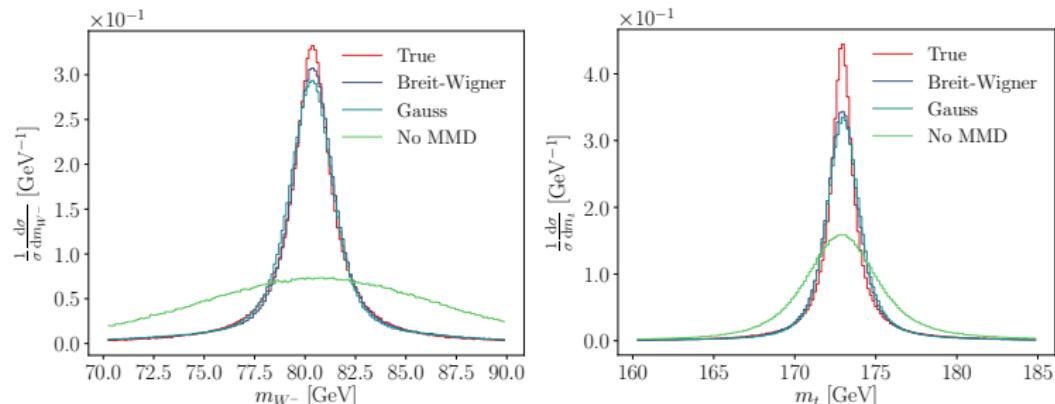
Invariant Mass Peaks

Including the MMD Loss



Invariant Mass Peaks

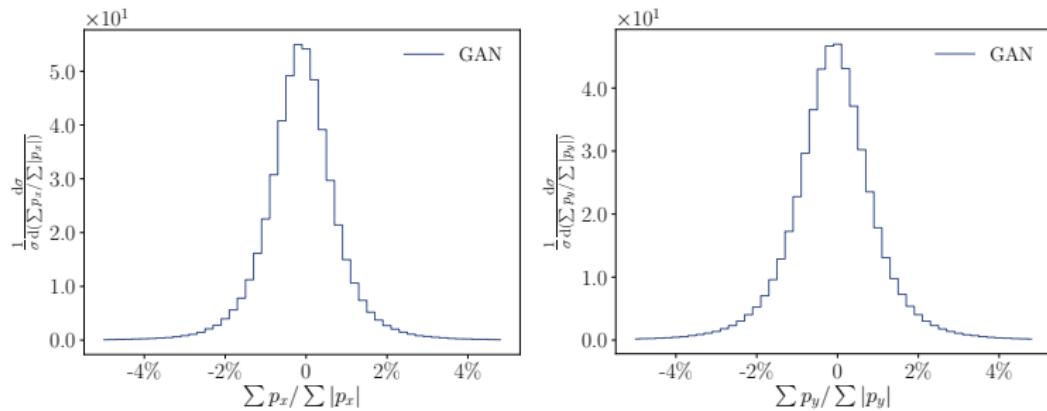
Including the MMD Loss



$$\text{MMD}^2(P_T, P_G) = \langle k(x, x') \rangle_{x, x' \sim P_T} + \langle k(y, y') \rangle_{y, y' \sim P_G} - 2 \langle k(x, y) \rangle_{x \sim P_T, y \sim P_G}$$

- free **kernel** choice → stable results
- **no** knowledge of m and Γ needed

Momentum Conservation by the Network



The generator learns to conserve momentum at a 1% level.

Outlook

- The GAN is able to reproduce the full phase space structure of a realistic LHC process
- Flat distributions can be reproduced at arbitrary precision, limited only by statistics
- Using the MMD loss, we can even describe rich peaking resonances properly
- The same setup will allow us to generate events from an actual LHC event sample
- The GAN does not require any event unweighting

Network Parameters

Parameter	Value
Input dimension G	$18 + 6$
Layers	10
Units per layer	512
Trainable weights G	2382866
Trainable weights D	2377217
λ_D	10^{-3}
λ_G	1
Batch size	1024
Epochs	1000
Iterations per epoch	1000
Training time	26h
Size of trainings data	10^6