

Invertible Networks for Unfolding

Tilman Plehn

Universität Heidelberg

CMS Unfolding 2/2021



How to GAN away detector effects

Goal: invert standard simulation [Bellagente, Butter, Kasiczka, TP, Winterhalder]

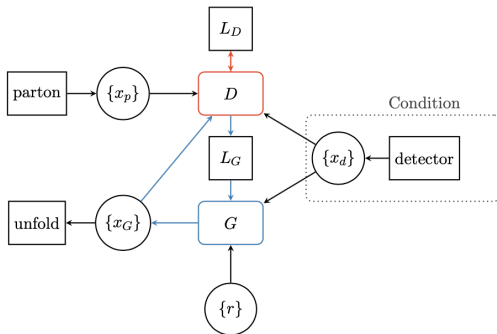
- detector simulation typical Monte Carlo, random-number-driven
- inversion possible, in principle [MEM, but entangled convolutions]
- GAN task

partons $\xrightarrow{\text{DELPHES}}$ detector $\xrightarrow{\text{GAN}}$ partons

⇒ Full phase space unfolded

Conditional GAN

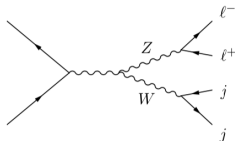
- random numbers to parton level
hadron level as condition
matched event pairs



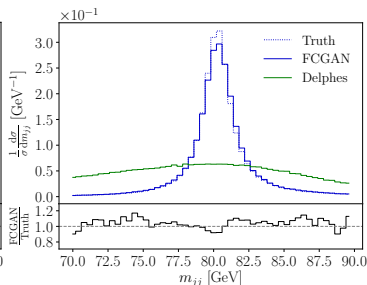
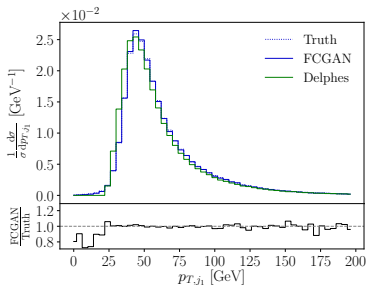
Detector unfolding

Reference process $pp \rightarrow ZW \rightarrow (\ell\ell)(jj)$

- broad jj mass peak
narrow $\ell\ell$ mass peak
modified $2 \rightarrow 2$ kinematics
fun phase space boundaries
- GAN same as **event generation** [with MMD]



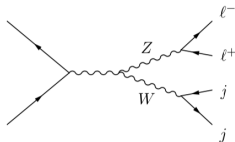
Model (in)dependence



Detector unfolding

Reference process $pp \rightarrow ZW \rightarrow (\ell\ell)(jj)$

- broad jj mass peak
narrow $\ell\ell$ mass peak
modified 2 \rightarrow 2 kinematics
fun phase space boundaries
- GAN same as **event generation** [with MMD]

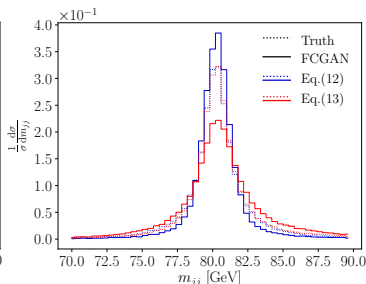
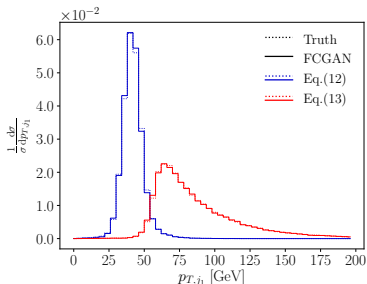


Model (in)dependence

- detector-level cuts [14%, 39% events, no interpolation, MMD not conditional]

$$p_{T,j_1} = 30 \dots 50 \text{ GeV} \quad p_{T,j_2} = 30 \dots 40 \text{ GeV} \quad p_{T,\ell^-} = 20 \dots 50 \text{ GeV} \quad (12)$$

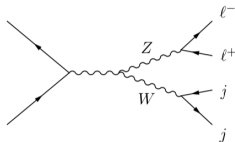
$$p_{T,j_1} > 60 \text{ GeV} \quad (13)$$



Detector unfolding

Reference process $pp \rightarrow ZW \rightarrow (\ell\ell)(jj)$

- broad jj mass peak
narrow $\ell\ell$ mass peak
modified $2 \rightarrow 2$ kinematics
fun phase space boundaries
- GAN same as **event generation** [with MMD]



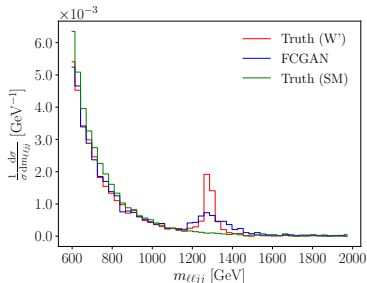
Model (in)dependence

- detector-level cuts [14%, 39% events, no interpolation, MMD not conditional]

$$p_{T,j_1} = 30 \dots 50 \text{ GeV} \quad p_{T,j_2} = 30 \dots 40 \text{ GeV} \quad p_{T,\ell^-} = 20 \dots 50 \text{ GeV} \quad (12)$$

$$p_{T,j_1} > 60 \text{ GeV} \quad (13)$$

- model dependence of unfolding
 - train: SM events
test: 10% events with W' in s -channel
- ⇒ **Working fine, but ill-defined**



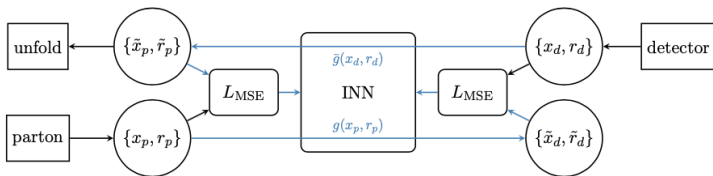
Invertible networks

Standard invertible networks [Bellagente, Butter, Kasieczka, TP, Rousselot, Winterhalder, Ardizzone, Köthe]

- network as bijective transformation — normalizing flow
 Jacobian tractable [specifically: coupling layer]
 evaluation in both directions — INN [Ardizzone, Rother, Köthe]
- mapping parton and detector phase spaces
 padding with random numbers [eINN, dimensionality, sampling for poor]

$$\begin{pmatrix} x_p \\ r_p \end{pmatrix} \begin{matrix} \xleftarrow{\text{PYTHIA, DELPHES: } g} \\ \xrightarrow{\text{unfolding: } \bar{g}} \end{matrix} \begin{pmatrix} x_d \\ r_d \end{pmatrix}$$

- training on event pairs (MSE) or samples (MMD) [thank you to Jessica]

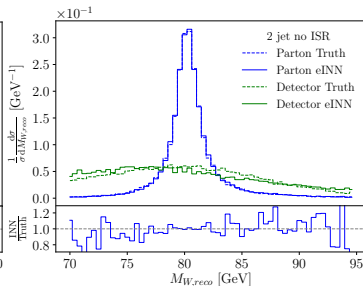
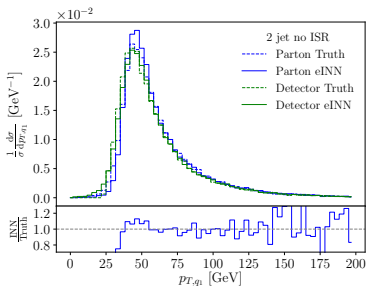


Standard invertible networks [Bellagente, Butter, Kasieczka, TP, Rousselot, Winterhalder, Ardizzone, Köthe]

- network as bijective transformation — normalizing flow
Jacobian tractable [specifically: coupling layer]
evaluation in both directions — INN [Ardizzone, Rother, Köthe]
- mapping parton and detector phase spaces
padding with random numbers [eINN, dimensionality, sampling for poor]

$$\begin{array}{ccc} & \text{PYTHIA, DELPHES: } g \rightarrow & \\ \left(\begin{array}{c} X_p \\ r_p \end{array} \right) & \longleftrightarrow & \left(\begin{array}{c} X_d \\ r_d \end{array} \right) \\ & \leftarrow \text{unfolding: } \tilde{g} & \end{array}$$

- training on event pairs (MSE) or samples (MMD) [thank you to Jessica]
- same task as FCGAN, similar performance



Invertible networks

Standard invertible networks [Bellagente, Butter, Kasieczka, TP, Rousselot, Winterhalder, Ardizzone, Köthe]

- network as bijective transformation — normalizing flow

Jacobian tractable [specifically: coupling layer]

evaluation in both directions — INN [Ardizzone, Rother, Köthe]

- mapping parton and detector phase spaces

padding with random numbers [eINN, dimensionality, sampling for poor]

$$\begin{pmatrix} x_p \\ r_p \end{pmatrix} \xleftarrow[\leftarrow \text{unfolding: } \bar{g}]{\text{PYTHIA, DELPHES: } g \rightarrow} \begin{pmatrix} x_d \\ r_d \end{pmatrix}$$

- training on event pairs (MSE) or samples (MMD) [thank you to Jessica]

- same task as FCGAN, similar performance

⇒ Working okay, still ill-defined



Proper inverting with cINN

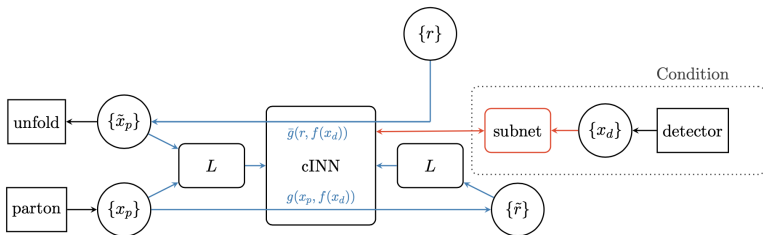
Statistical inversion [Bellagente, Butter, Kasieczka, TP, Rousselot, Winterhalder, Ardizzone, Köthe]

– task: construct parton-level pdf for (single) detector-level event

1- conditional INN: parton-level events from $\{r\}$

2- maximum likelihood loss

$$\begin{aligned}
 L &= - \langle \log p(\theta | x_p, x_d) \rangle_{x_p, x_d} \\
 &\approx - \left\langle \log p(g(x_p, x_d)) + \log \left| \frac{\partial g(x_p, x_d)}{\partial x_p} \right| \right\rangle_{x_p, x_d} - \log p(\theta) \\
 &= - \left\langle - \frac{\|g(x_p, x_d)\|_2^2}{2} + \log \left| \frac{\partial g(x_p, x_d)}{\partial x_p} \right| \right\rangle_{x_p, x_d} - \log p(\theta)
 \end{aligned}$$



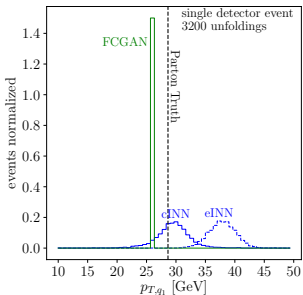
Proper inverting with cINN

Statistical inversion [Bellagente, Butter, Kasieczka, TP, Rousselot, Winterhalder, Ardigzone, Köthe]

- task: construct parton-level pdf for (single) detector-level event
- 1- conditional INN: parton-level events from $\{r\}$
- 2- maximum likelihood loss

Again $pp \rightarrow ZW \rightarrow (\ell\ell) (jj)$

- performance like FCGAN
- distribution: single pair (x_p, x_d) , unfolded many times [FCGAN is out]



Proper inverting with cINN

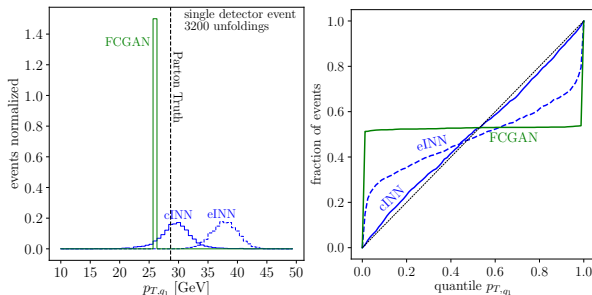
Statistical inversion [Bellagente, Butter, Kasieczka, TP, Rousselot, Winterhalder, Ardizzone, Köthe]

- task: construct parton-level pdf for (single) detector-level event
- 1- conditional INN: parton-level events from $\{r\}$
- 2- maximum likelihood loss

Again $pp \rightarrow ZW \rightarrow (\ell\ell)(jj)$

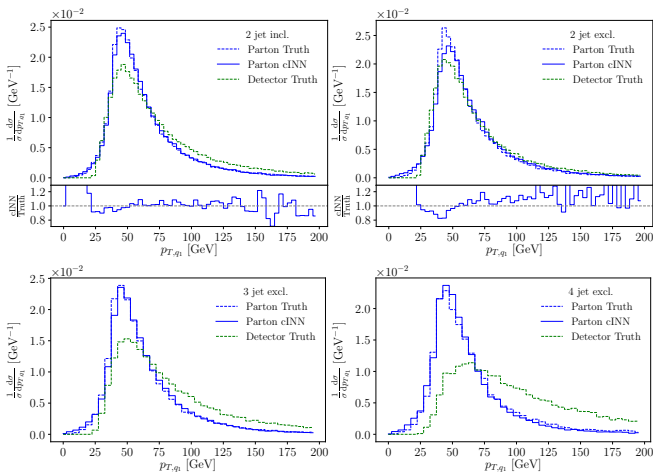
- performance like FCGAN
- distribution: single pair (x_p, x_d) , unfolded many times [FCGAN is out]
- calibration: 1500 pairs (x_p, x_d) , each unfolded 60 times, check for truth

⇒ **cINN well-defined!**



What theorists want: unfolding ISR

- detector-level process $pp \rightarrow ZW + \text{jets}$ [variable number of objects]
- ME vs PS jets decided by network
- training jet-inclusively or jet-exclusively
parton-level hard process extracted as $2 \rightarrow 2$



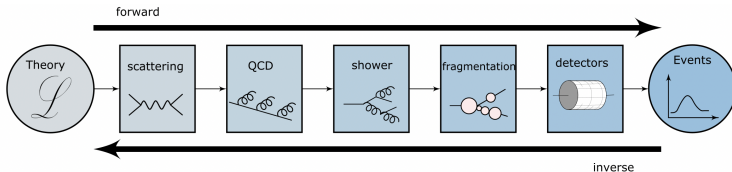
Inverting to hard process

What theorists want: unfolding ISR

- detector-level process $pp \rightarrow ZW + \text{jets}$ [variable number of objects]
- ME vs PS jets decided by network
- training jet-inclusively or jet-exclusively
parton-level hard process extracted as $2 \rightarrow 2$

Towards systematic inversion

- detector unfolding on way
 - QCD parton from jet algorithm standard
 - jet radiation possible
- ⇒ **Hard matrix element a proper goal?**



Things are moving

Machine learning for LHC theory

- big data for fundamental physics
- GANs the cool kid
- INNs the theory hope
- Full inversion in reach?



The poster features a scenic view of Heidelberg, Germany, with the Old Bridge over the Neckar River in the foreground. The background shows the city's architecture and the Neckar Hills under a clear sky. The text is overlaid on the image.

ML4Jets hybrid
July 6-8 2021

INSTITUTE FOR
THEORETICAL PHYSICS

UNIVERSITÄT
HEIDELBERG
ZUKUNFT
SEIT 1386

Local Organizers
Anja Butter
Barry Dillon
Ullrich Köthe
Tilman Plehn
Hans-Christian Schultz-Coulon

International Organization Committee
Kyle Cranmer (NYU)
Ben Nachman (LBNL)
Maurizio Pierini (CERN)
Tilman Plehn (Heidelberg)
Jesse Thaler (MIT)

 <https://indico.cern.ch/event/980214>

Photo: Syntexis / Florida / Adobe Stock, Composition: Kuba Hejzlmann

