

658 **4** | Results

659 This chapter presents the results obtained from the GAN in three channels: $pp \rightarrow t\bar{t}$,
660 $gg \rightarrow ZZ$ and $gg \rightarrow HH$. Within each channel, the Mandelstam variables s and t
661 and the rapidity y are used to identify the event and to train the GAN with an
662 increasing number of events from 10k to 1M. The performance of the generator is
663 evaluated in terms of the accordance between the true and the generated distribution
664 evaluated with the KL divergence, JS divergence, and the ratio of these distributions.
665 To improve the results, an optimization on a set of hyperparameters is done using
666 the tool Hyperopt [refhyperopt].

667 **4.1 Sample processing**

668 The events within each channel are unweighted events produced using MadGraph5
669 and written in a Les Houches Events (LHE) [lhe] file. The selected channels are
670 three different $2 \rightarrow 2$ hadron scattering where the beam energy is $E_{beam} = 6500$
671 GeV:

- 672 – $pp \rightarrow t\bar{t}$, a proton-proton scattering into a top quark pair at LO;
673 – two loop-induced channels at LO, $gg \rightarrow ZZ$ and $gg \rightarrow HH$.

From a given LHE file the events are translated in an array of dimension $[3, N]$ where N is the number of events in the file. The rows of this matrix are filled with three kinematical variables that define the event. The chosen variables are the Mandelstam variables s , t and the rapidity y in the parton scattering reference frame. They are defined by the equations:

$$s = (p_1 + p_2)^2 = (p_3 + p_4)^2 \quad (4.1)$$

$$t = (p_1 - p_3)^2 = (p_2 - p_4)^2 \quad (4.2)$$

$$y = \frac{1}{2} \log \left(\frac{x_1}{x_2} \right) \quad (4.3)$$

674 where p_i is the momentum of the i -th particle and $x_i = E_i/E_{beam}$ is the energy
675 fraction of the beam energy carried by the scattering parton.

Before feeding the GAN with the obtained events, a preprocessing is applied using the **Scikit-Learn** package. Two possible preprocessing are considered: a standardization or a Yeo-Johnson power transform followed by a minmax scaling in the range $[-1, 1]$ of the input features. An example of the effects of these transformations on a sample of 10k $pp \rightarrow t\bar{t}$ events is shown in Fig. 1.1. In the first case, the distributions are not limited therefore the output layer of the generator is a linear function. Instead, in the second case the output is contained in the range $[-1, 1]$, therefore the hyperbolic tangent activation function is more suitable in the output layer. Then, the generated events are post-processed using the inverse transformation of the used scaler.

To rebuild the momenta of the particles from the kinematical variables, assuming massless incoming partons, in the laboratory reference frame, use Eqs. 1.1 and 1.3 to find the energy fractions of the incoming particles:

$$x_2 = \frac{\sqrt{s}}{2E_{beam}} e^{-y}, \quad (4.4)$$

$$x_1 = x_2 e^{2y}, \quad (4.5)$$

where E_{beam} is the beam energy. Thus, the incoming momenta in the lab frame are:

$$p_1 = [x_1 E_{beam}, 0, 0, x_1 E_{beam}], \quad (4.6)$$

$$p_2 = [x_2 E_{beam}, 0, 0, -x_2 E_{beam}]. \quad (4.7)$$

$$(4.8)$$

The other two momenta can be found using the Eqs. 1.1 and 1.2. In the parton scattering reference frame, the energy E_3^* , the scattering angle θ^* and the longitudinal momentum p_z^* of the third particle are:

$$E_3^* = \frac{\sqrt{s}}{2}, \quad (4.9)$$

$$\theta^* = 1 + 2 \frac{t}{s}, \quad (4.10)$$

$$p_z^* = E_3^* \cos \theta^*. \quad (4.11)$$

Finally, a Lorentz boost is needed in the longitudinal direction to obtain the momentum in the laboratory reference frame:

$$E_3 = E_3^* \cosh y + p_z^* \sinh y, \quad (4.12)$$

$$p_z = E_3^* \sinh y + p_z^* \cosh y, \quad (4.13)$$

$$p_T = \sqrt{E_3^2 - (p_z^2 + m^2)}, \quad (4.14)$$

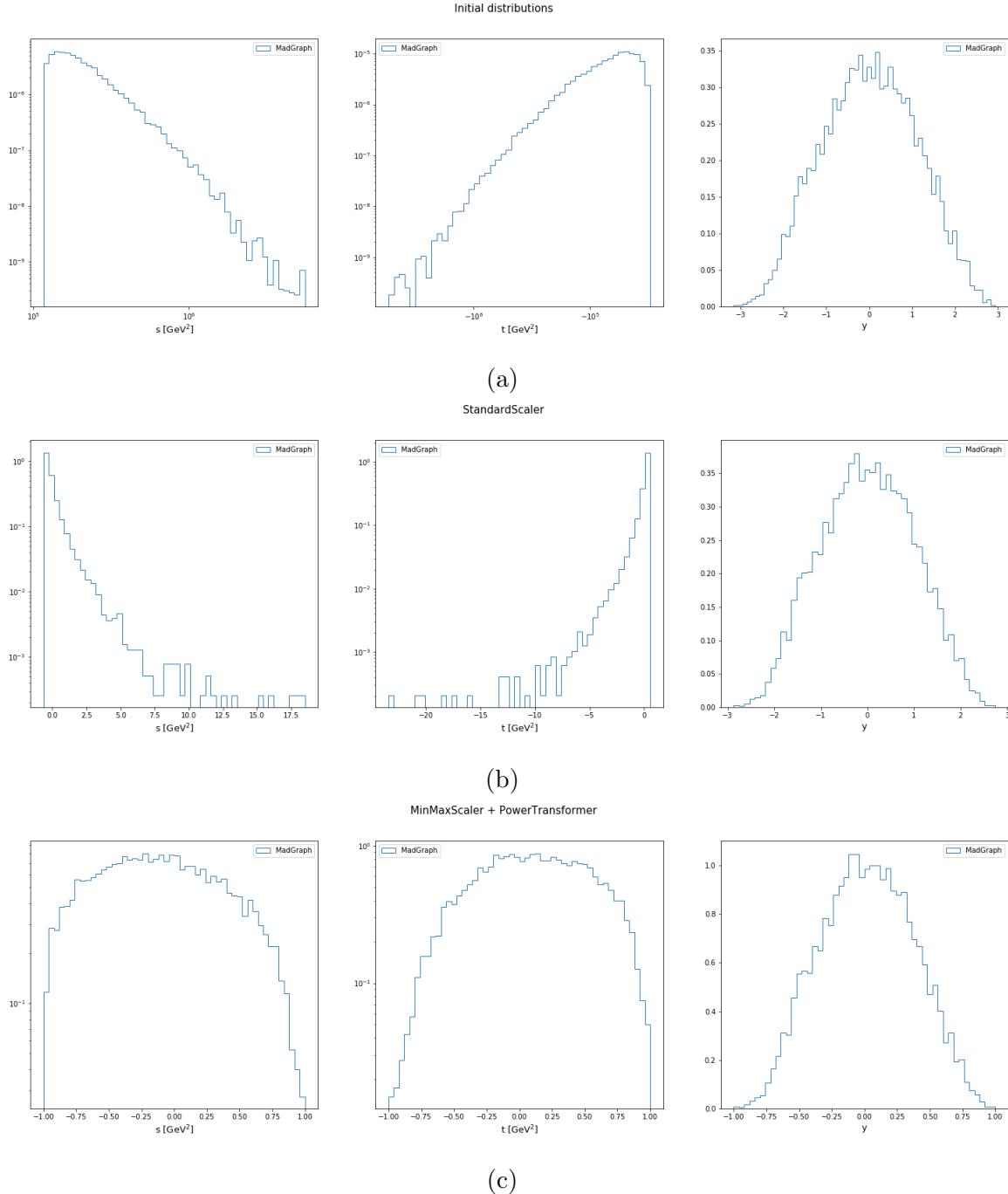


Figura 4.1: Preprocessing of the input features. (a) Distribution without preprocessing. (b) Standard scaling. (c) Power Transformer followed by a MinMax scaling in range $[-1, 1]$.

where p_T is the transverse momentum, found by energy conservation, and m is the mass of the particle. Applying the momentum conservation law, it is possible to find the related quantities of the final outgoing particle:

$$p_3 = [E_3, p_T, 0, p_z], \quad (4.15)$$

$$p_4 = [E_1 + E_2 - E_3, -p_T, 0, E_1 - E_2 - p_z]. \quad (4.16)$$

686 4.2 GAN test and hyperparameter optimization

687 Before starting to train the models in different channels, benchmark training is
 688 done to test the effectiveness of the GAN. The test consists of training the GAN
 689 using data sampled from three Gaussians, one for each input feature, and measuring
 690 the ability of the GAN to reproduce these distributions. The input sample is firstly
 691 scaled in the range $[-1, 1]$ and then feed into the neural networks. The methods
 692 used to evaluate the performance are the KL divergence and the ratio of the GAN
 693 to the expected distribution in each bin. The optimizer used is SGD with default
 694 learning rate and the total number of epochs utilized is 100k, a snapshot of the
 695 obtained distributions at different epochs is shown if Fig. 1.2.

696 This test shows that the proposed GAN is working and it is able to reproduce
 697 the initial core distribution already after 10k epochs where the error in the ratio
 698 plot stabilize within few percent. A deterioration of the output can be seen in
 699 the tails of the distribution where the ratio error falls at 50% due to lack of data.
 700 However, increasing the number of training epoch improve the result also in the
 701 aforementioned region as it can be seen in Fig. 1.2 (c).

702 Subsequently, the optimization through Hyperopt has been performed. The
 703 optimization process trains different models selecting the hyperparameters with
 704 a Bayesian method called Tree-structure Parzen Estimator (TPE). At the end of
 705 the training, the mean of the KL divergence between the distributions of the three
 706 input features is calculated and used as a metric to evaluate the performance of
 707 the model. The first search space defined includes the number of epochs, the batch
 708 size, the learning rate plus the distinction in the two possible preprocessing defined
 709 in the above section.

710 In Figs. 1.3 is shown the loss metric for each iteration obtained using 10k
 711 $pp \rightarrow t\bar{t}$ events and the SGD optimizer. The results show that the minimax
 712 preprocessing is less spread and thus preferable over the standard one. Regarding
 713 the other parameters, a higher batch size reduces the fluctuations introduced by
 714 the stochastic optimizer while, with both preprocessing, a higher learning rate
 715 gave better results. Finally, the number of epochs there is no improvement in the
 716 performances over 30k epochs.

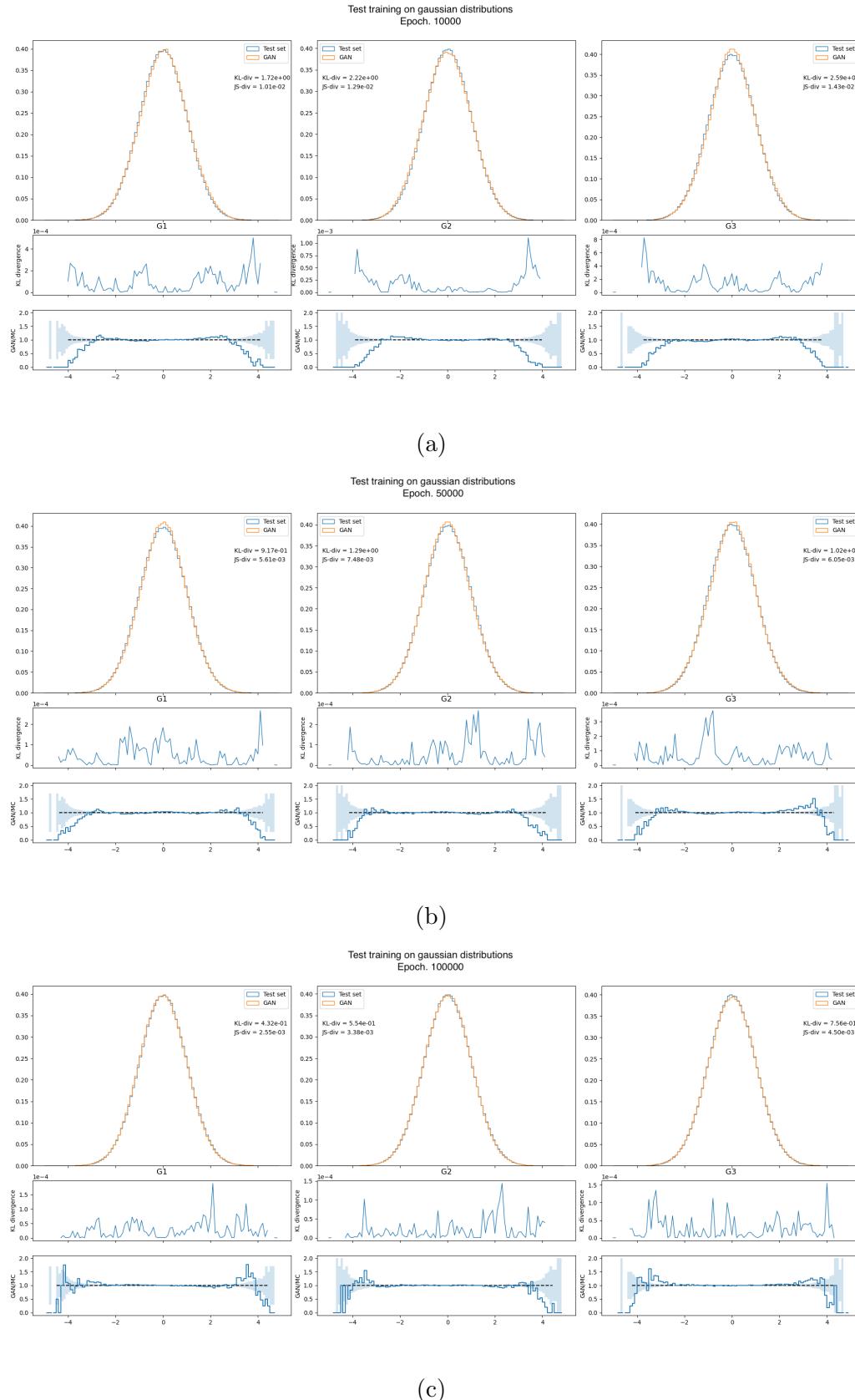


Figura 4.2: Training on three gaussian distributions. The first subplot shows the KL divergence for each bin, while the second one shows the ratio between the generated and the true distribution. The snapshot is done at 10k (a), 50k (b) and 100k (c) epochs.

717 A second optimization is done focusing on the optimizer and the learning rate.
718 The possible choices and the resulting loss metrics are shown in Fig. 1.4. From the
719 results, Adadelta is the most consistent and performing optimizer and a learning
720 rate below 0.05 deteriorates the predictions of the model. After the hyperparameters
721 optimization, all trainings have been done with a minmax preprocessing, with batch
722 size 512, using the Adadelta optimizer with a learning rate of 0.3. However, due to
723 the usage of the stochastic approach, the output distributions present fluctuations
724 around the best local minimum found. In order to select the best model, the
725 training is prolonged to 50k epochs and the best iteration, evaluated in terms of
726 the KL divergence, is the reference for the best model to save. These results for 10k
727 events are rescaled for samples of higher size keeping constant the ratio between
728 the total number of events fed into the network and the sample size.

729 The loss functions of all these trainings are similar and an example one is
730 reported in Fig. 1.5. The loss functions show that the optimal loss value is reached
731 after \sim 1000 epochs. Nevertheless, the training continues after that point because
732 the correlation between the input features keeps improving up to tens of thousands
733 epochs, where the models start to be saved. The training time for such architecture
734 goes from \sim 2 hours to \sim 6 hours depending on the sample size, done using an
735 NVIDIA Tesla P100-PCIe-12GB GPU.

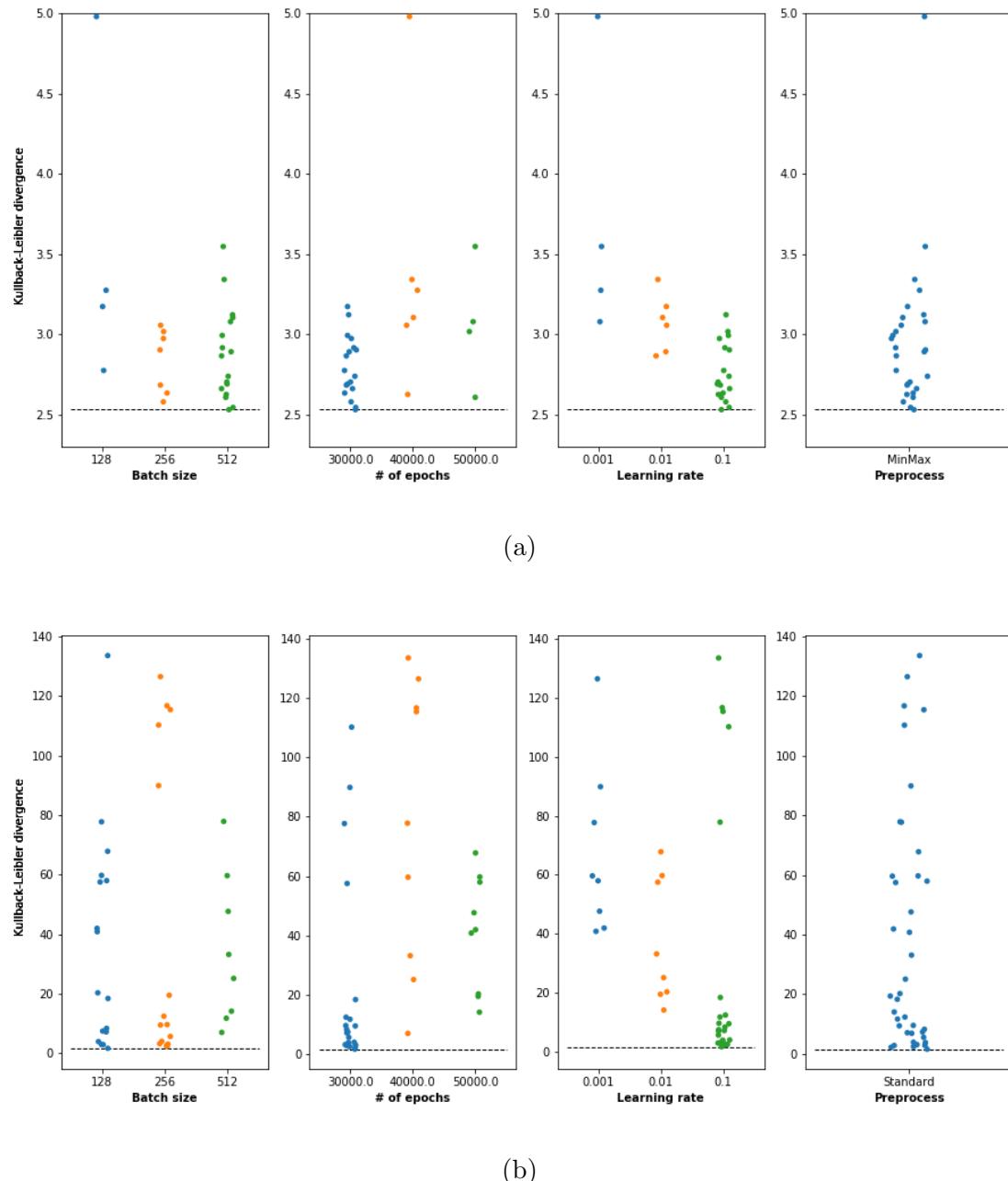


Figura 4.3: First hyperparameters optimization. The search space includes: the batch size, number of epochs, learning rate and the preprocessing used. The y-axis indicates the mean of the KL divergence in the output distributions which is the metric to minimize.

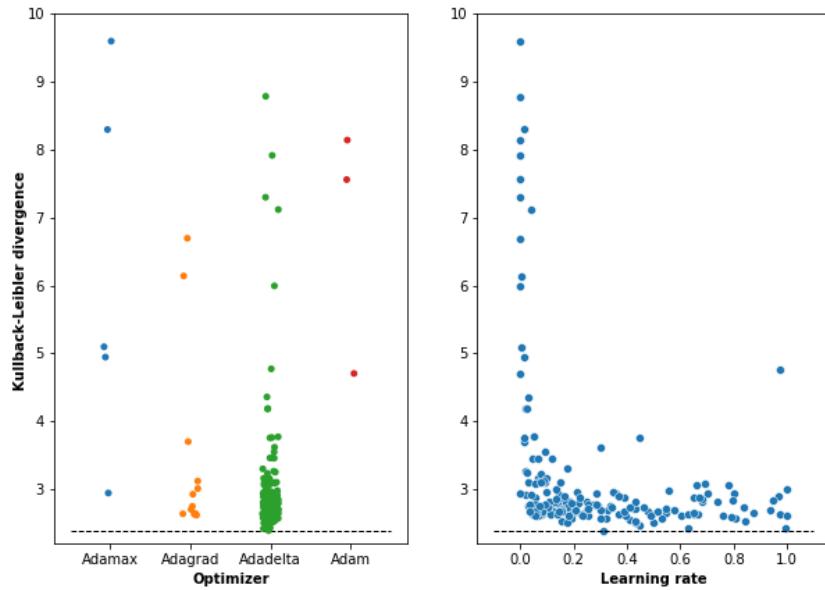


Figura 4.4: Second hyperparameters optimization. The search space includes: the optimizer and a continuous range of possible learning rates. The metric used is the same of Fig.1.3.

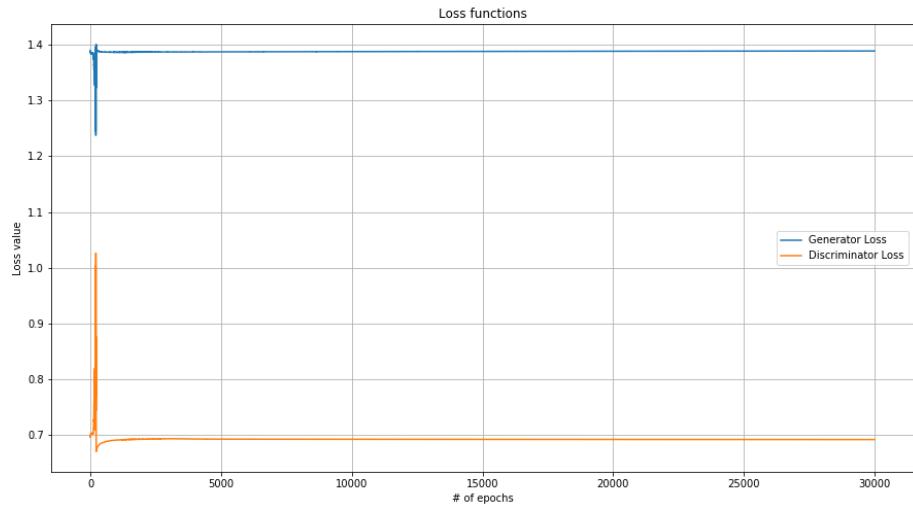


Figura 4.5: Loss functions for a training with 30k epochs.

736 4.3 Generated samples

737 Using the training setup defined in the above section, the first target is to reproduce
738 the same statistics generated with MadGraph using the GAN. To this purpose,
739 two samples for each channel have been generated, a low statistics sample with 10k
740 events and a high statistics sample with 1M events. The following plots present
741 a comparison between the generated and the true distributions for both samples,
742 the KL divergence, and the ratio have been calculated for each bin. The error on
743 the ratio subplot is the statistical counting error calculated as the square root of
744 the number of events. Furthermore, the correlation between invariants has been
745 checked through three 2D scatter plots and the agreement between the plots has
746 been evaluated binning the space and calculating the counting error. Finally, the
747 distributions of the standard deviations formed by all bins are produced and fitted
748 to a gaussian,

749 $pp \rightarrow t\bar{t}$ channel

750 The obtained distributions in the $pp \rightarrow t\bar{t}$ channel are shown in Fig. 1.6. To
751 do not neglect the tails region the s and t histograms have log-scaled bins. The
752 first histogram shows that the ratio error, for the sample of 10k events, goes
753 from 10% in the core region to 20% in the tail region and that the result is
754 everywhere compatible within 3σ . As shown in Fig. 1.7, the correlations between
755 the kinematical variables are well reproduced. In particular, the subplot (b) shows
756 that the model uniformly learned the input distribution with few bins that have an
757 error, in terms of standard deviations, above 3σ . These bin-wise errors are collected
758 into the (c) histogram which shows that this distribution is well centered in zero
759 with variance 1.5.

760 Regarding the sample with 1M events, the ratio error is below 5% in the core
761 region for all input features and it degrades up to 10% before the high energy
762 region. Here and for high rapidity points, the model has difficulties in finding the
763 proper parameters due to the lack of data. The studies on the correlations are
764 summarized in Fig. 1.8. Here, the same conclusions of the sample with 10k events
765 can be drawn with the exception of the correlation between s and t . The models
766 struggle to parametrize the cut imposed by the Mandelstam invariants, resulting
767 in the leakage of events over the boundary and its worst representation.

768 $gg \rightarrow ZZ$ channel

769 The results in this channel are in accordance with the ones presented above. Fig. 1.9
770 shows that the ratio between the true and generated distributions is compatible
771 within 3σ in the majority of the bins for the 10k and 1M events samples and the
772 KL divergence subplots show that the distribution is uniformly learned. The only

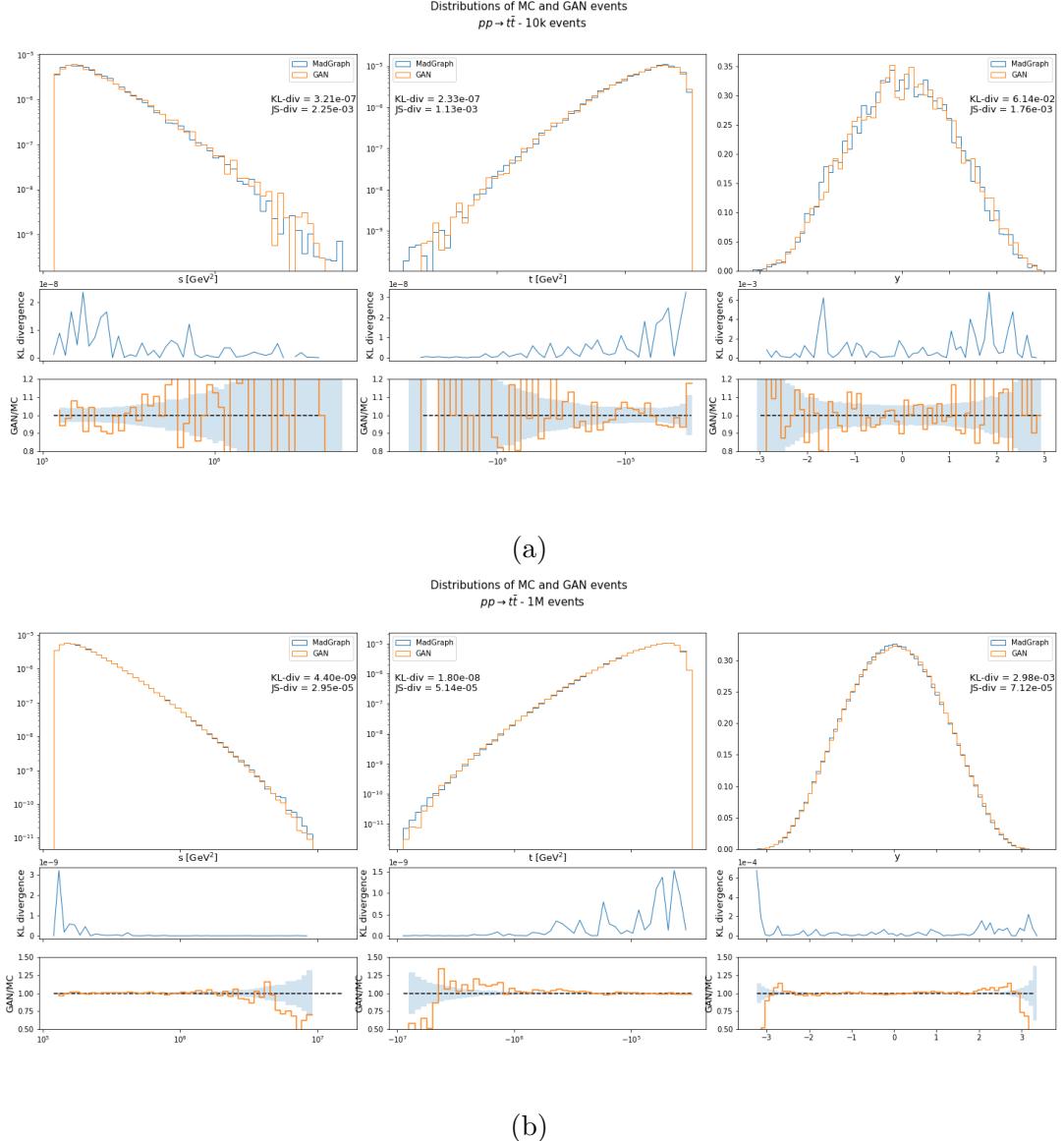


Figura 4.6: Histograms for the $pp \rightarrow t\bar{t}$ channel with (a) 10k events and (b) 1M events. the first and the second subplots show respectively the bin-wise KL divergence and the ratio of the histograms.

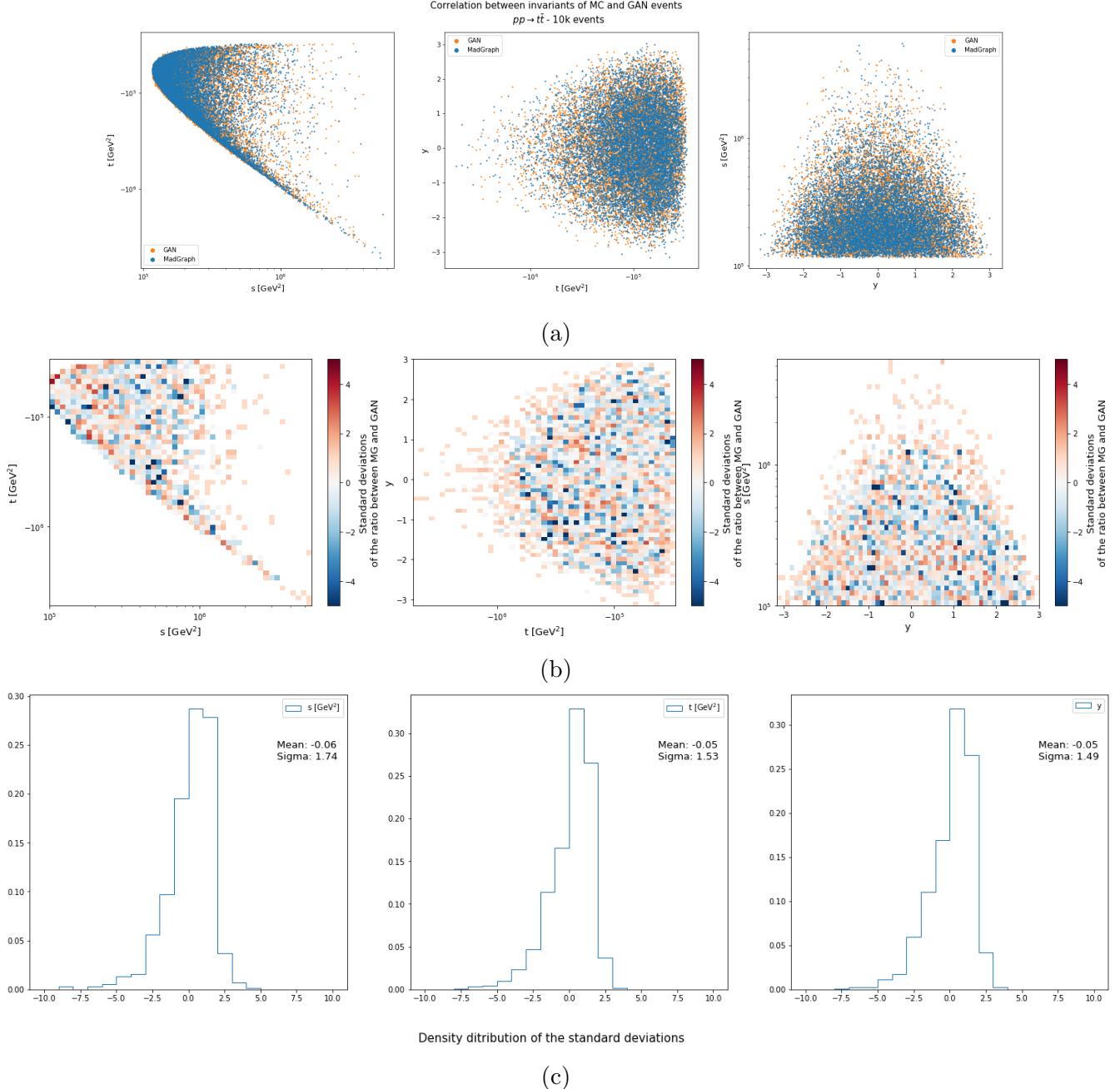


Figura 4.7: Correlations plot for the channel $pp \rightarrow t\bar{t}$ with 10k events. (a) Scatter plot of the two distributions, (b) bin-wise counting error in units of standard deviations, (c) distribution of the errors in units of standard deviations for all bins.

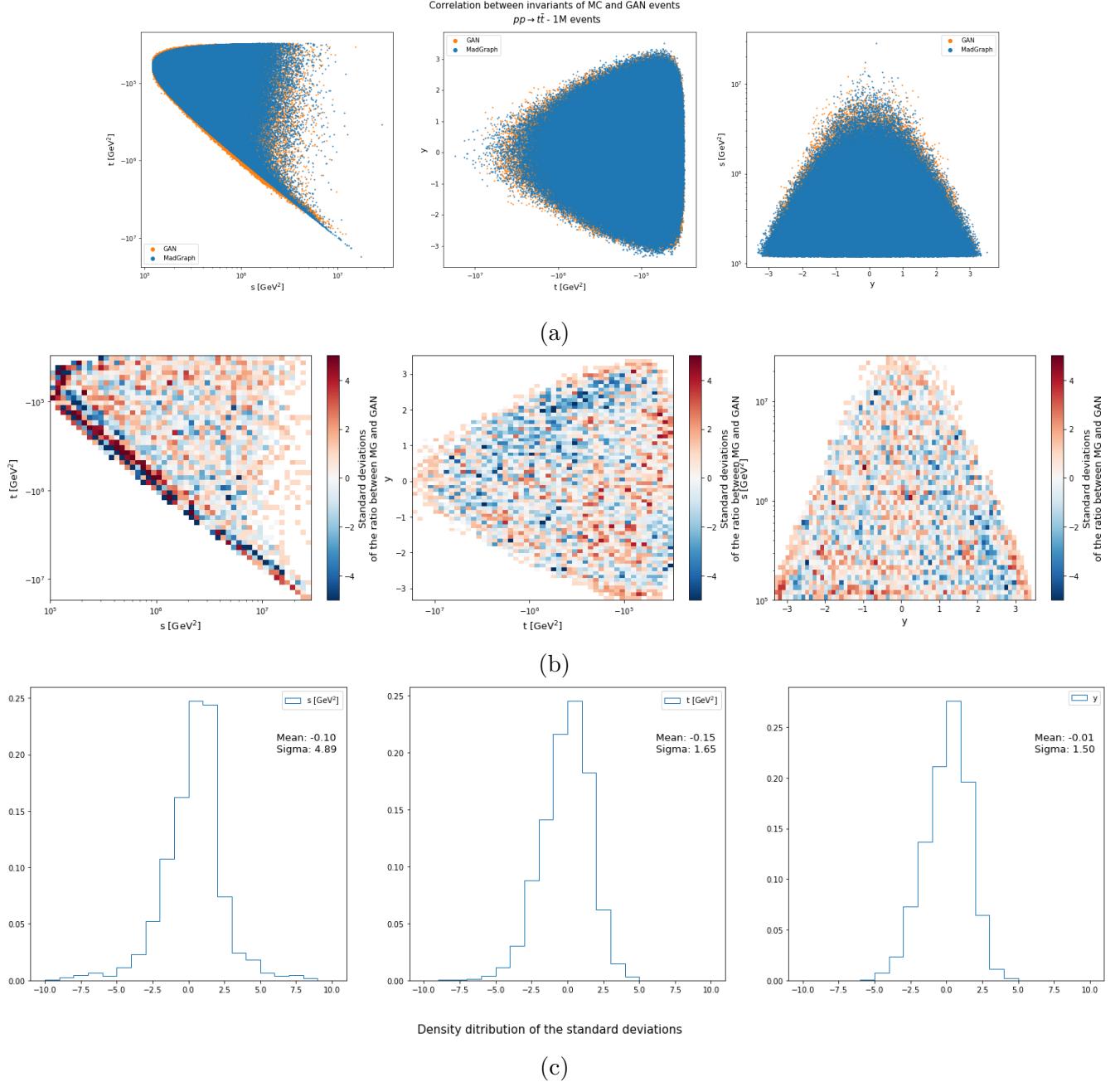


Figura 4.8: Correlations plot for the channel $pp \rightarrow t\bar{t}$ with 1M events. (a) Scatter plot of the two distributions, (b) bin-wise counting error in units of standard deviations, (c) distribution of the errors in units of standard deviations for all bins.

773 region not well reproduced is the tails region where the model falls faster than the
774 true distribution.

775 The correlations plots in Figs. 1.10, 1.11 show that the correct relation between
776 kinematics is uniformly learned for both samples. Indeed, the error distributions
777 are zero-centered with a variance of ~ 1.5 . The noticeable areas with deviations
778 from the true correlations are in the sample with 1M events. These areas are the
779 cut between s and t , which shows some events leakage, and the high-rapidity region,
780 which is not symmetrical.

781 **$gg \rightarrow HH$ channel**

782 This last channel has a slightly different phase space where the matrix element
783 suppresses the Higgs di-boson at the production threshold. These differences can
784 be seen in the histogram plots in Fig. 1.12 for the samples of 10k and 1M events.
785 This effect impacts the s and t distributions especially in the case of 1M events.
786 Indeed, the histograms of the sample with 10k events show a ratio error within 20%
787 and compatible with the statistical error of the true sample. On the other hand,
788 the sample with 1M events presents a ratio error of 5% in the core region above
789 the statistical error. The rapidity is not affected and shows similar results to the
790 other channel with errors below 5%. Regarding the tail region, the model struggles
791 to reproduce the true distribution, as in the other channels, due to lack of data.

792 The correlations plot in Fig. ?? refers to the sample of 10k events. In this
793 case, a minor leakage is present in the cut between the Mandelstam invariants,
794 while the other plots show that the correlations are correctly uniformly reproduced.
795 As in the other channels, the distributions of the errors are zero-centered with a
796 variance of ~ 1.5 . On the contrary, the correlations of the high statistics sample,
797 presented in Fig. ??, show better where the distributions are wrongly reproduced.
798 The interested areas are the cut between the two Mandelstam variables, which
799 shows a major leak of events, and the low energy region, which is, compared to the
800 other channels, less populated.

801 To study the effects of cuts in the training sample, an additional training has
802 been done with the application of two cuts in the $pp \rightarrow t\bar{t}$ channel:

- 803 – a cut on the transverse momentum: $p_T > 250$ GeV;
- 804 – a cut on the rapidity of the outgoing particles: $-2.5 < \eta < 2.5$.

805 Fig. 1.15 shows the histograms of the input features and the cuts reconstructed by
806 the model. As expected, the histogram of the input distribution is well reproduced.
807 Regarding the imposed cuts, the model identifies and tries to reproduce them.
808 Indeed, only a few percent of events are beyond the imposed cuts.

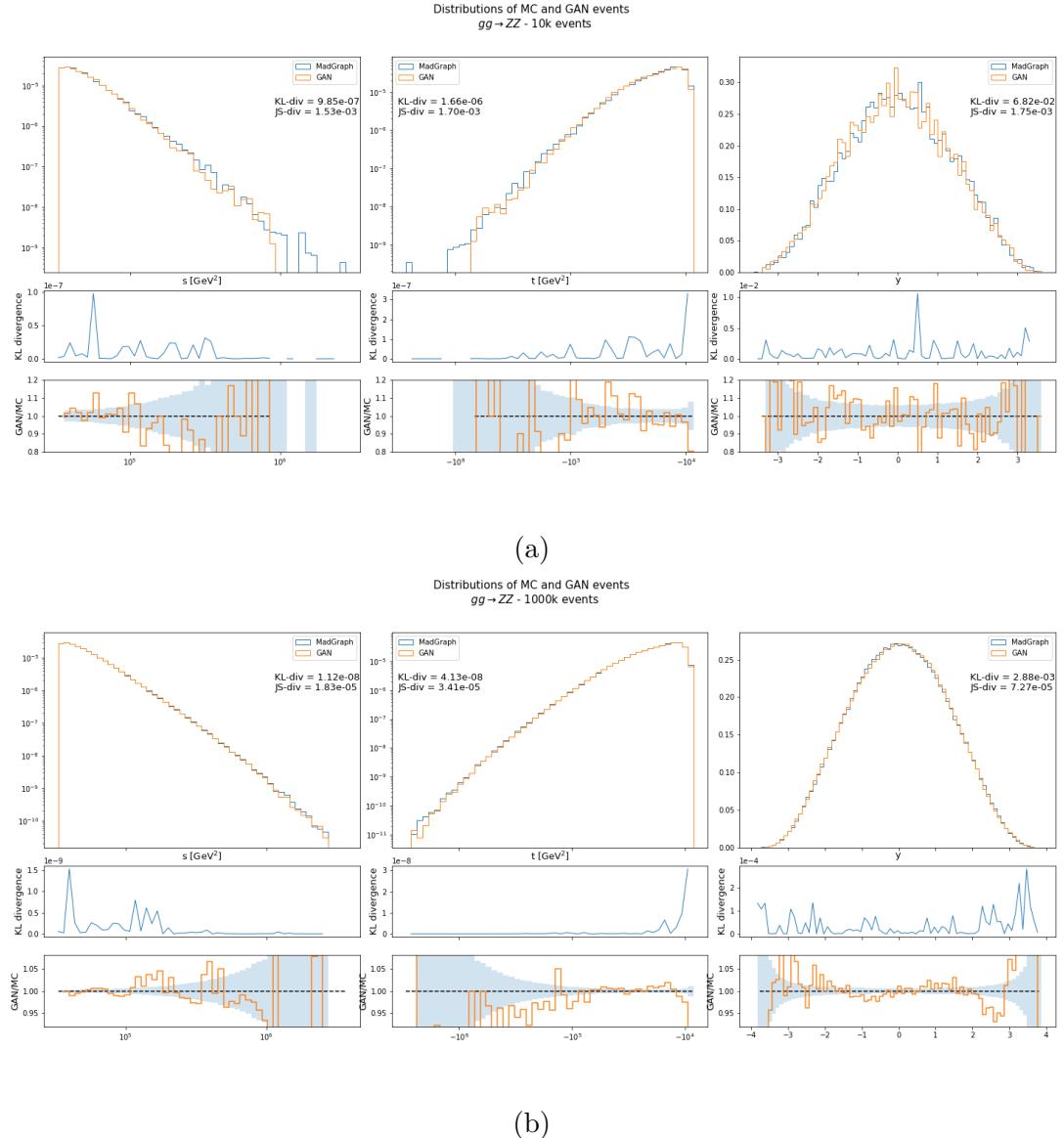


Figura 4.9: Histograms for the $gg \rightarrow ZZ$ channel with (a) 10k events and (b) 1M events. the first and the second subplots show respectively the bin-wise KL divergence and the ratio of the histograms.

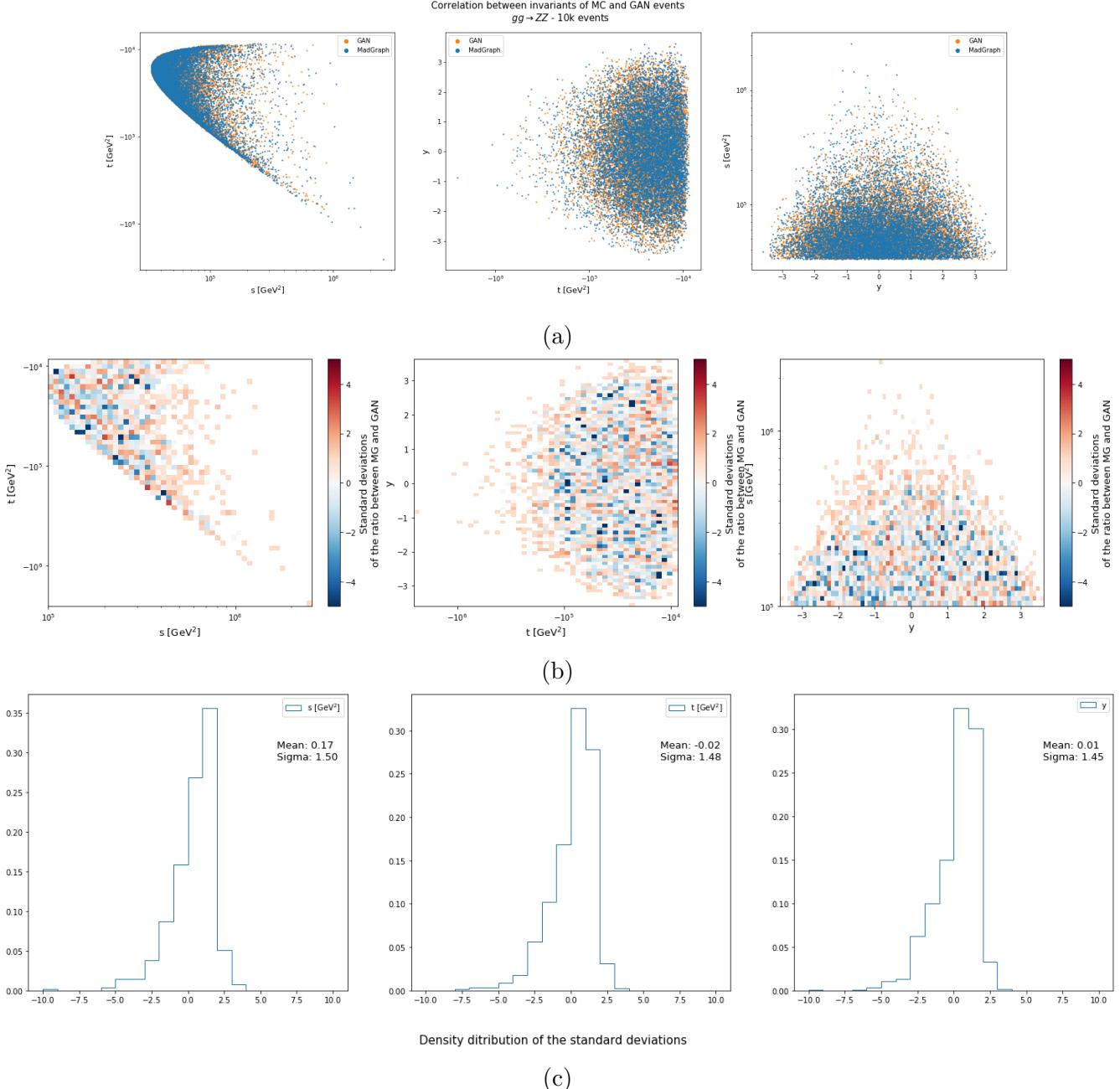


Figura 4.10: Correlations plot for the channel $gg \rightarrow ZZ$ with 10k events. (a) Scatter plot of the two distributions, (b) bin-wise counting error in units of standard deviations, (c) distribution of the errors in units of standard deviations for all bins.

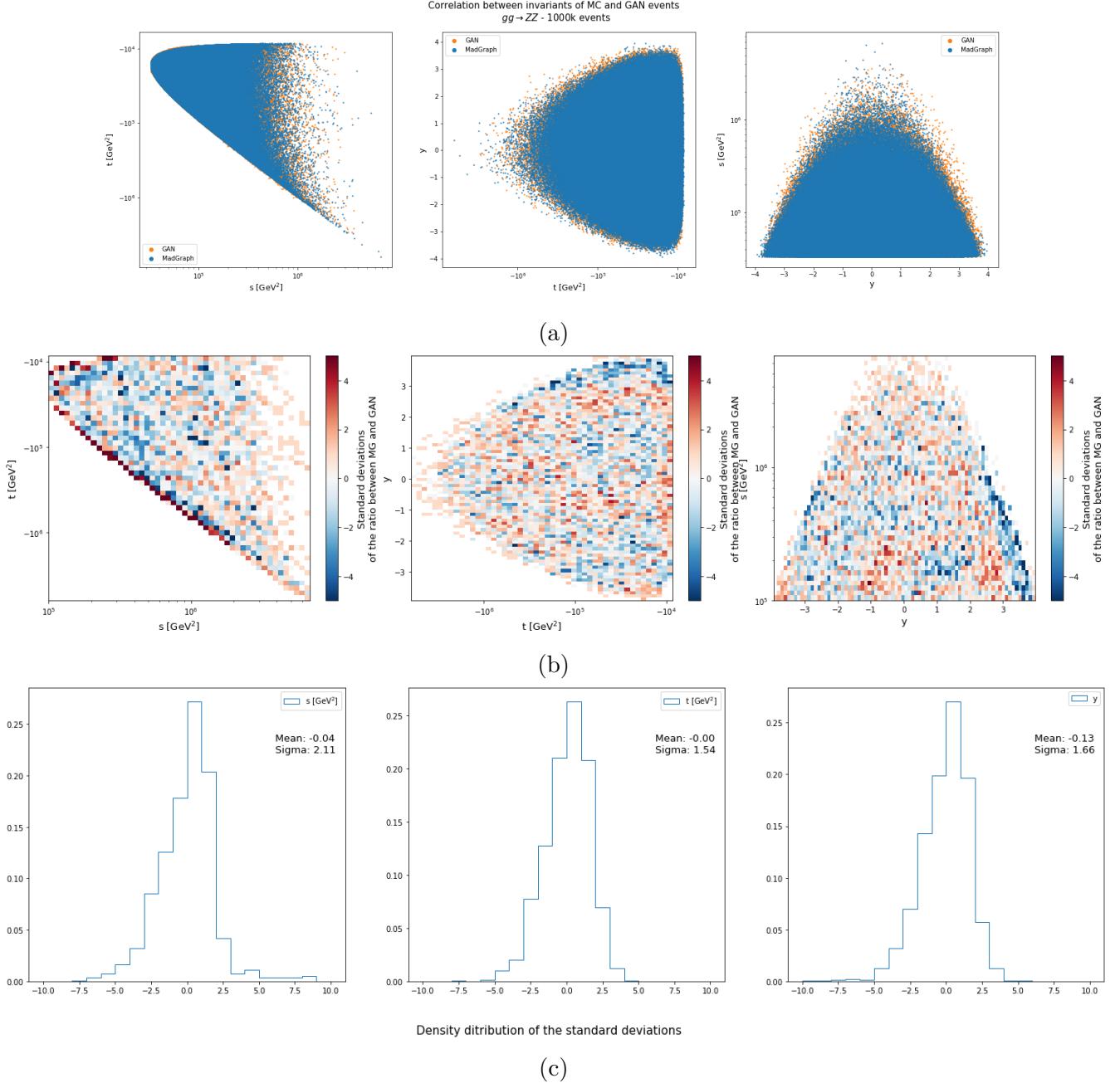


Figura 4.11: Correlations plot for the channel $gg \rightarrow ZZ$ with 1M events. (a) Scatter plot of the two distributions, (b) bin-wise counting error in units of standard deviations, (c) distribution of the errors in units of standard deviations for all bins.

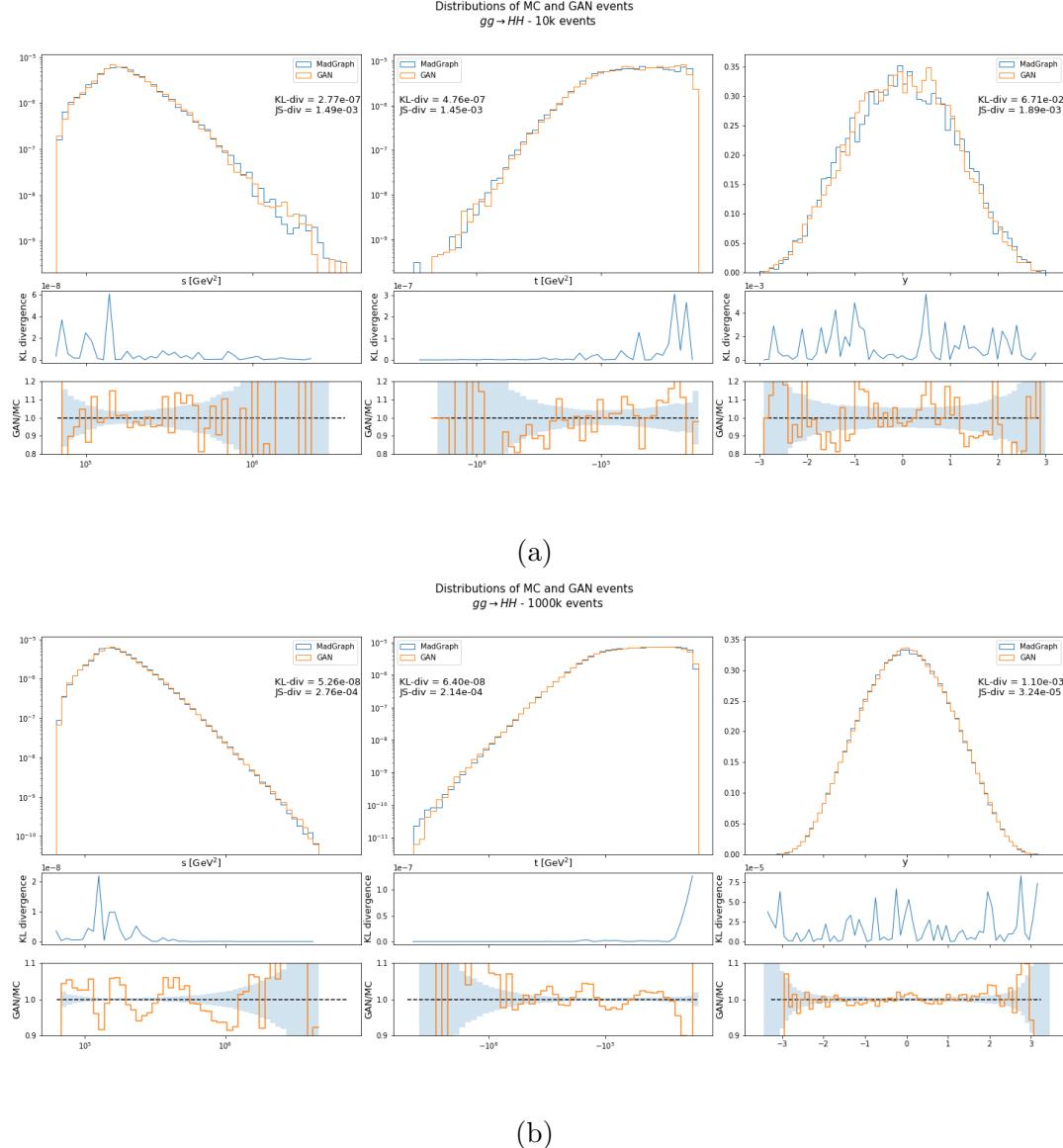


Figura 4.12: Histograms for the $gg \rightarrow HH$ channel with (a) 10k events and (b) 1M events. the first and the second subplots show respectively the bin-wise KL divergence and the ratio of the histograms.

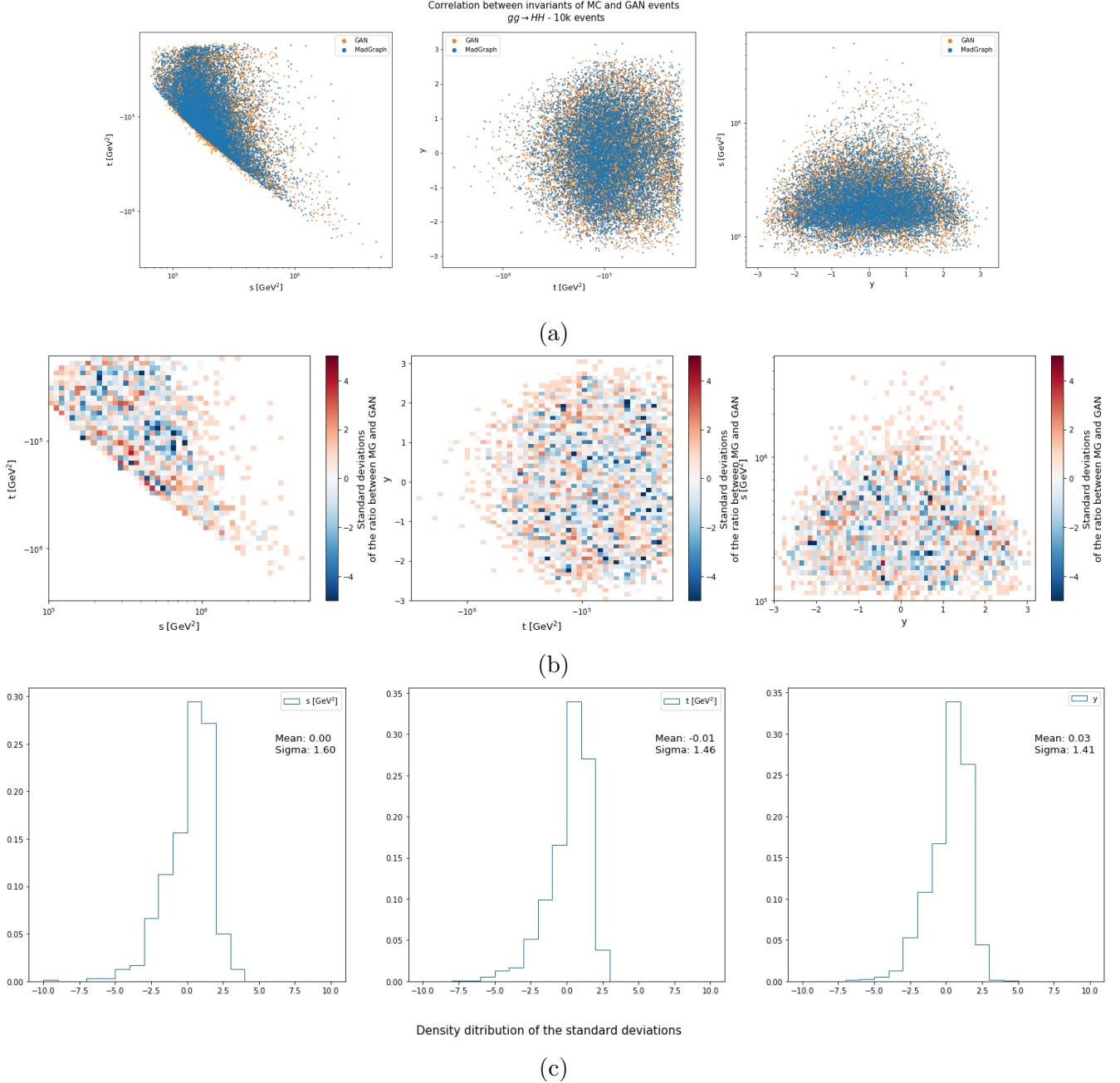


Figura 4.13: Correlations plot for the channel $gg \rightarrow HH$ with 10k events. (a) Scatter plot of the two distributions, (b) bin-wise counting error in units of standard deviations, (c) distribution of the errors in units of standard deviations for all bins.

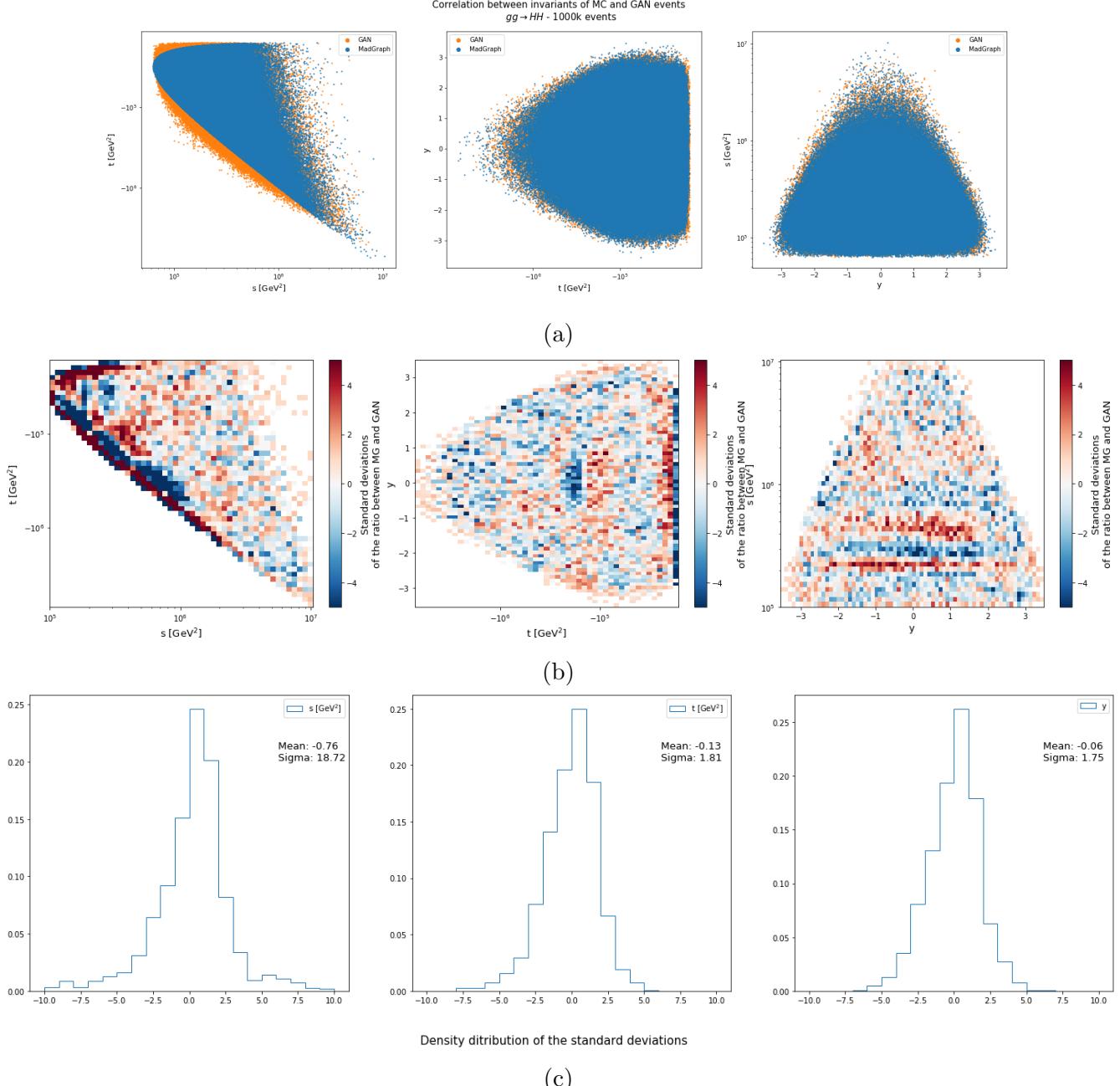


Figura 4.14: Correlations plot for the channel $gg \rightarrow HH$ with 10k events. (a) Scatter plot of the two distributions, (b) bin-wise counting error in units of standard deviations, (c) distribution of the errors in units of standard deviations for all bins.

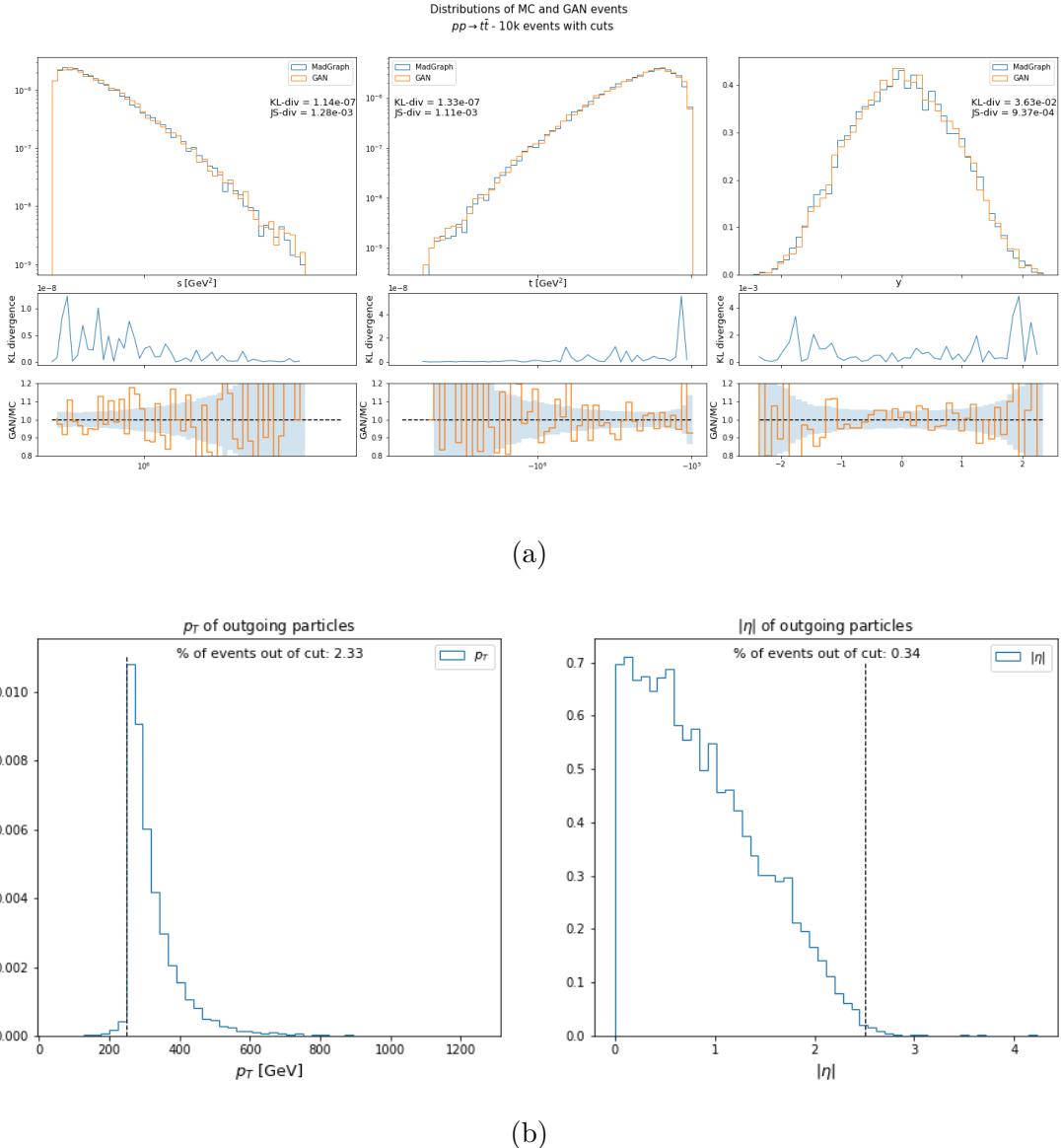


Figura 4.15: Histograms of the distributions of the input feature with imposed cuts on external particles (a). Histograms of the imposed cuts on the external particles (b).

Overall, these plots show that the defined model is able to correctly reproduce the input distribution for each channel. The difficulties found are mainly related to the correlation between the Mandelstam variables and the tails region. The first issue is related to the presence of a cut imposed by Mandelstam invariants, while the second is caused by the low number of events on which the training is done in these regions. Having in mind these drawbacks, these models are a possible way to speed up the generation process of events and decrease the CPU usage¹.

4.4 Data augmentation

The next goal of this thesis is to study if the proposed models can generate an events sample with higher statistics than used for training. This would mean that the GAN can learn the underlying distribution of the kinematical variables without reproducing the statistical noise present in the training sample. To this purpose, the data augmentation capability of the models has been tested with four different events samples, on the same channel of the previous section, with increasing size: 10k events, 50k events, 100k events, and 500k events. The ceiling number of events fixed to test these models is a sample of 1M events for each channel. The procedure consists of training five models in each channel, one for each sample size defined above. Then, each model generates a sample with 10k, 50k, 100k and 500k events which is compared to a slice of the same size extracted from the sample of 1M events.

The following plots show, for each channel, the histograms and the correlations of the input features. They are structured such that a row corresponds to a single trained model whereas a column corresponds to a fixed size of the sample. In this context, the comparison between the true distribution and the generated one, obtained from the same number of events, appears on the diagonal of the 5×5 matrix plot.

$pp \rightarrow t\bar{t}$ channel

The histograms of the input features organized as described above are presented in Figs. 1.16, 1.17 and 1.18. From these plots emerge that the ratio error between the true and the generated distributions is always within the 20%. The more noticeable areas where the discrepancy between the two distributions arise are the tails region for the Mandelstam invariants where, in case of low statistics of the training sample, the model falls off without reproducing the true high statistics tail. Moreover, the rapidity learned at 10k events shows an increasing asymmetry with respect to the true rapidity at bigger samples. The correlations plot of these

¹The time needed for the generation of 1M events was ~ 2 minutes on laptop

histograms are shown in Figs. 1.19, 1.20, 1.21 together with the summary of the errors distribution in Fig. 1.22. These plots confirm the difficulty of the models to reproduce the cut of the Mandelstam variables and that a model trained on a low statistics sample is not able to reproduce an arbitrarily large sample. A possible estimate of data augmentation capability, considering the error distributions, is an increase of factor 10 from the starting sample size.

$gg \rightarrow ZZ$ and $gg \rightarrow HH$ channels

The histogram of the other two processes are presented in Figs. 1.23, 1.24, 1.25 for $gg \rightarrow ZZ$ channel and in Figs. 1.30, 1.31, 1.32 for $gg \rightarrow HH$ channel. For both channels, similar points can be highlighted. As in the $pp \rightarrow t\bar{t}$ channel, the ratio error is within 20% in the majority of the bins and becomes lower in the core of the distributions for training with a higher training sample. The worst reconstruction of the true distribution of the augmented samples is also in the tail region because, during the training, these events are not fed into the GAN. On the other hand, the augmented rapidity distributions are better reproduced as can be seen by the lower asymmetry and the ratio error within 10% in the core region. The correlations plot and the connected errors distribution plot of the $gg \rightarrow ZZ$ process, shown in Figs. 1.26, 1.27, 1.28, and 1.29, highlight again that the cut is difficult to parametrize whereas the correlations between the other input features are uniformly learned. The same conclusion can be drawn from the $gg \rightarrow HH$ channel, the correlations between s and t are not well reproduced due to the different phase space as stated in the previous section. On the other hand, the correlations of the other input features uphold the data augmentation capability of factor 10. The plots are shown in Figs. 1.33, 1.34, 1.35, 1.36.

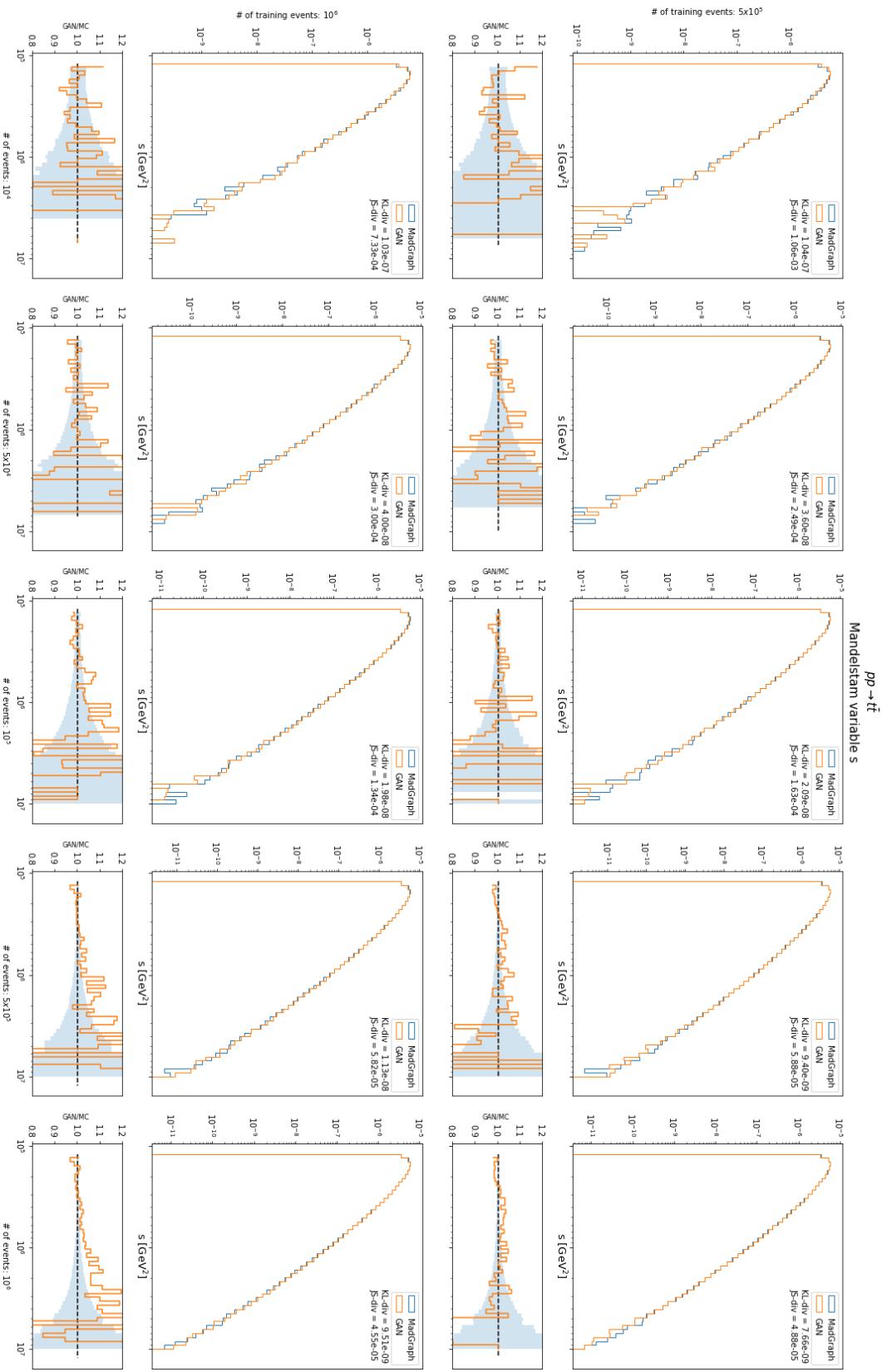


Figura 4.16 (b): Matrix of histograms of the Mandelstam variable s for the channel $pp \rightarrow t\bar{t}$. Models trained on 500k, 1M events.

CAPITOLO 4. RESULTS

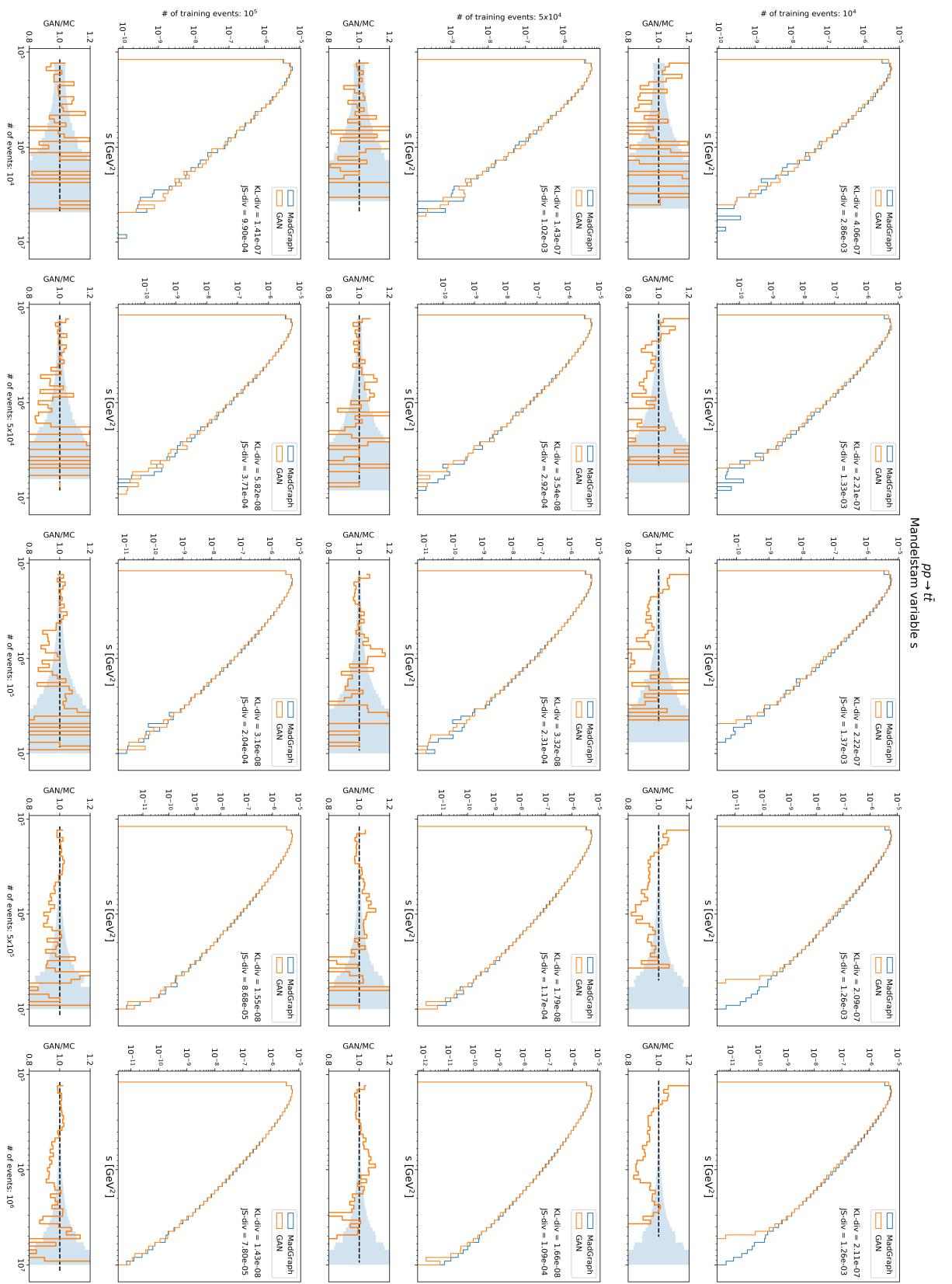


Figura 4.16 (a): Matrix of histograms of the Mandelstam variable s for the channel $pp \rightarrow t\bar{t}$. Models trained on 10k, 50k, 100k events.

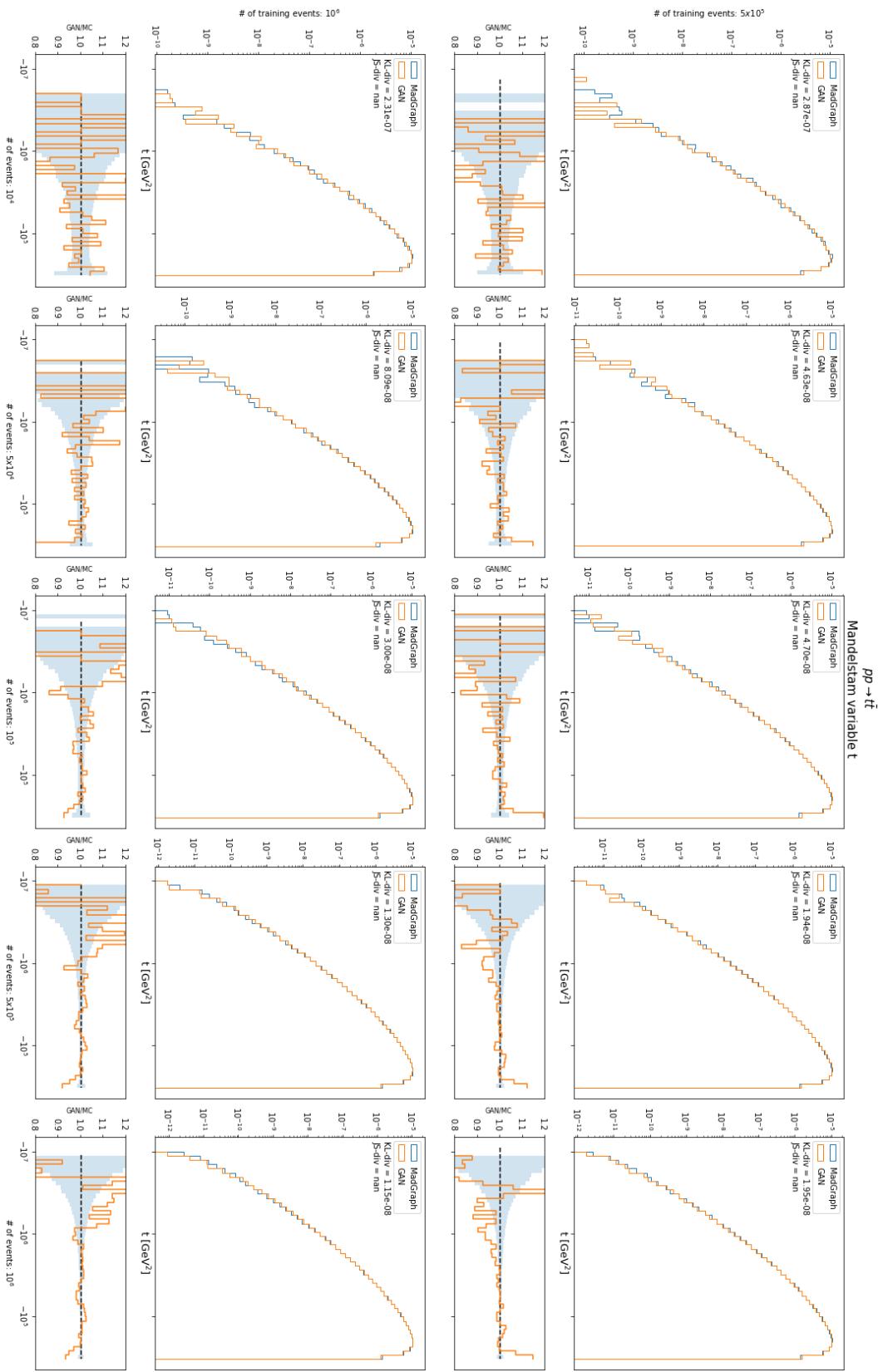


Figura 4.17 (b): Matrix of histograms of the Mandelstam variable t for the channel $pp \rightarrow t\bar{t}$. Models trained on 500k, 1M events.

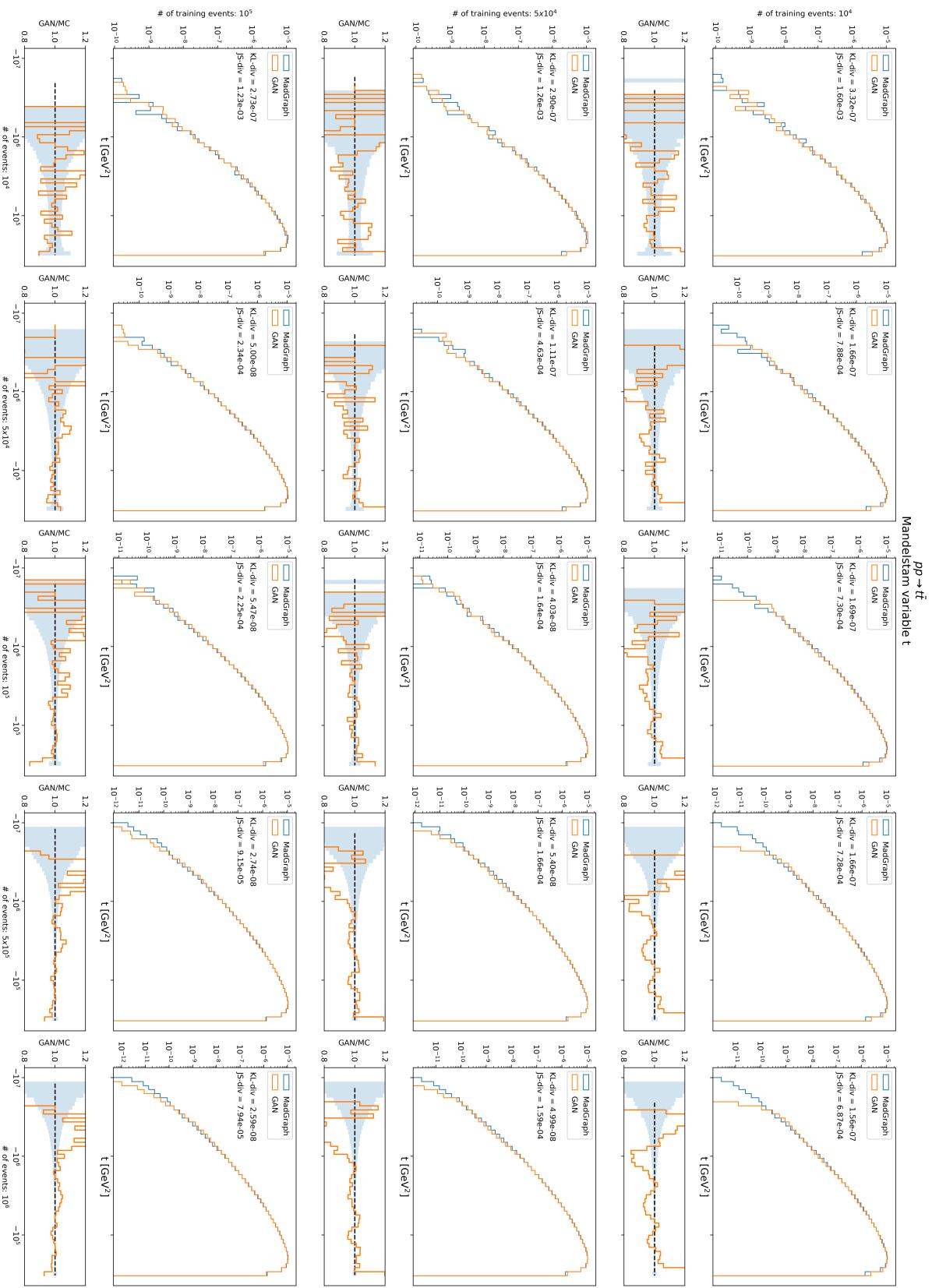


Figura 4.17 (a): Matrix of histograms of the Mandelstam variable t for the channel $pp \rightarrow t\bar{t}$. Models trained on 10k, 50k, 100k events.

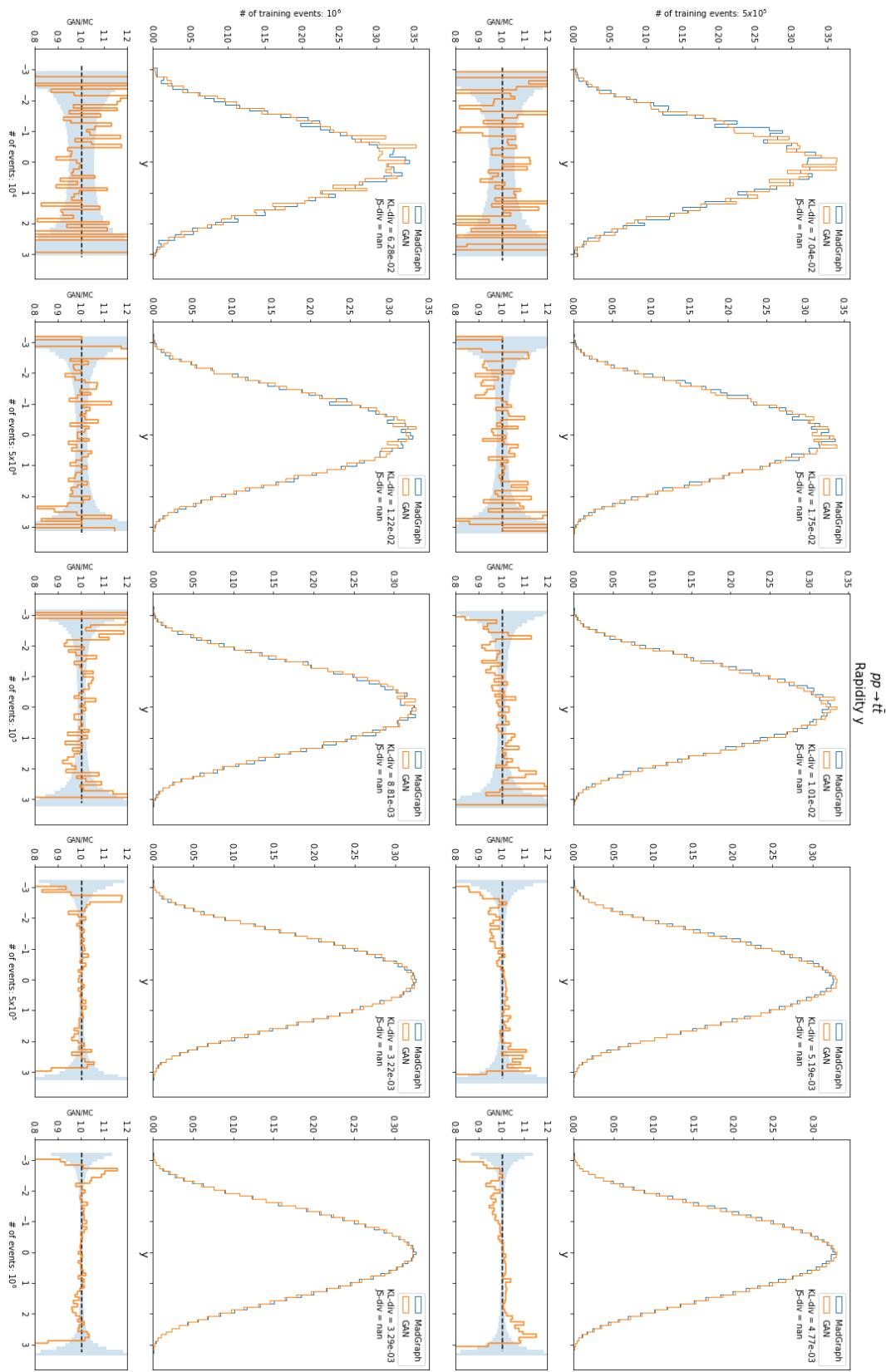


Figura 4.18 (b): Matrix of histograms of the rapidity y for the channel $pp \rightarrow t\bar{t}$. Models trained on 500k, 1M events.

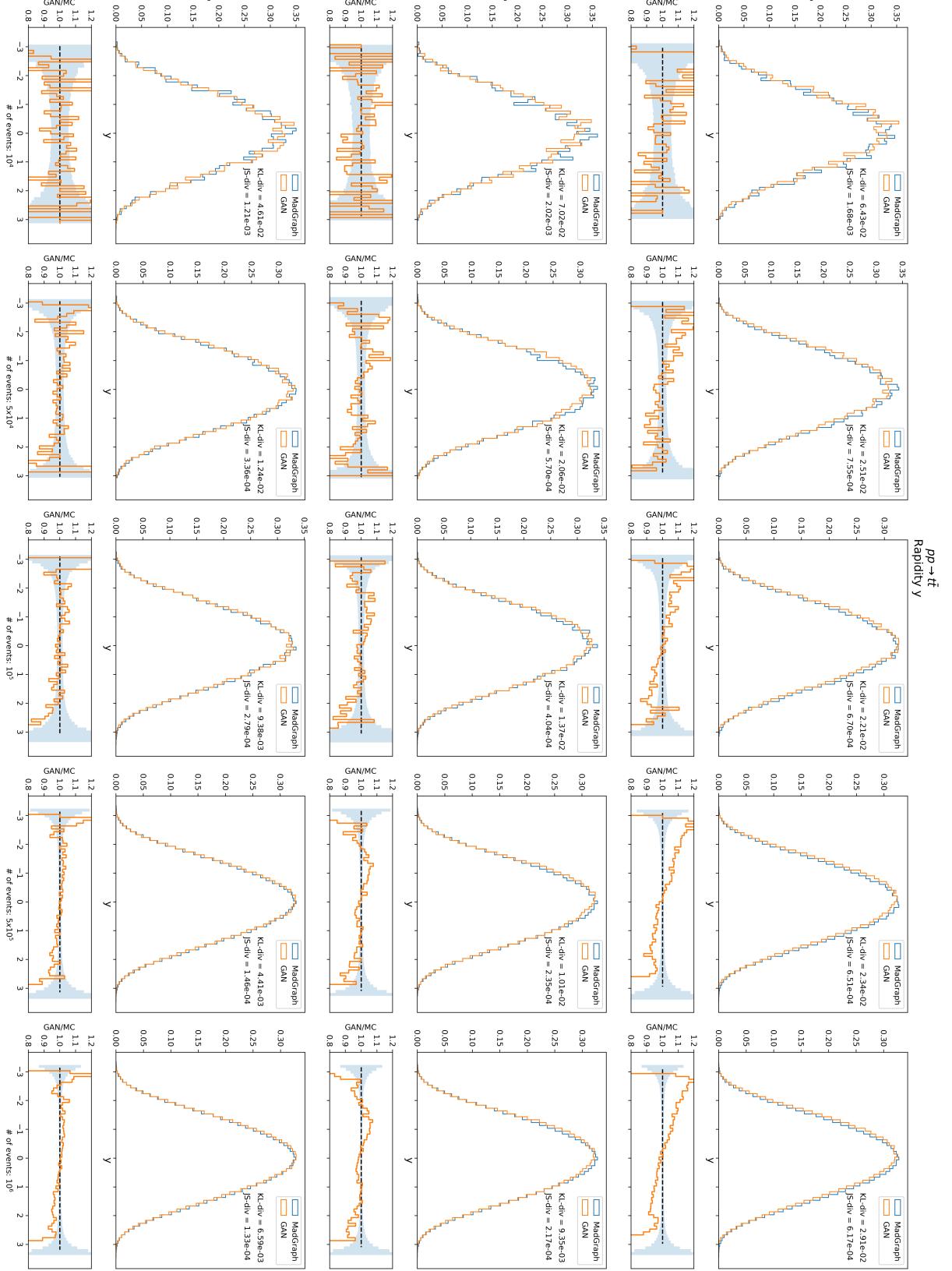


Figura 4.18 (a): Matrix of histograms of the rapidity y for the channel $pp \rightarrow t\bar{t}$. Models trained on 10k, 50k, 100k events.

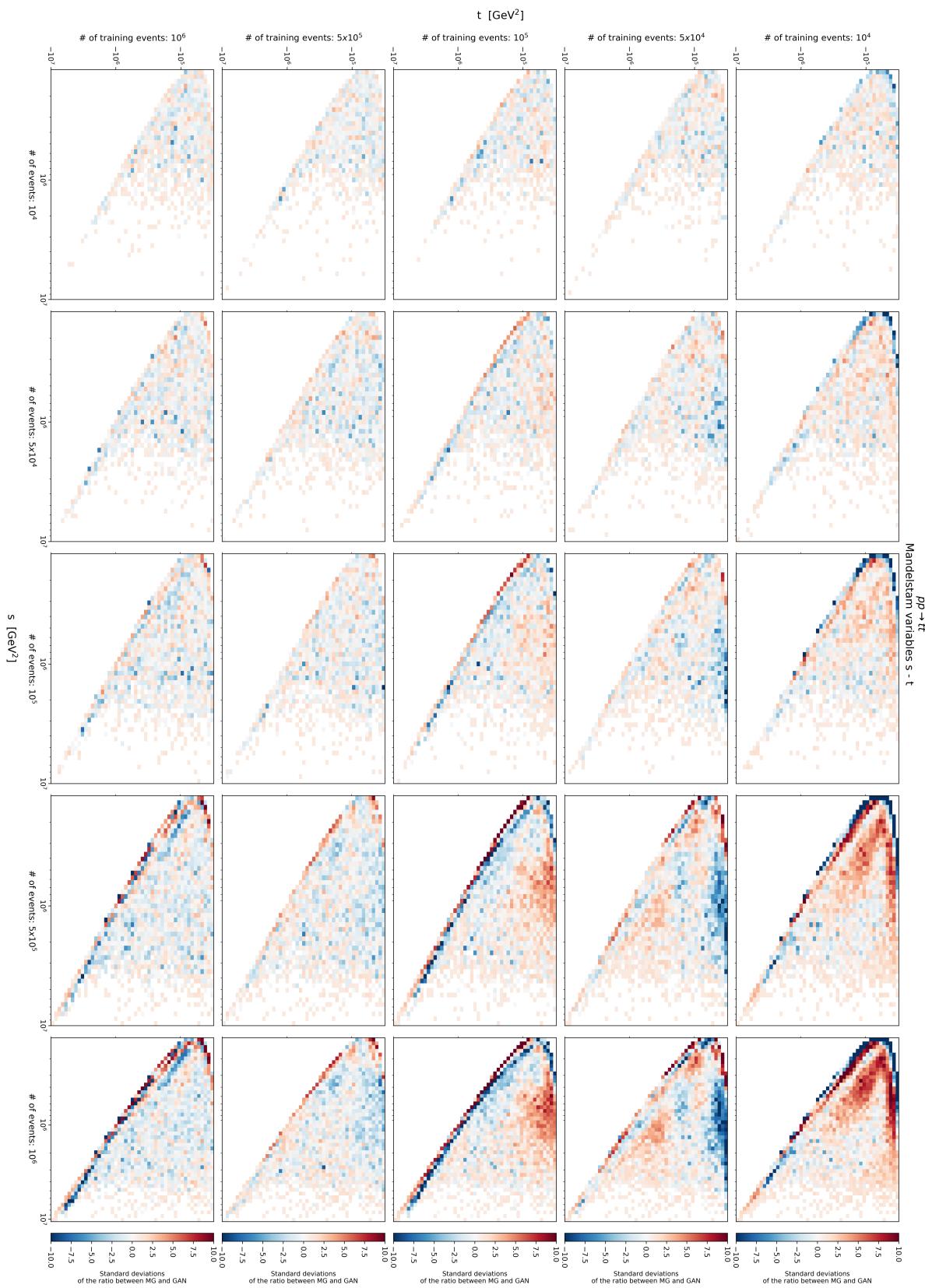


Figura 4.19: Matrix of correlations plot of the Mandelstam variable s for the channel $pp \rightarrow t\bar{t}$.

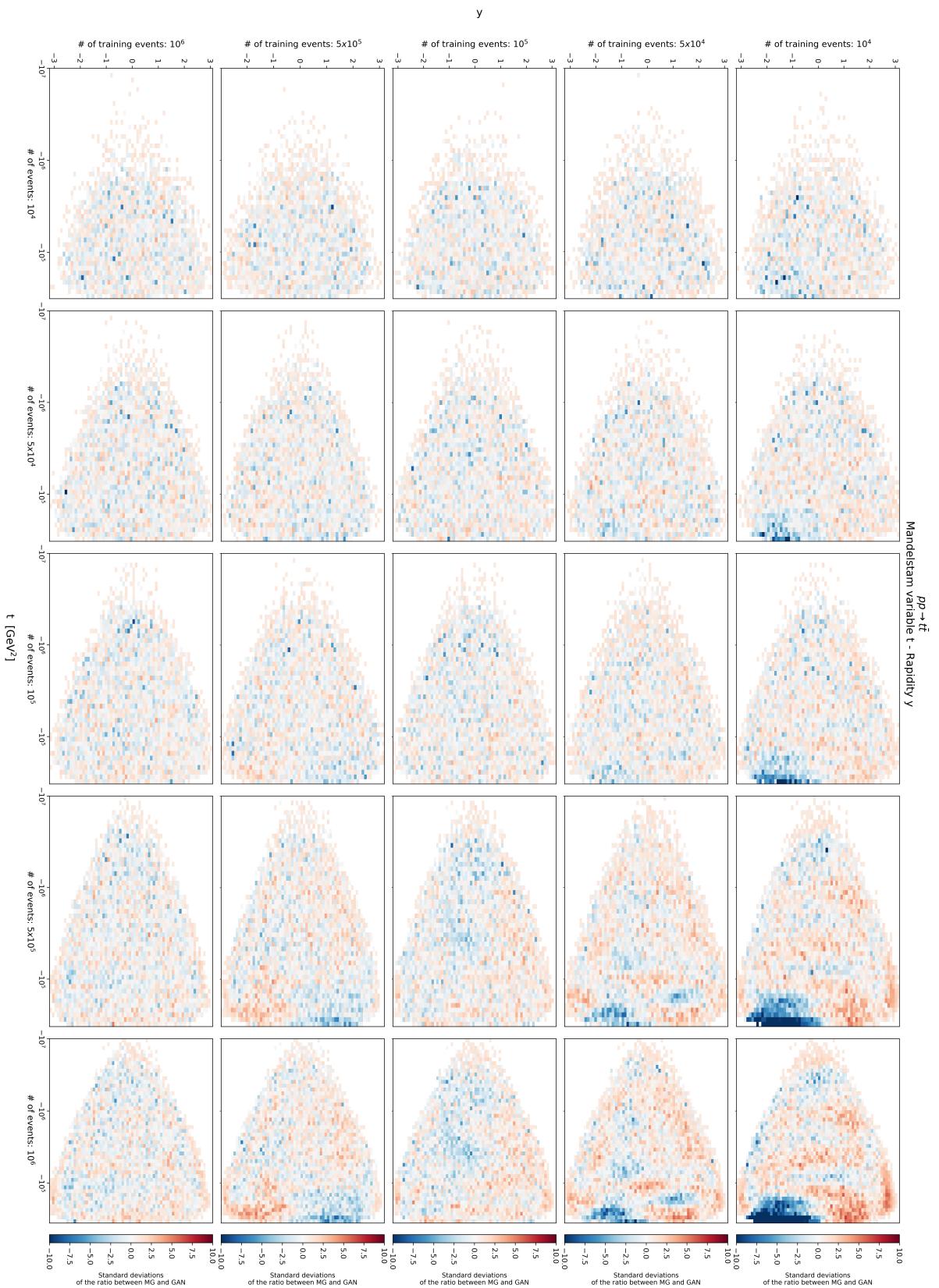


Figura 4.20: Matrix of correlations plot of the Mandelstam variable t for the channel $pp \rightarrow t\bar{t}$.

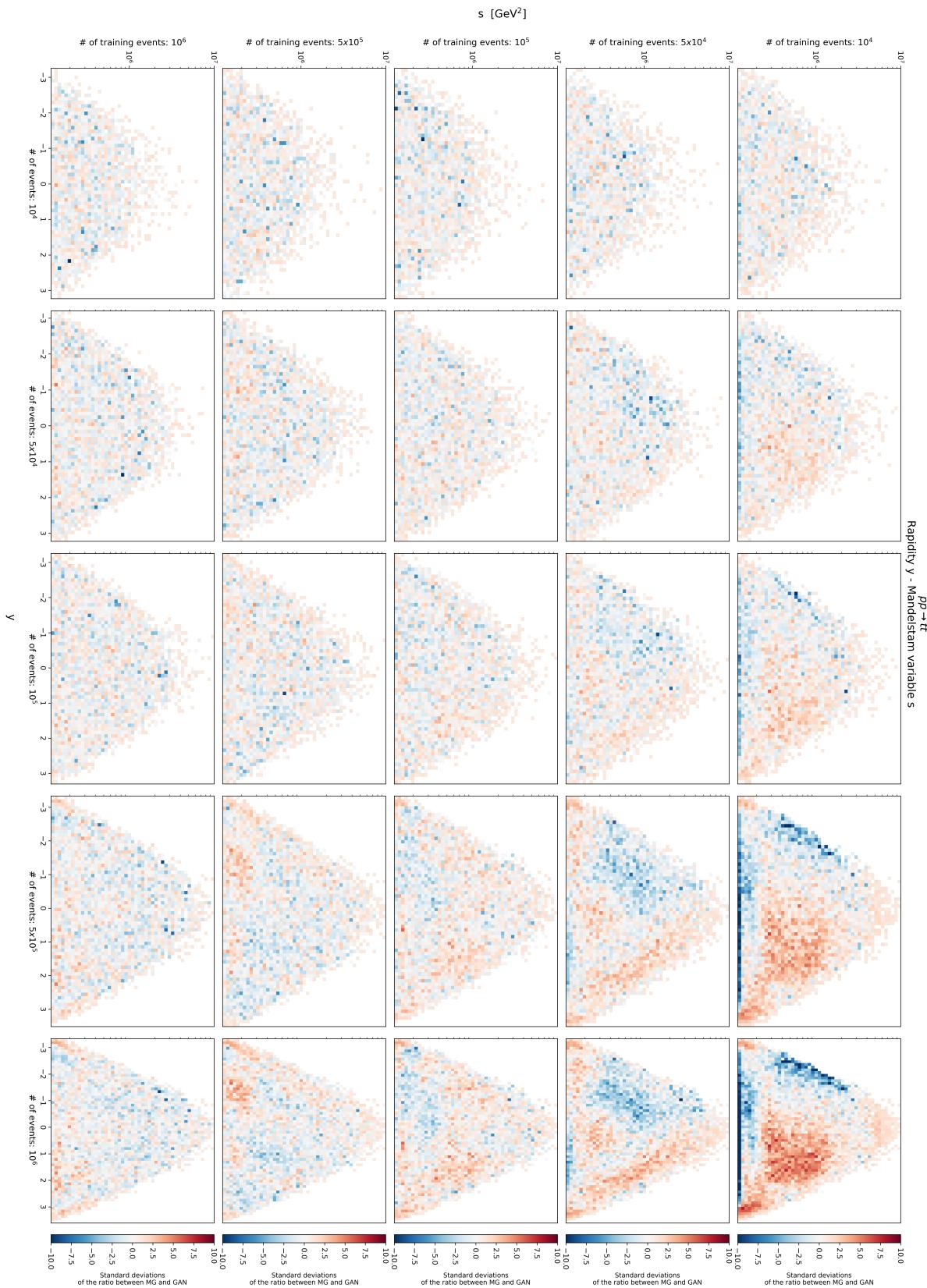


Figura 4.21: Matrix of correlations plot of the rapidity y for the channel $pp \rightarrow t\bar{t}$.

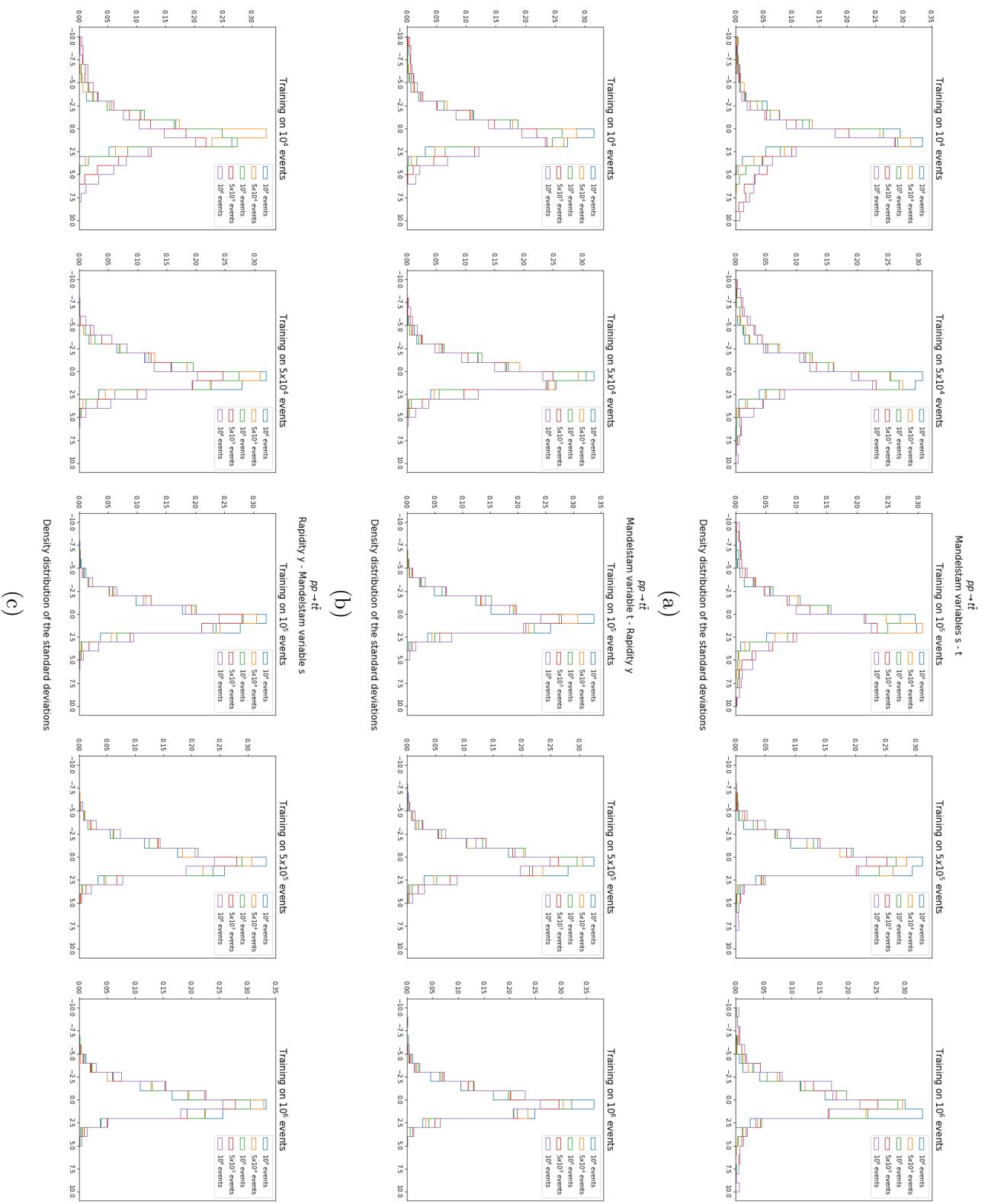


Figura 4.22: Distribution of the errors obtained from the correlations plot for the variable s (a), t (b) and y (c). A single plot correspond to a row of the correlations plot.

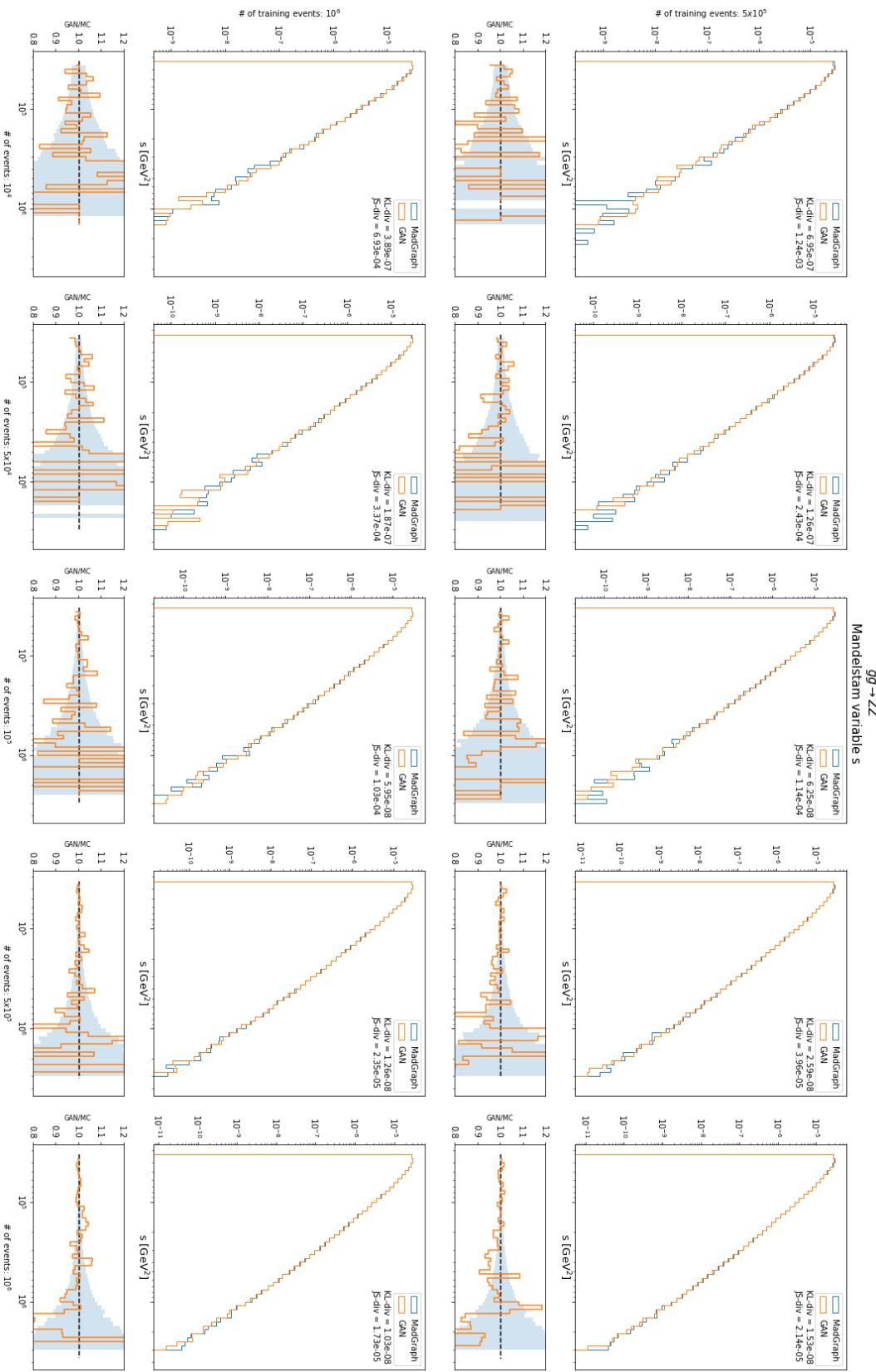


Figura 4.23 (b): Matrix of histograms of the Mandelstam variable s for the channel $gg \rightarrow ZZ$. Models trained on 500k, 1M events.

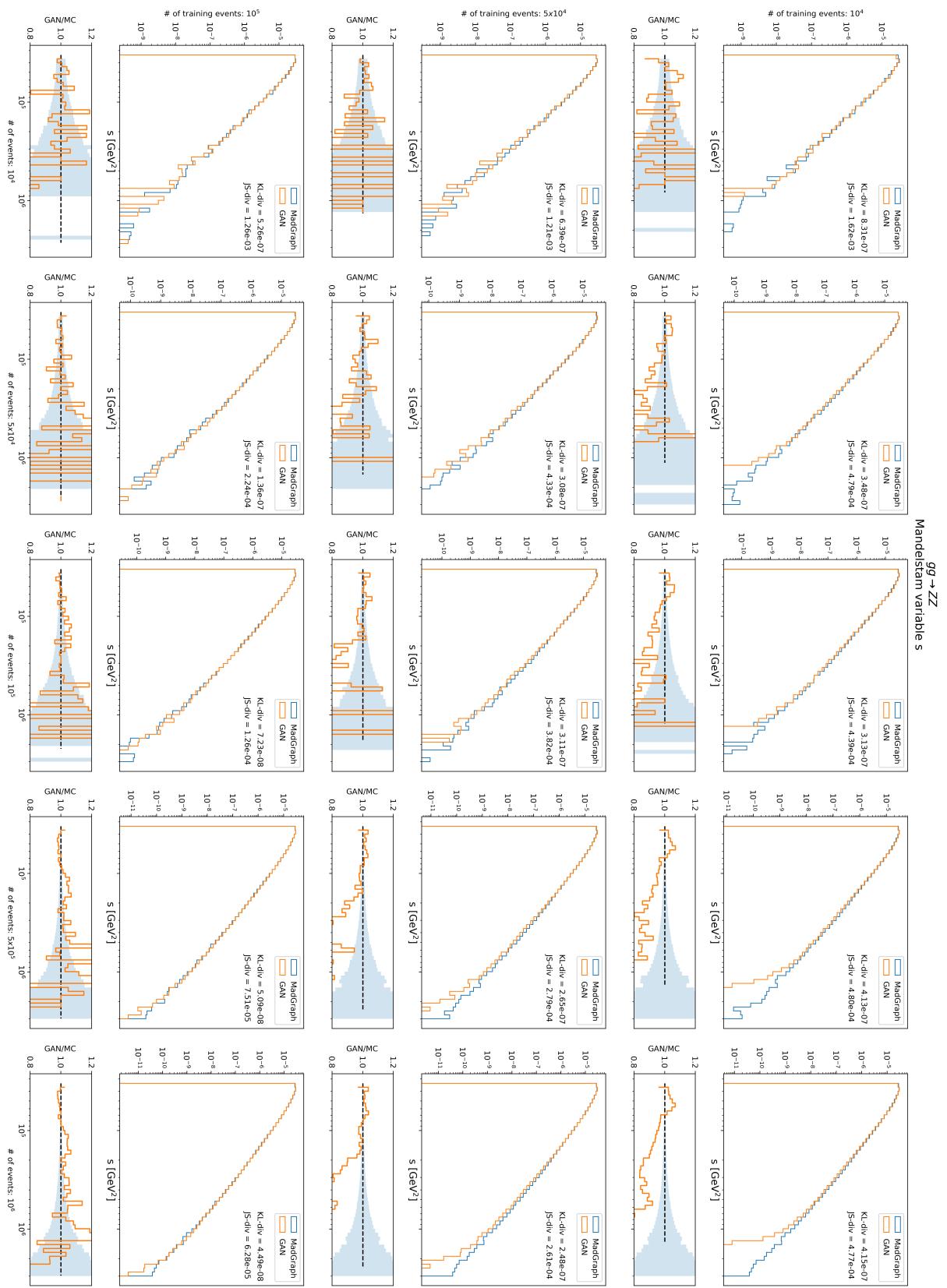


Figura 4.23 (a): Matrix of histograms of the Mandelstam variable s for the channel $gg \rightarrow ZZ$. Models trained on 10k, 50k, 100k events.

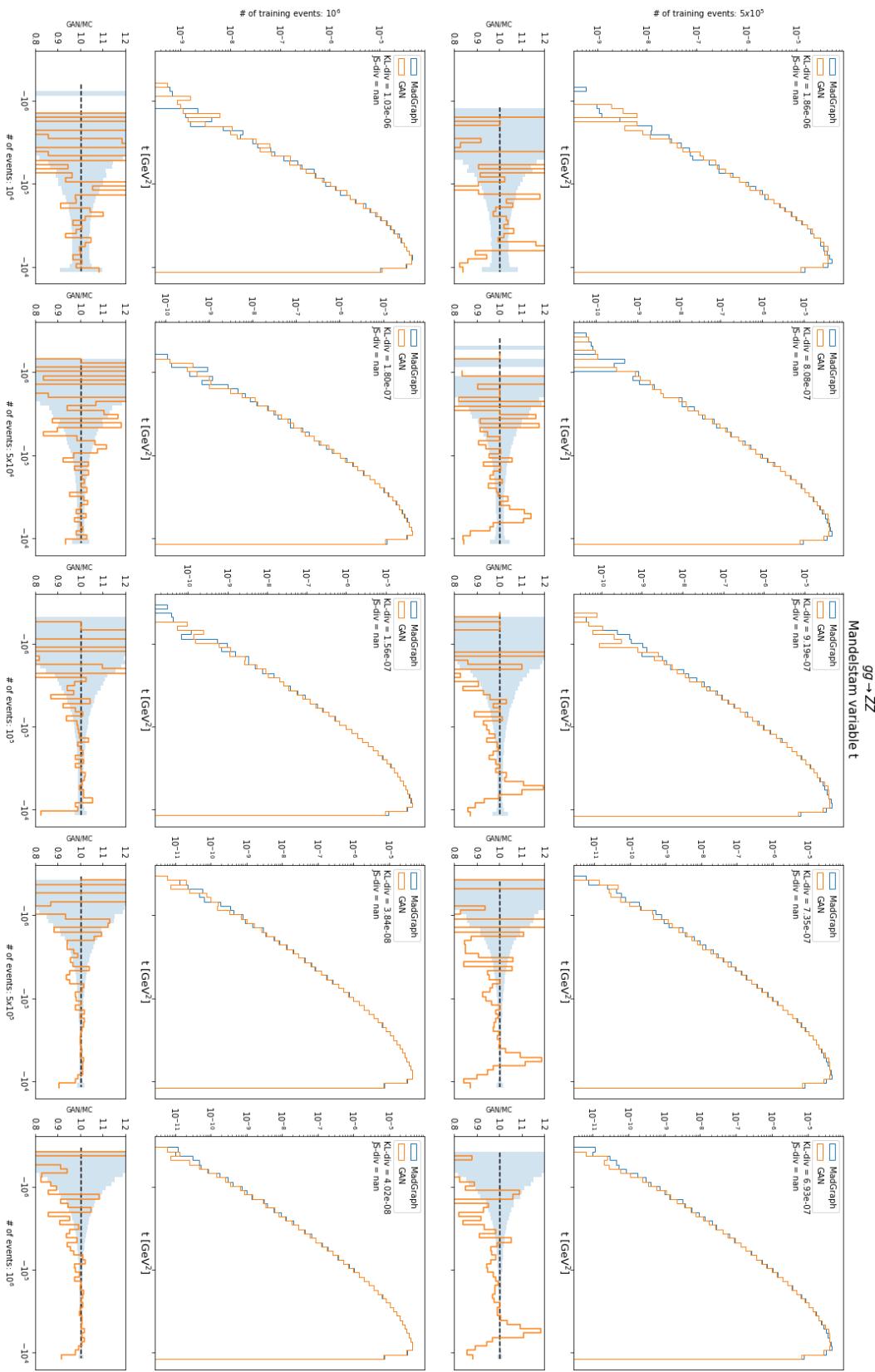


Figura 4.24 (b): Matrix of histograms of the Mandelstam variable t for the channel $gg \rightarrow ZZ$. Models trained on 500k, 1M events.

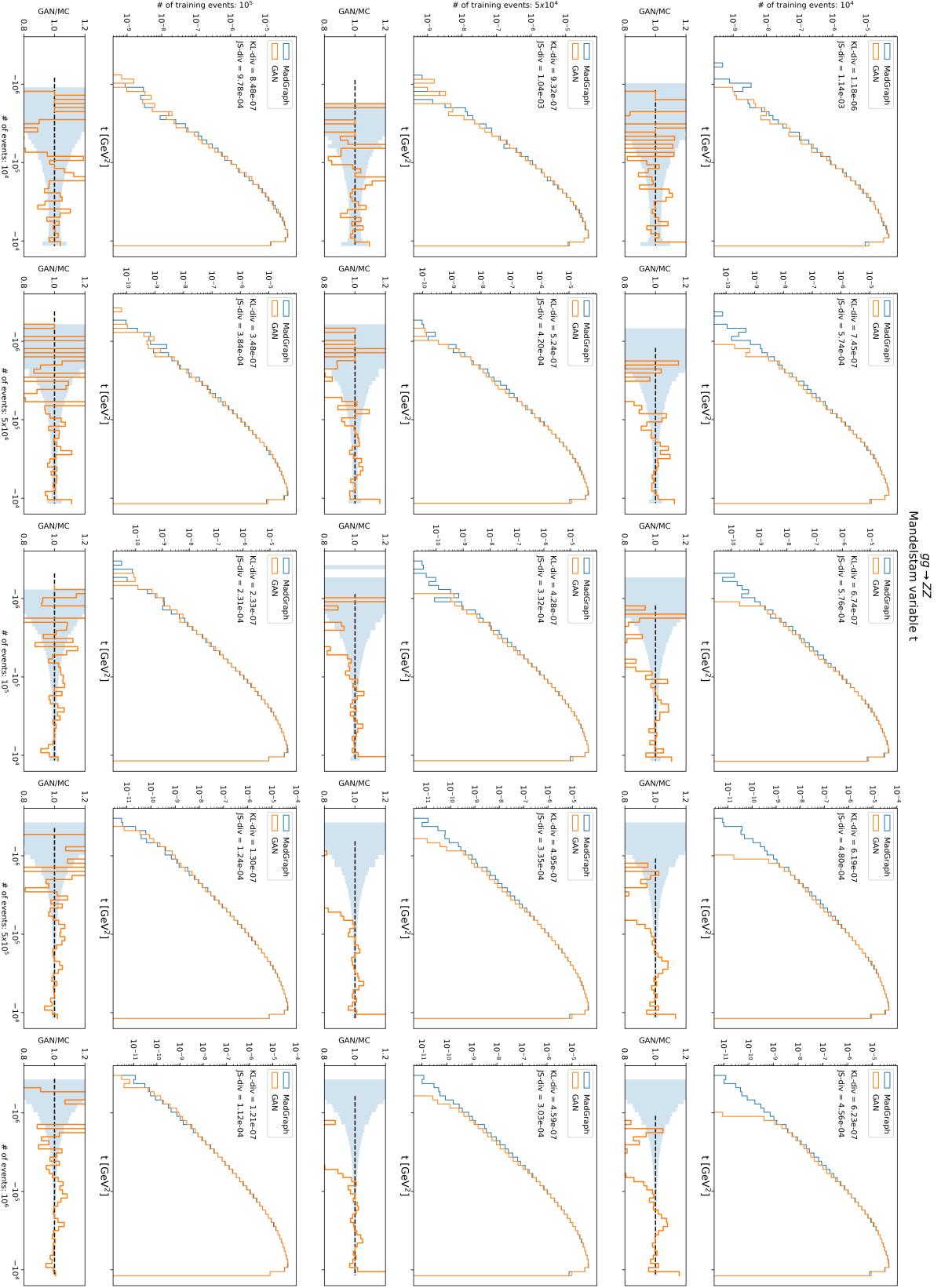


Figura 4.24 (a): Matrix of histograms of the Mandelstam variable t for the channel $gg \rightarrow ZZ$. Models trained on 10k, 50k, 100k events.

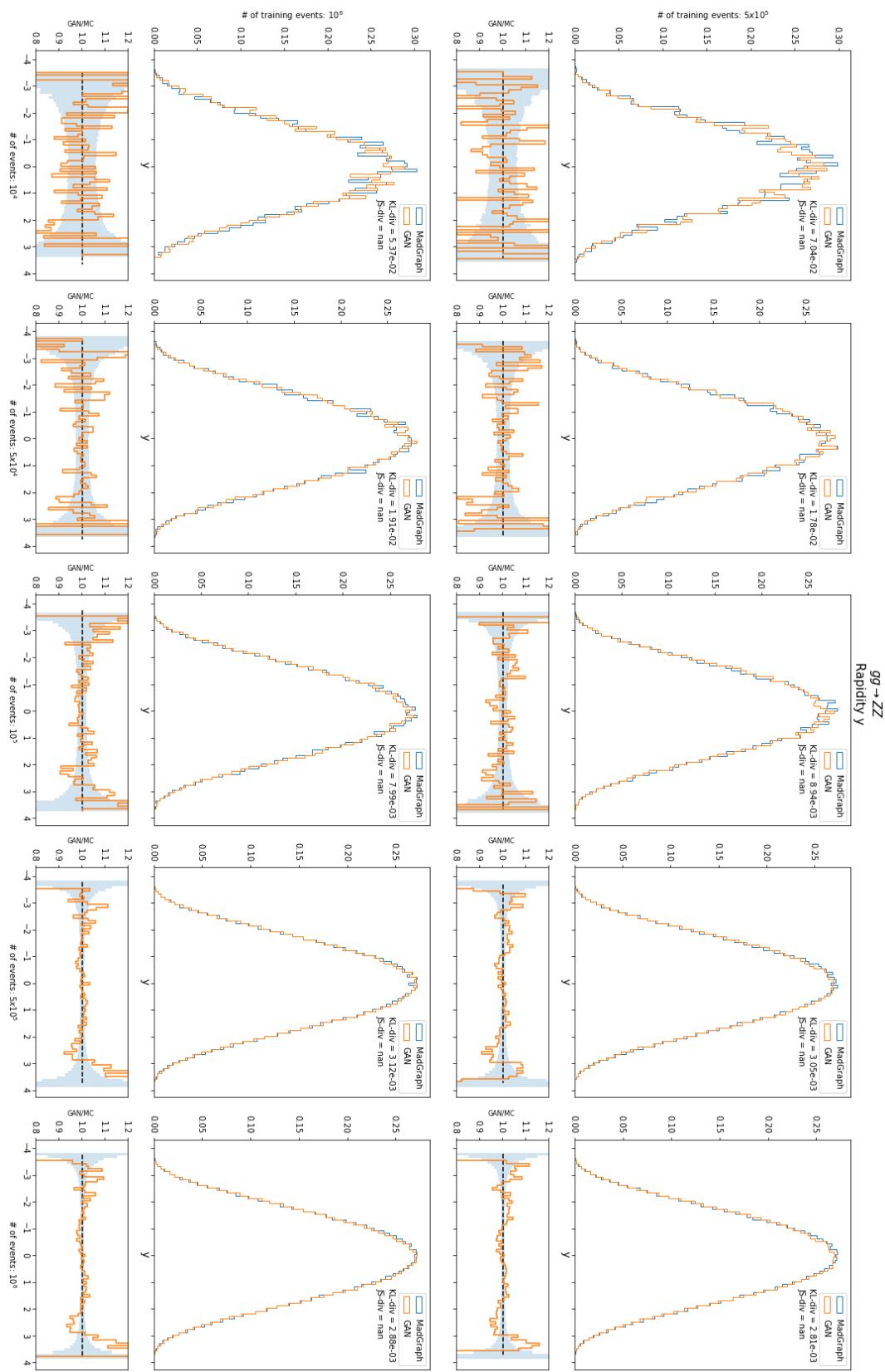


Figura 4.25 (b): Matrix of histograms of the rapidity y for the channel $gg \rightarrow ZZ$. Models trained on 500k, 1M events.

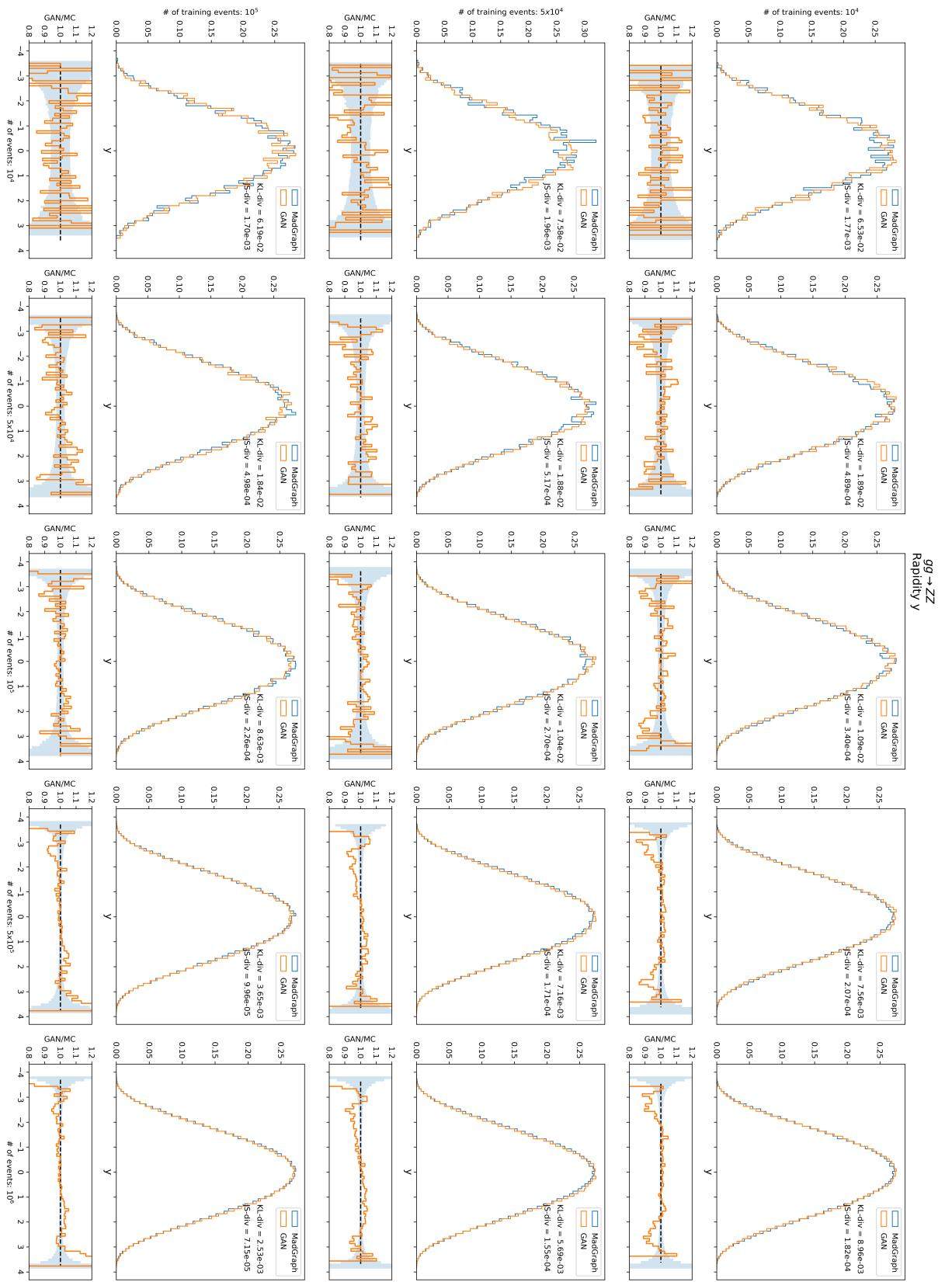


Figura 4.25 (a): Matrix of histograms of the rapidity y for the channel $gg \rightarrow ZZ$. Models trained on 10k, 50k, 100k events.

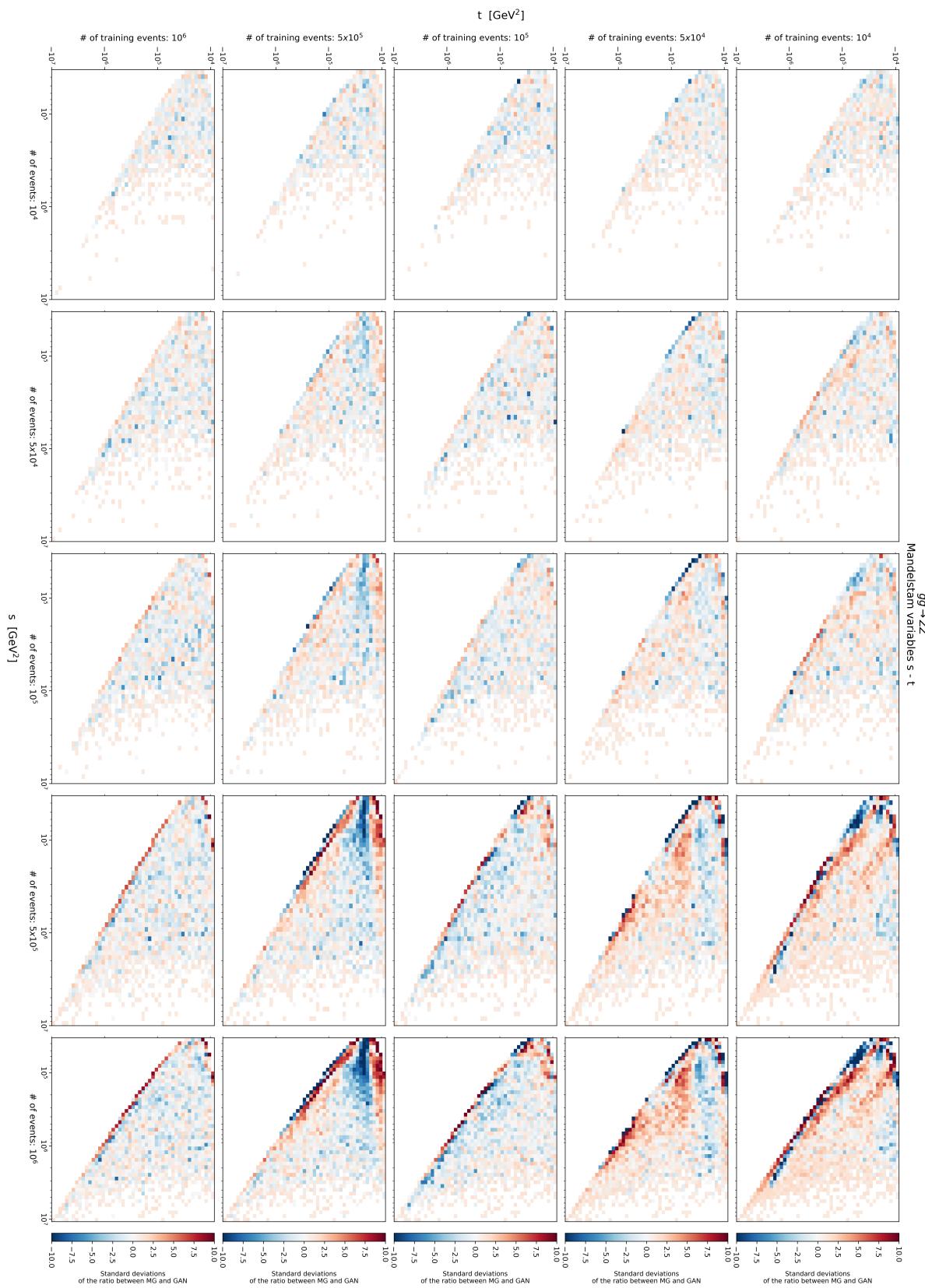


Figura 4.26: Matrix of correlations plot of the Mandelstam variable s for the channel $gg \rightarrow ZZ$.

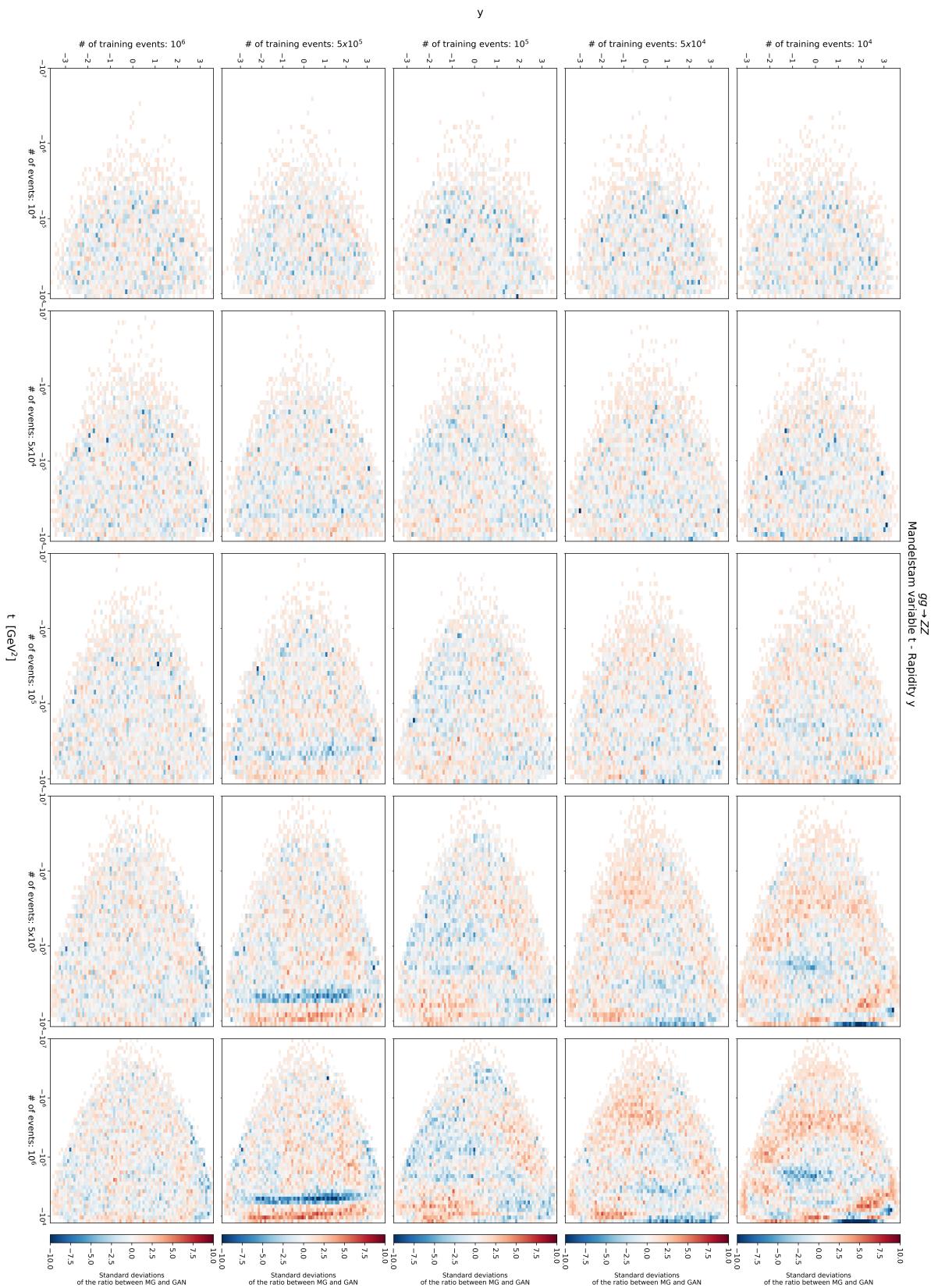


Figura 4.27: Matrix of correlations plot of the Mandelstam variable t for the channel $gg \rightarrow ZZ$.

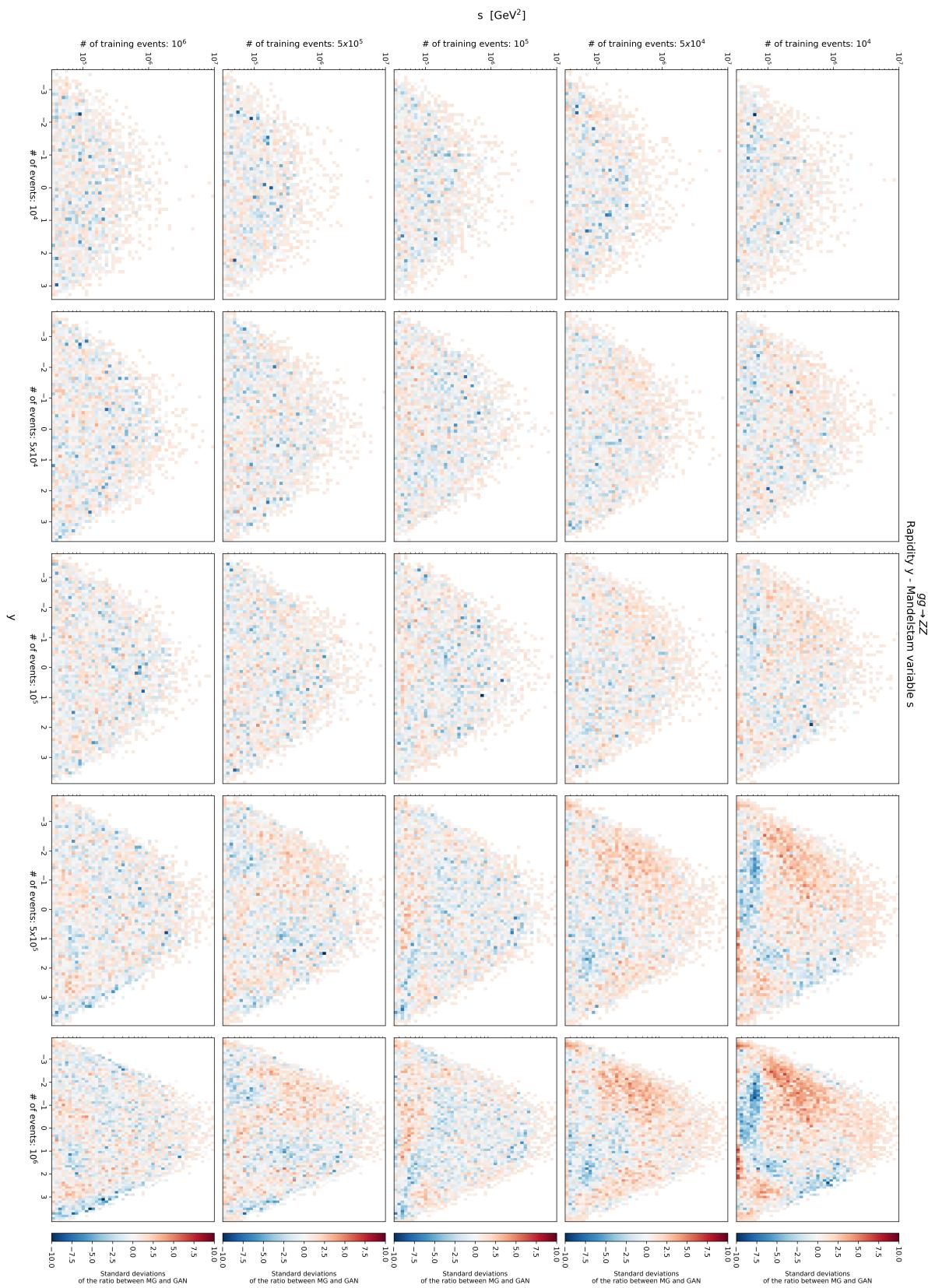


Figura 4.28: Matrix of correlations plot of the rapidity y for the channel $gg \rightarrow ZZ$.

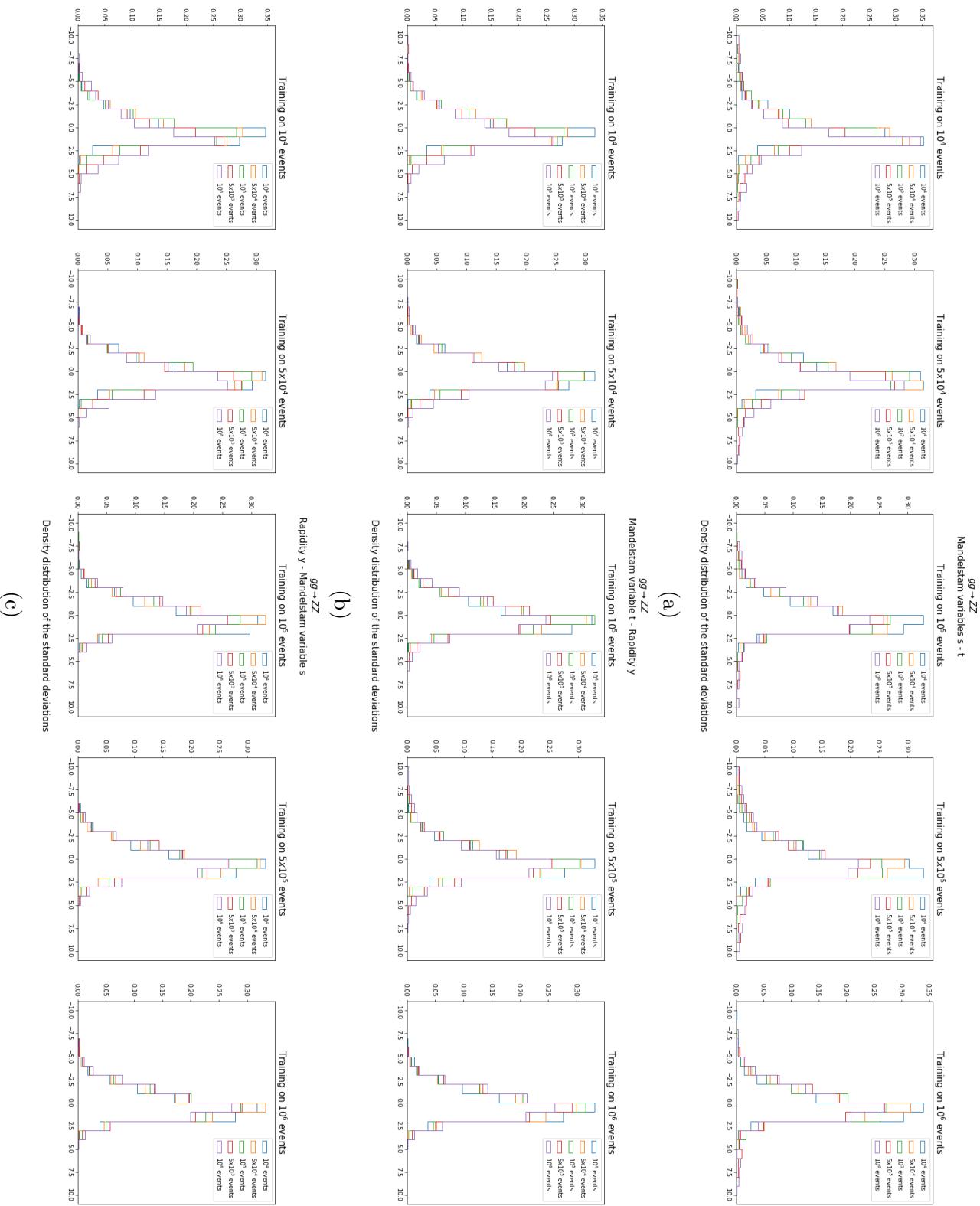


Figura 4.29: Distribution of the errors obtained from the correlations plot for the variable s (a), t (b) and y (c). A single plot correspond to a row of the correlations plot.

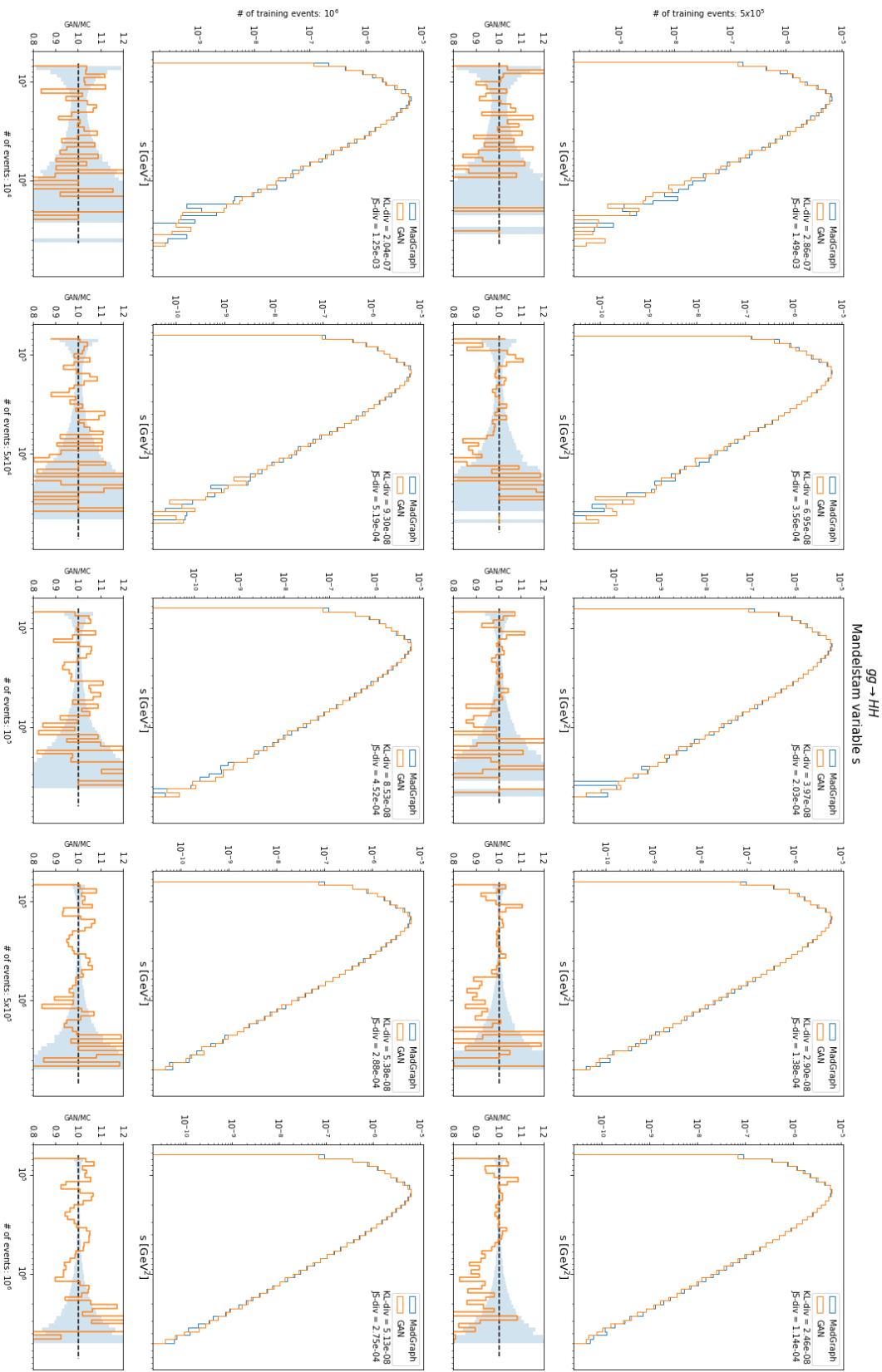


Figura 4.30 (b): Matrix of histograms of the Mandelstam variable s for the channel $gg \rightarrow HH$. Models trained on 500k, 1M events.

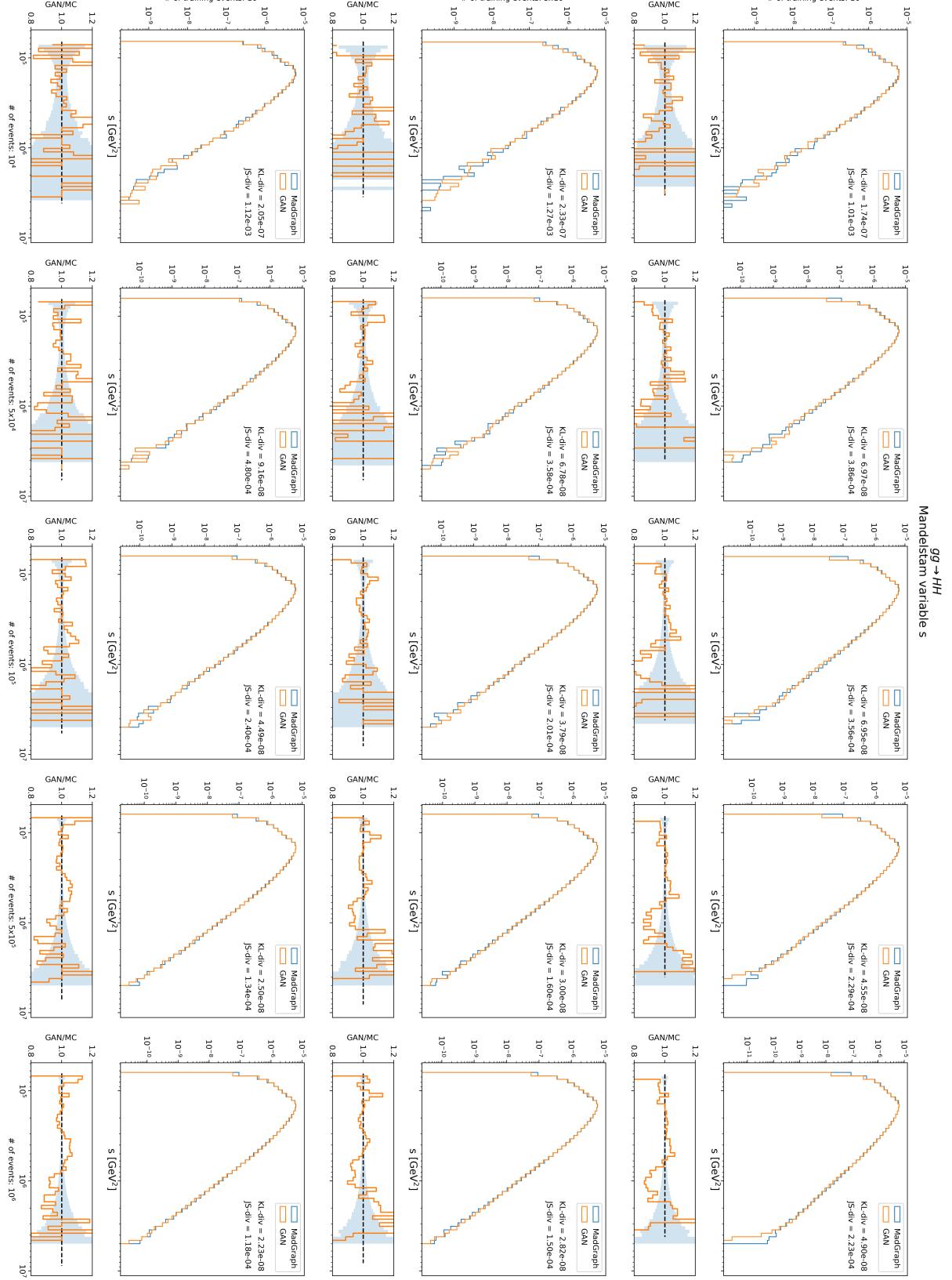


Figura 4.30 (a): Matrix of histograms of the Mandelstam variable s for the channel $gg \rightarrow HH$. Models trained on $10k$, $50k$, $100k$ events.

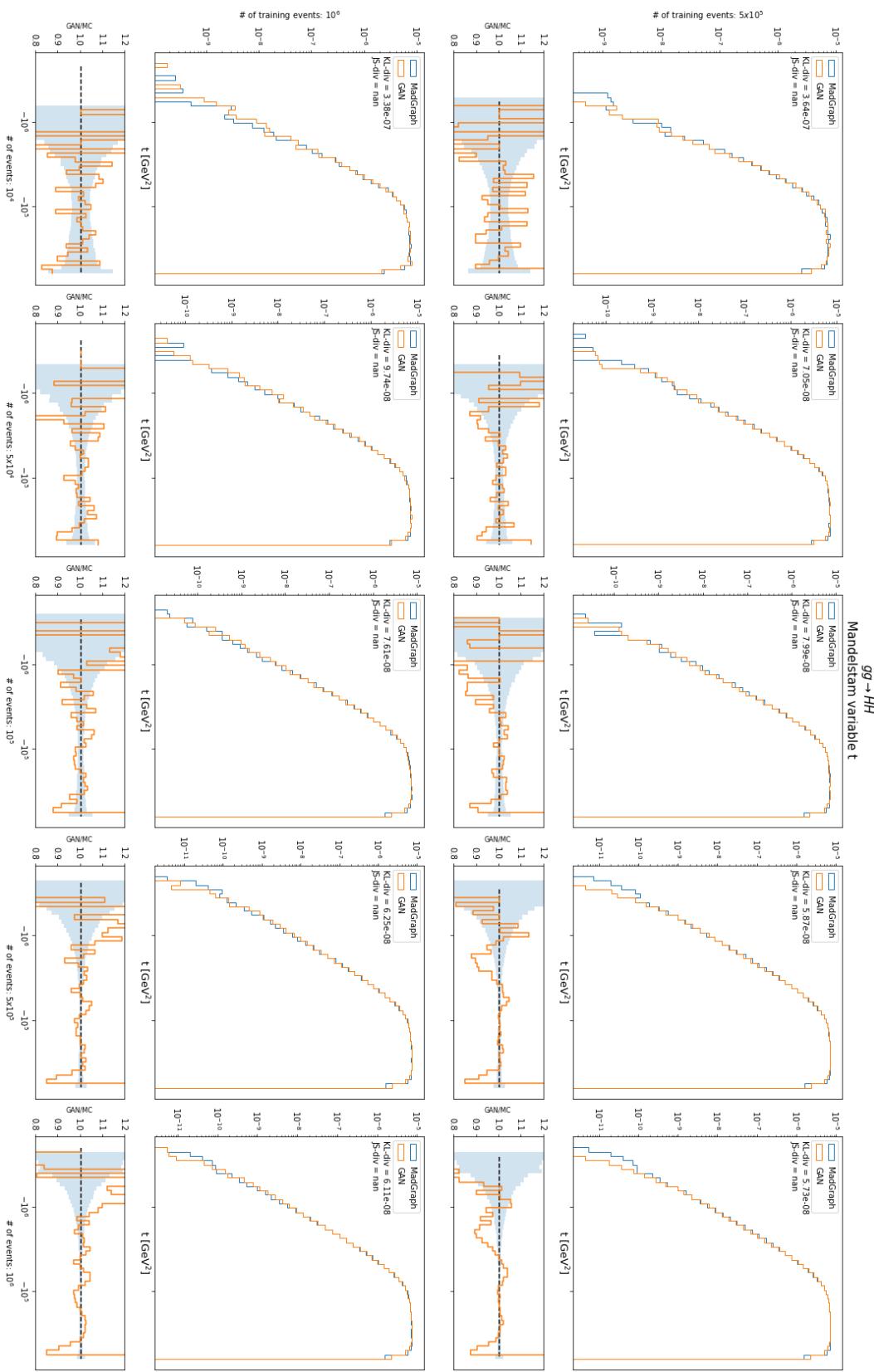


Figura 4.31 (b): Matrix of histograms of the Mandelstam variable t for the channel $gg \rightarrow HH$. Models trained on 500k, 1M events.

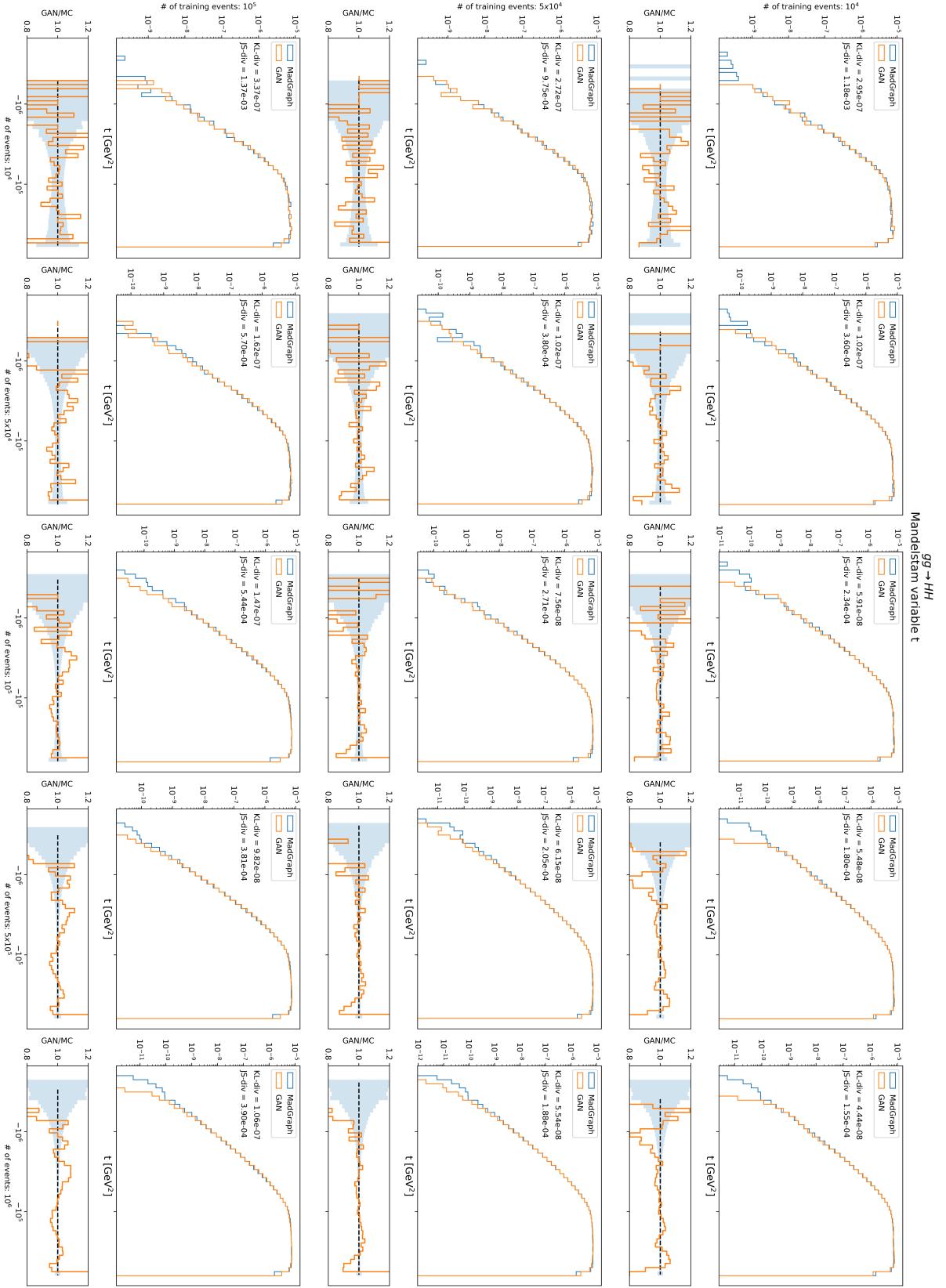


Figura 4.31 (a): Matrix of histograms of the Mandelstam variable t for the channel $gg \rightarrow HH$. Models trained on 10k, 50k, 100k events.

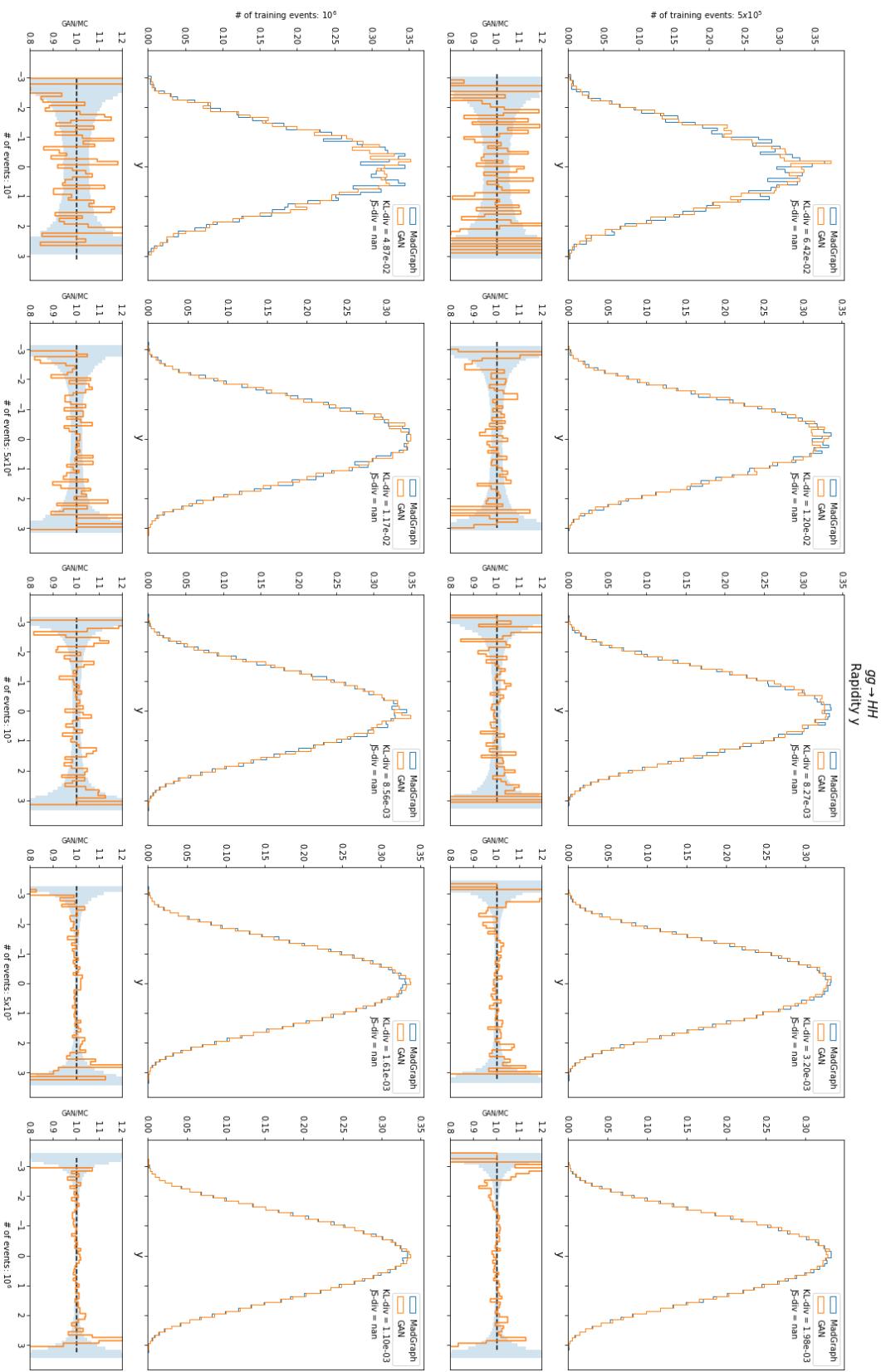


Figura 4.32 (b): Matrix of histograms of the rapidity y for the channel $gg \rightarrow HH$. Models trained on 500k, 1M events.

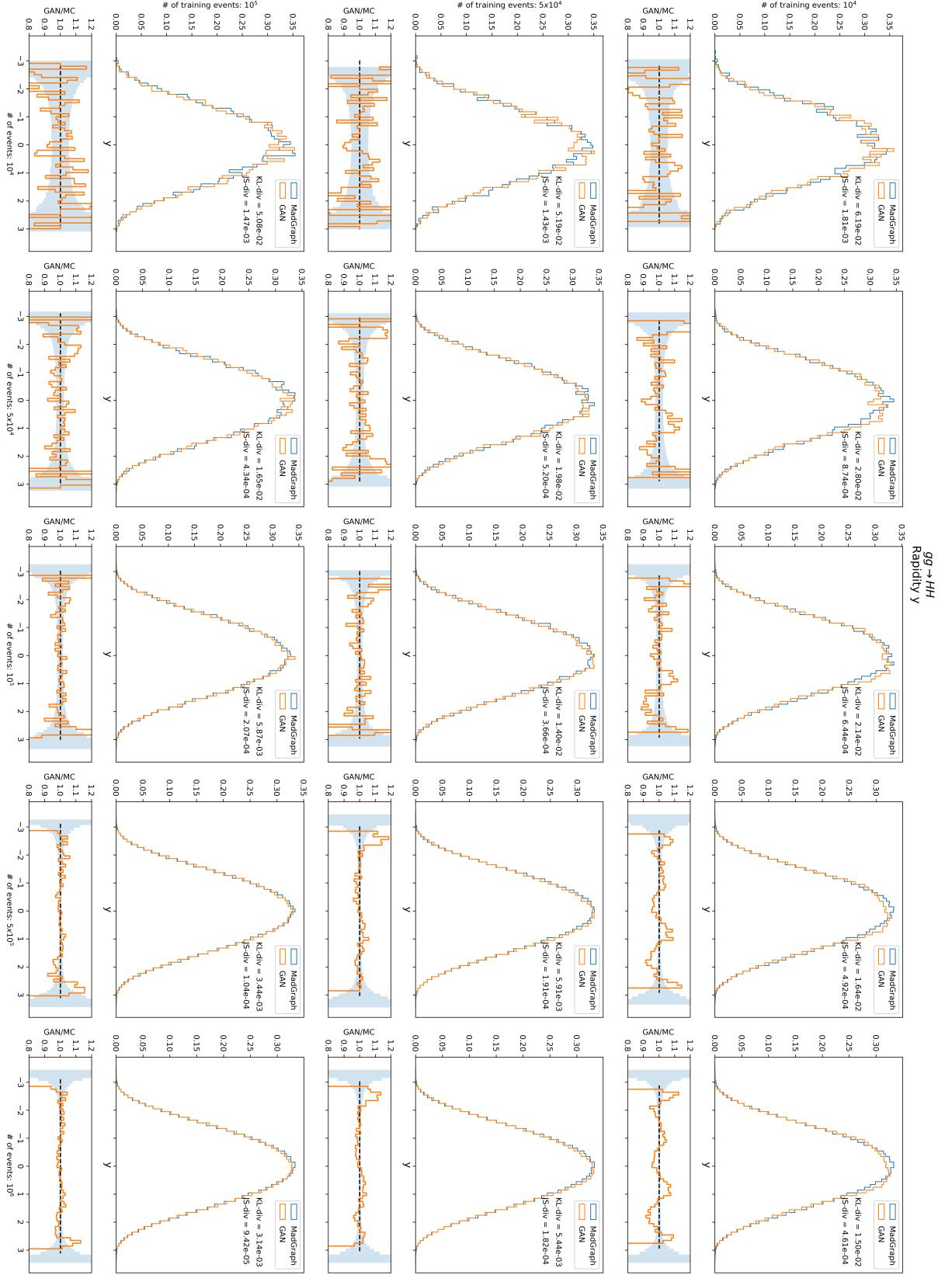


Figura 4.32 (a): Matrix of histograms of the rapidity y for the channel $gg \rightarrow HH$. Models trained on 10k, 50k, 100k events.

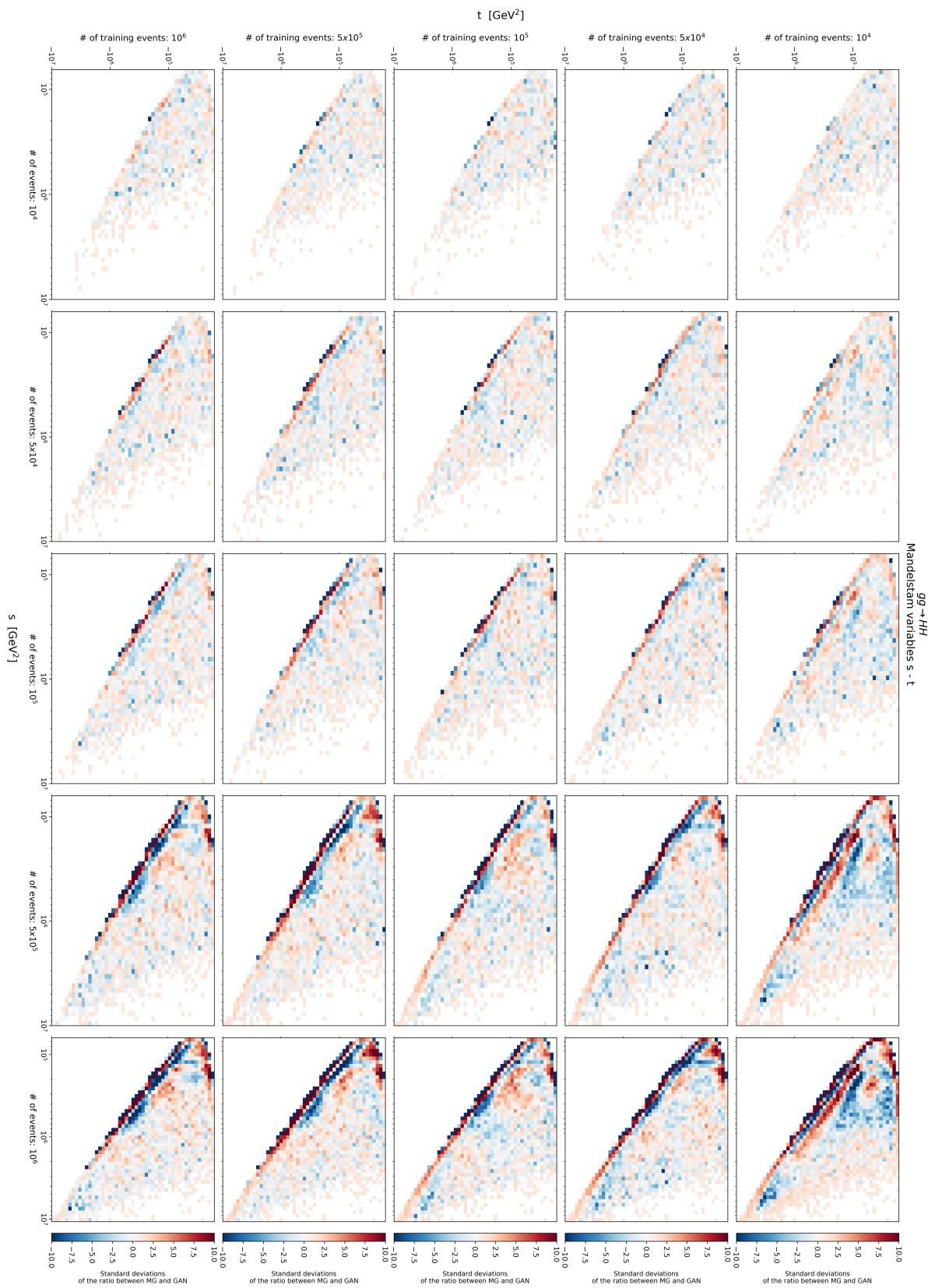


Figura 4.33: Matrix of correlations plot of the Mandelstam variable s for the channel $gg \rightarrow HH$.

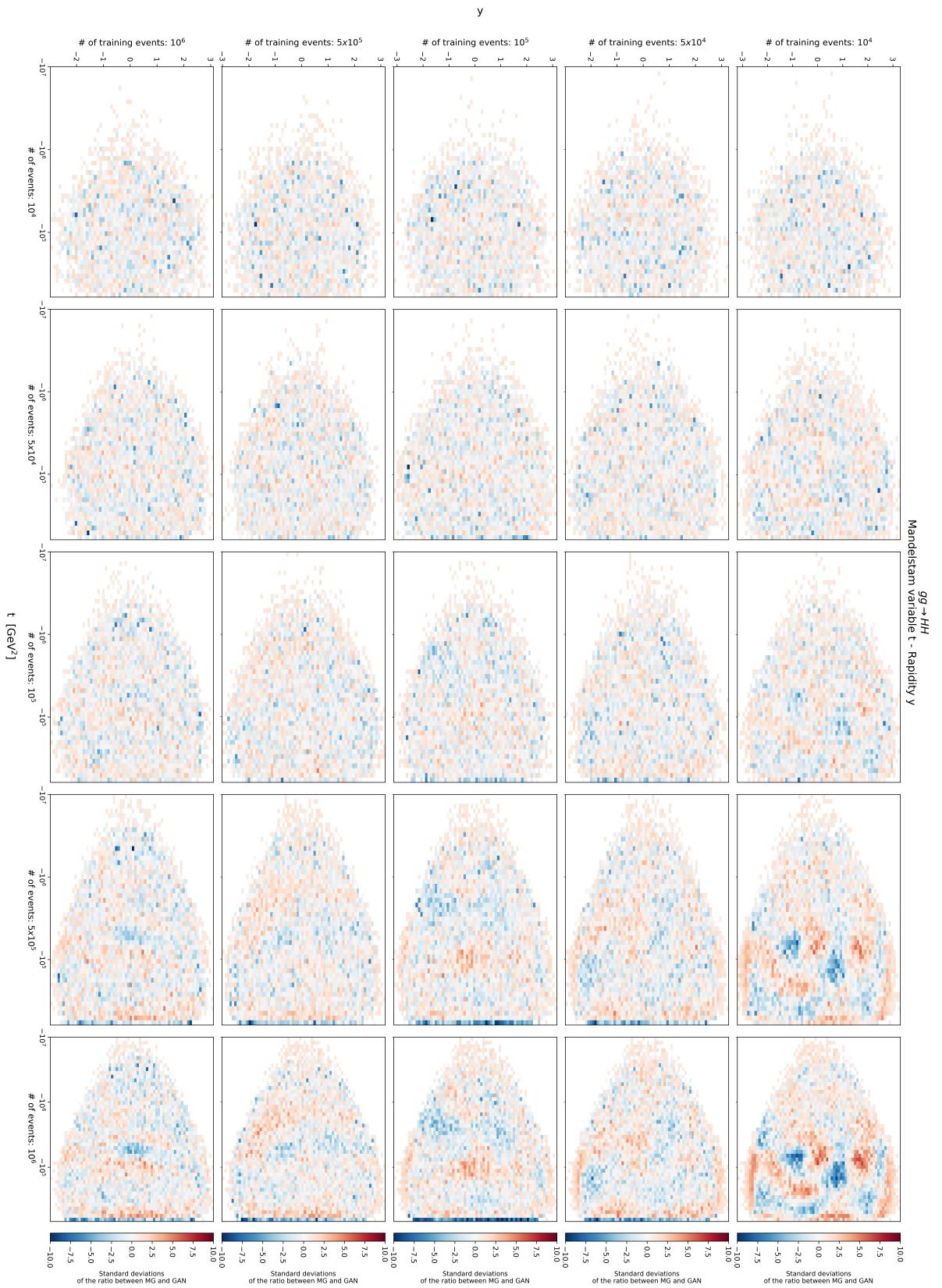


Figura 4.34: Matrix of correlations plot of the Mandelstam variable t for the channel $gg \rightarrow HH$.



Figura 4.35: Matrix of correlations plot of the rapidity y for the channel $gg \rightarrow HH$.

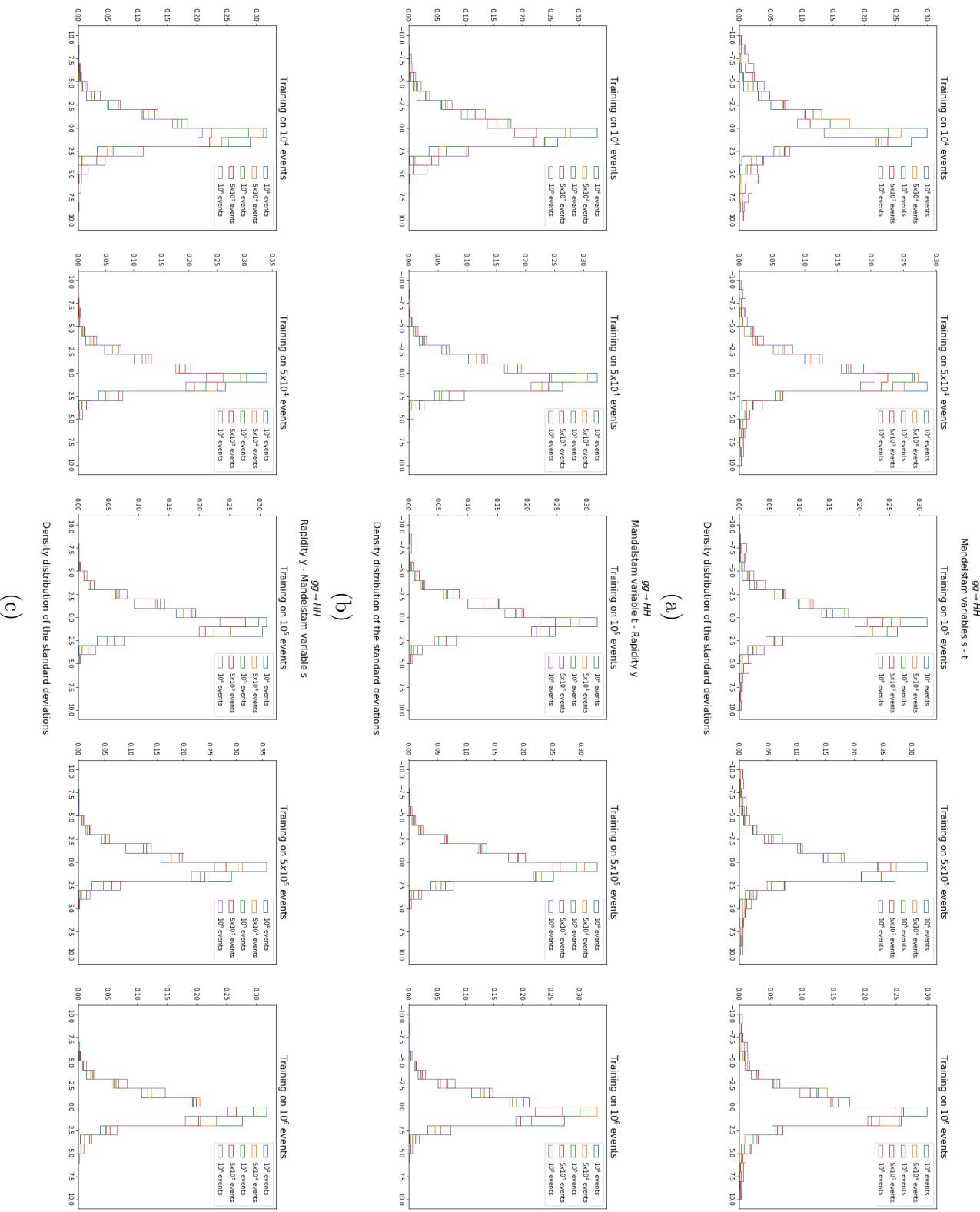


Figura 4.36: Distribution of the errors obtained from the correlations plot for the variable s (a), t (b) and y (c). A single plot correspond to a row of the correlations plot.