Adversarial Network in the search for SUSY in events with one lepton and multiple jets in proton-proton collisions

5th September 2019

Summer Student: Farouk Mokhtar

Zewail City of Science and Technology, Egypt

Supervised by: Ashraf Mohamed, Dirk Kruecker, Isabell Melzer-Pellmann

HELMHOLTZ SPITZENFORSCHUNG FÜR GROSSE HERAUSFORDERUNGEN

DESY.

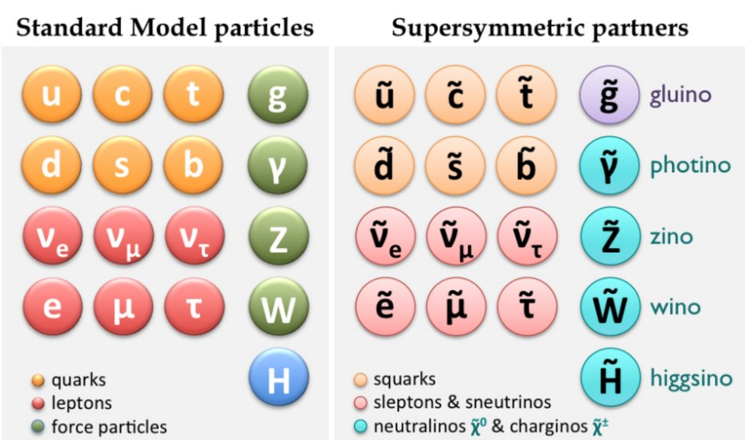
Supersymmetry (SUSY)

SM

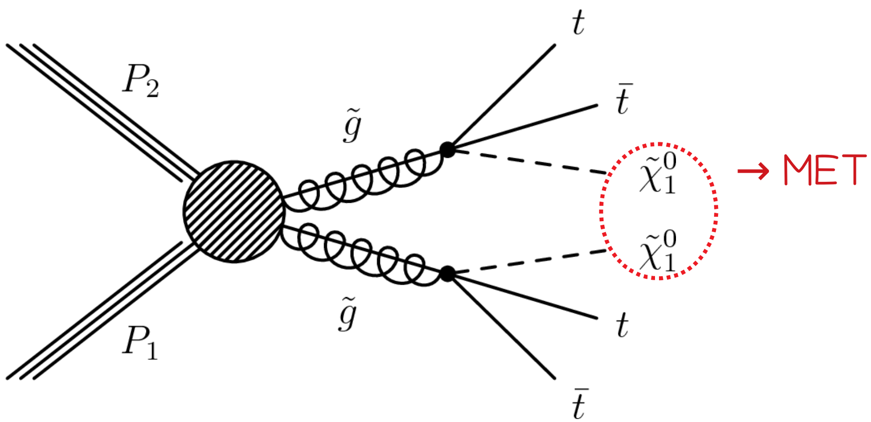
- Most successful model but has limitations like:
  (1) Hierarchy Problem (2) Dark Matter …

SUSY

- New space-time symmetry
- Extension to SM which add extra degrees of freedom
- Each fermion have supersymmetric partner~ boson (and vice-versa)
- Some hypothetical particles provide solution to (1)
- LSP ($\tilde{\chi}_0$) is a candidate for (2)

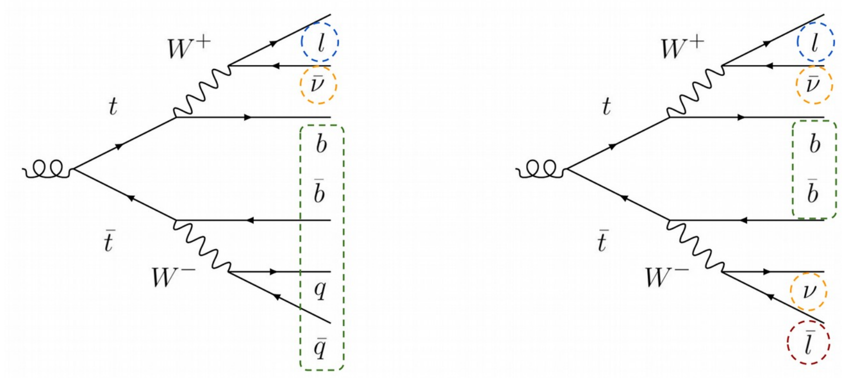


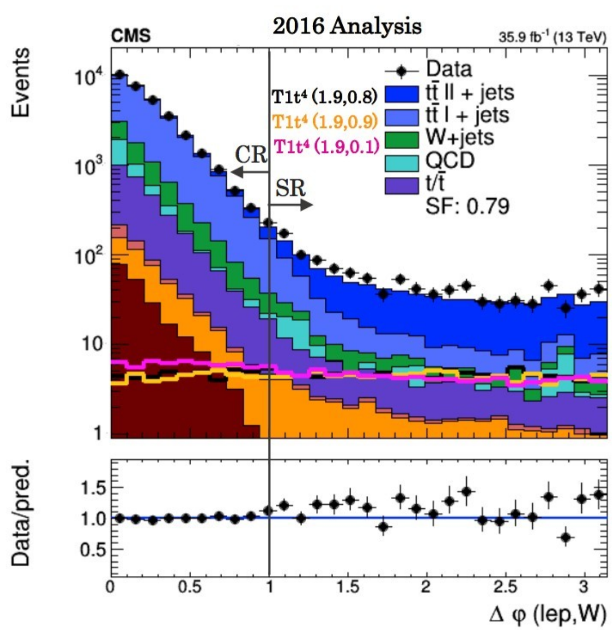
Physics Problem: SUSY 1-Lepton



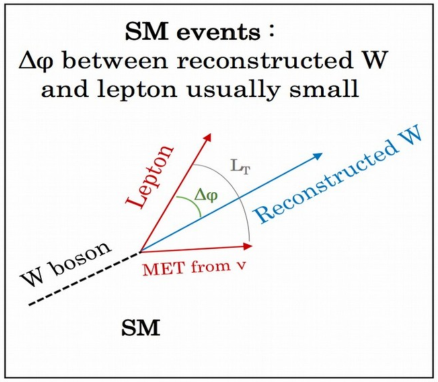
tt - semileptonic:

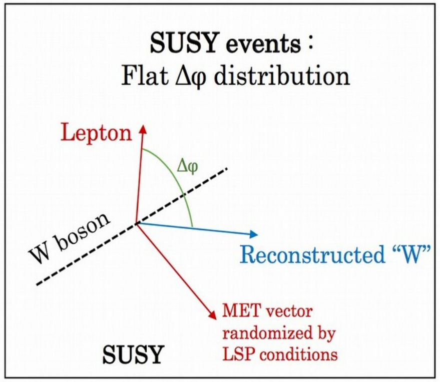
tt - dileptonic:





Main Search Variable: Δφ

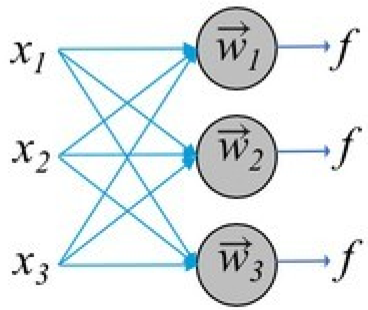




Target → Improve on the old analysis by using a Deep Neural Network (DNN) classifier

What is a Neural Network?

- Composed of layers of neurons
- One layer takes input x_i and produces output using activation function f

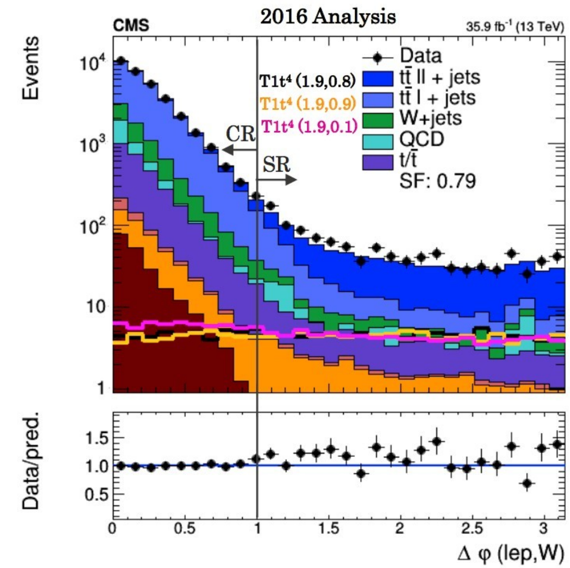


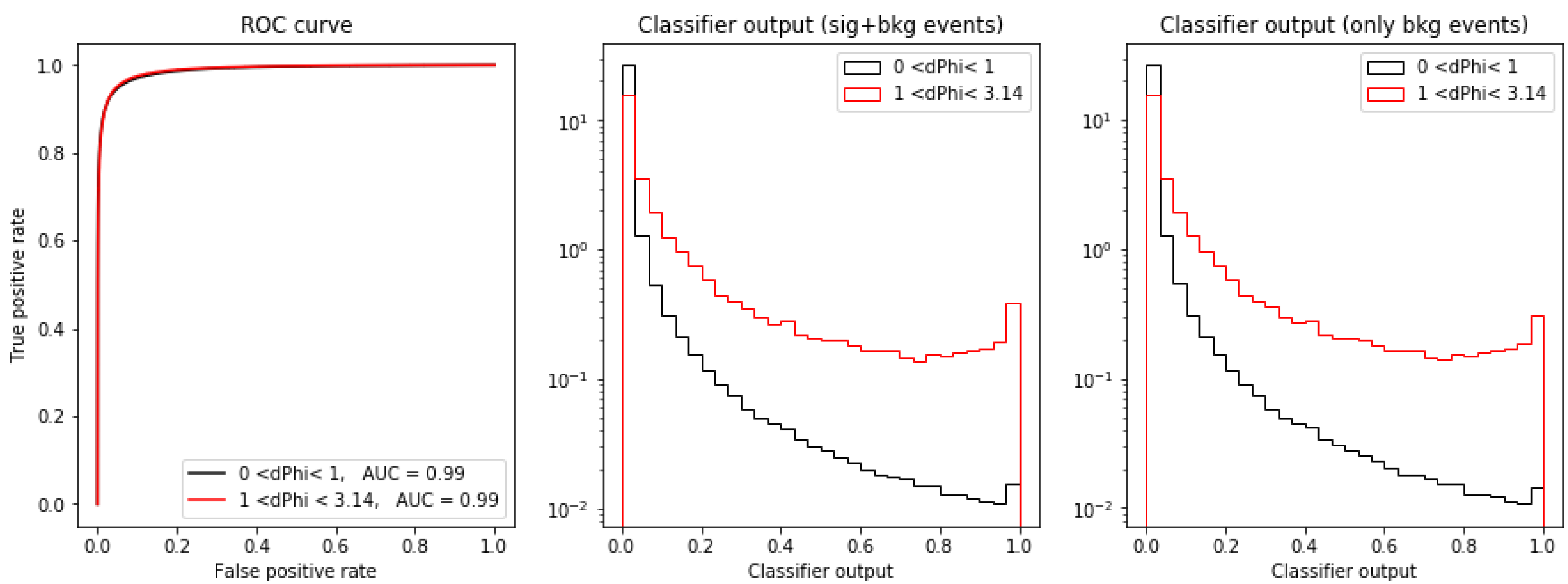
$$f(\sum_{k=1}^{3} w_k x_k + b)$$

- **Loss function**: translates the output → scalar (called loss) ~ which represents how far is the predicted result from the true one
- For **sig/bkg classification**, we use Binary Cross Entropy **(BCE)** as Loss function
- **Goal of training** → minimize the loss by adjusting all the weights and biases

DNN classifier performance

- ✔ Large area under the ROC curve
- ✔ Most events (especially small Δφ events) are classified as bkg
- ➢ **Note**: all events are weighted by their cross-section weight



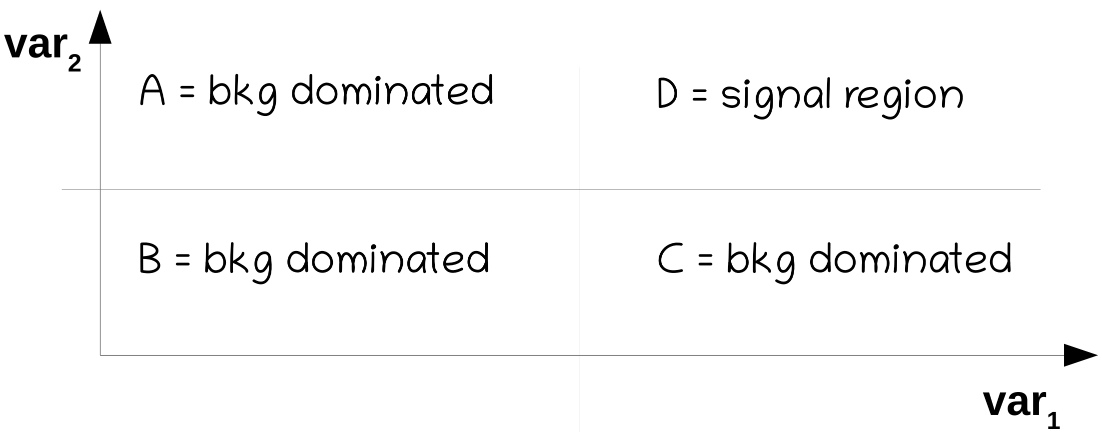


Project Description

- **AIM**: to use ***Data-Driven Background Estimation (ABCD)*** method to extrapolate and predict the background in the signal region

ABCD

If variables var_1 & var_2 are uncorrelated → background in signal region is predicted by the ratio: $\mathbf{N_D = N_C * (N_A / N_B)}$



- **KEY: Decorrelating the relation between the classifier output and Δφ**

- **HOW**? Using **Adversarial Network** to classify sig/bkg events

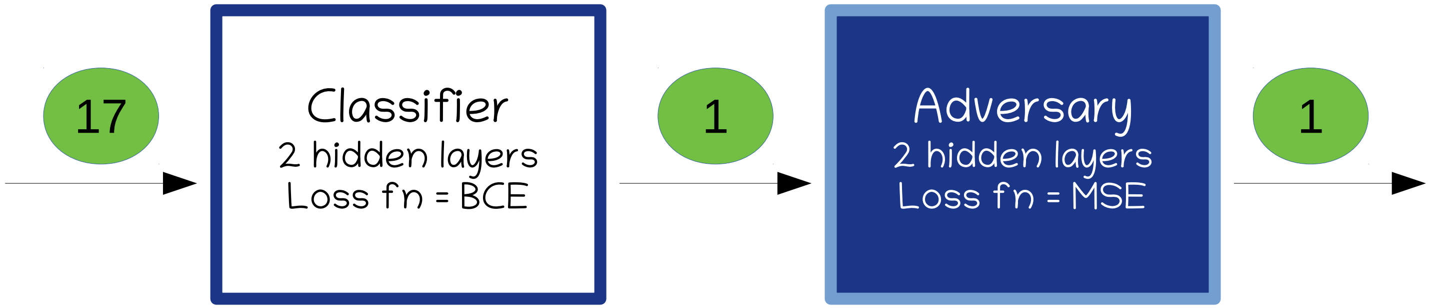
Adversarial Training

What is an Adversarial Network?

https://arxiv.org/abs/1611.01046 ~ Uses Adversarial Network in LHC analysis

- **Adversarial Network** consists of 2 NN(s) ~ an Adversary which competes with the Classifier
- Training is done simultaneously
- **Goal of Adversary →** confuse the classifier and provide feedback until the classifier output is independent of Δφ

✗ This decorrelation comes at the expense of classification efficiency



$$loss = clf_{loss} - \lambda * adv_{loss}$$

$$loss = clf_{loss} - \lambda * adv_{loss}$$

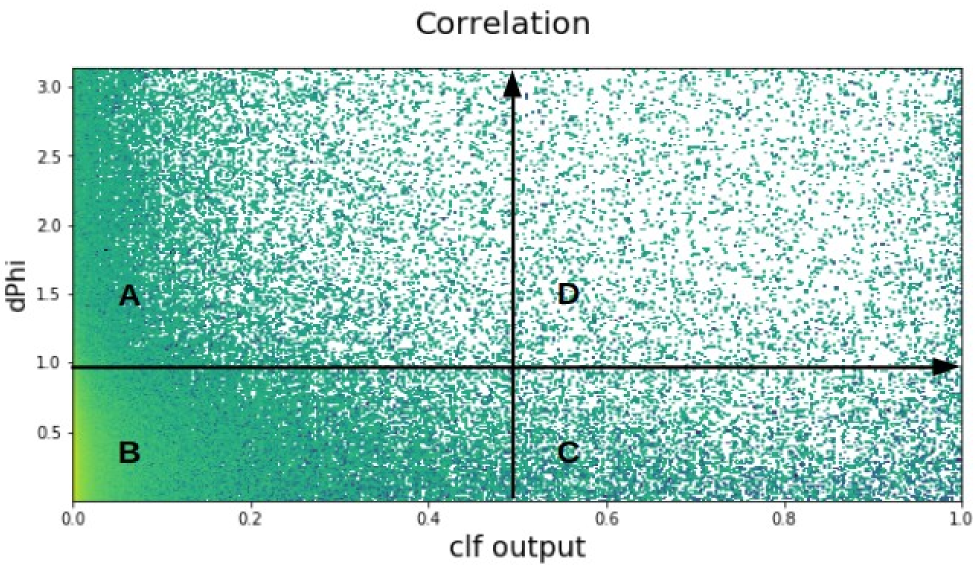
AIM:

Find the optimal value of λ which makes:

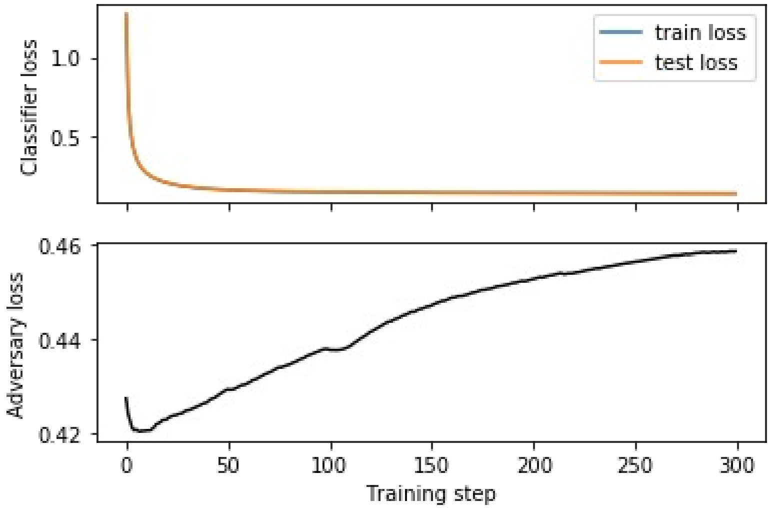
$$\text{Ratio} = \frac{N_A/N_B}{N_D/N_C} = 1$$

Remember that:

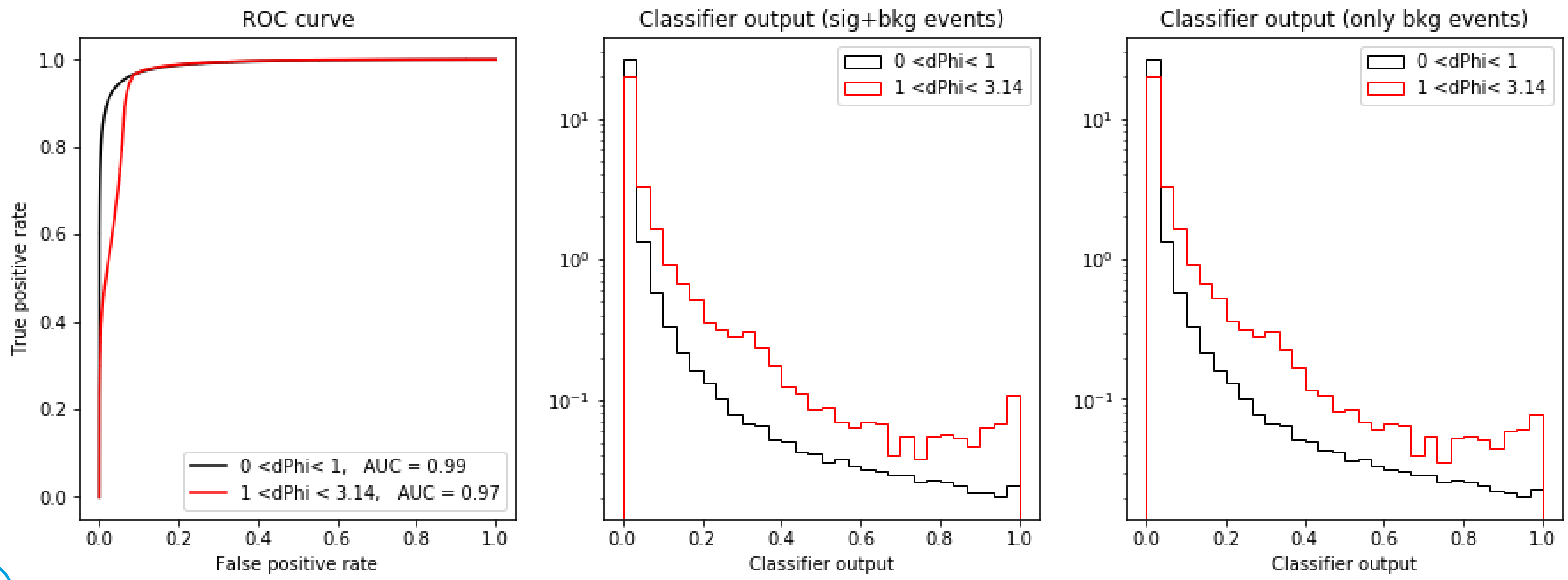
- We **only use bkg** events to get the ratio
- We weight each event by the cross-section weight



Training for λ=0.85



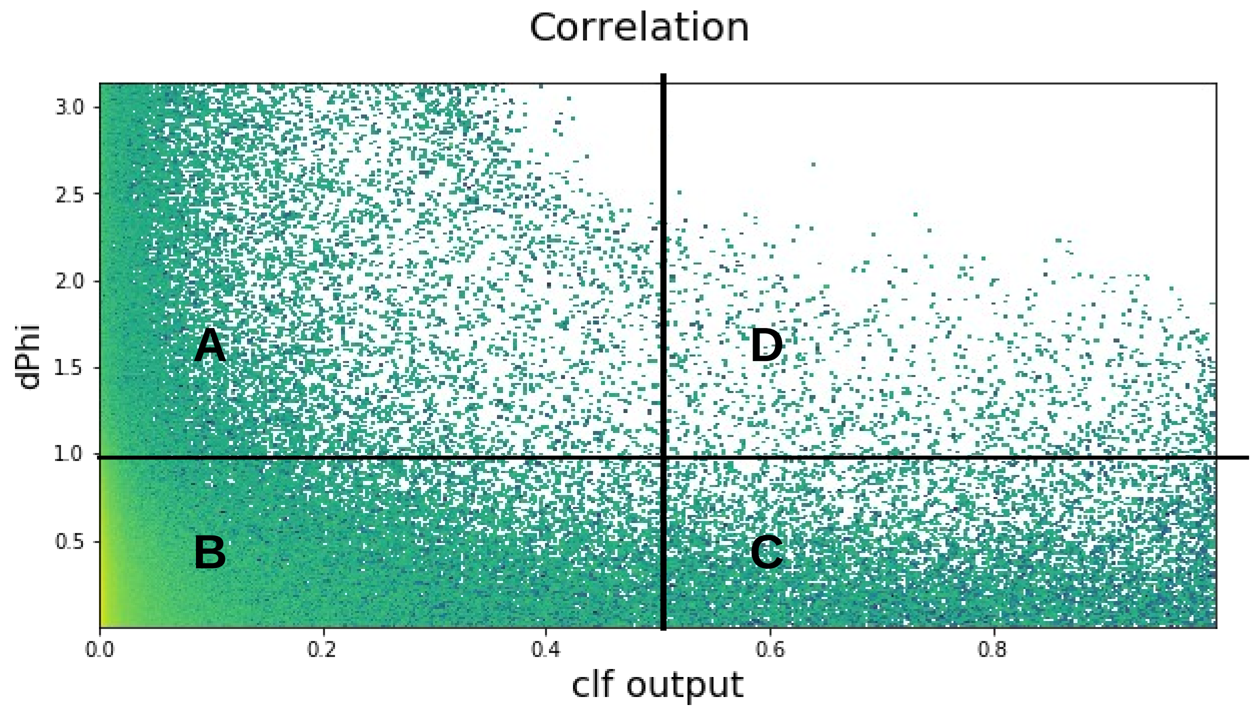
- ✔ Still: Large area under the ROC curve
- ✔ Still: Most events (especially small Δφ events) are classified as bkg
- ✔ Clf output is getting **more independent** of Δφ
- ✗ Small kink in ROC curve indicates small confusion



Training for λ=0.85

Ratio = $\frac{N_A/N_B}{N_D/N_C}$

Ratio ≈ 0.887



- ✔ **Ratio is close to 1**

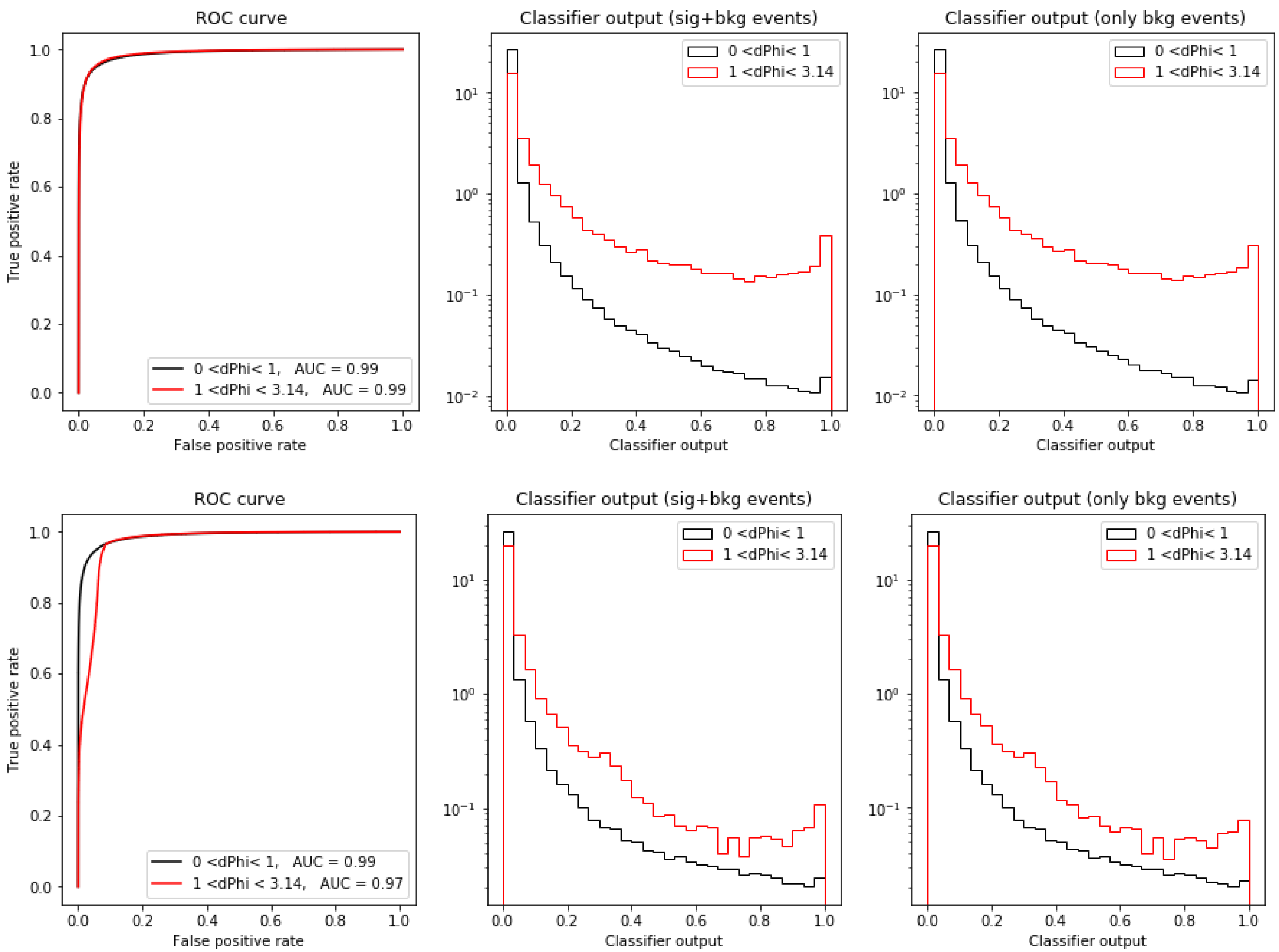
For Comparison

Normal Classifier

Ratio ≈ 0.079

With Adversarial

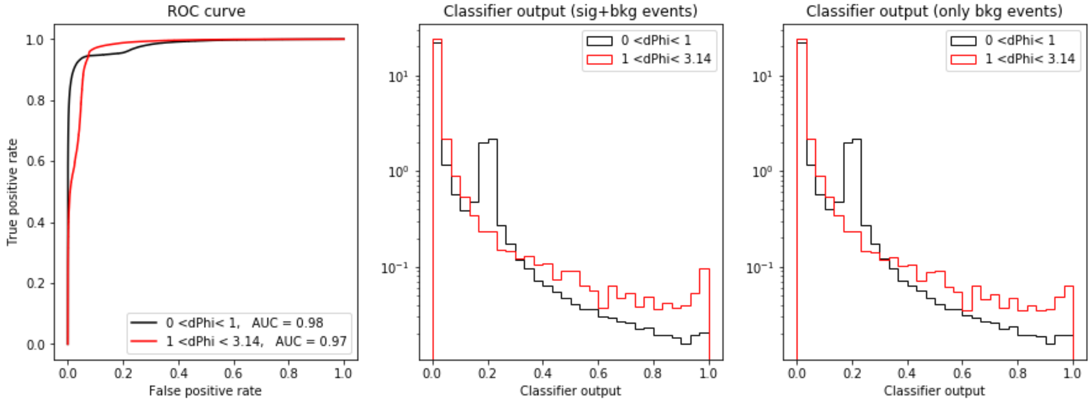
Ratio ≈ 0.887



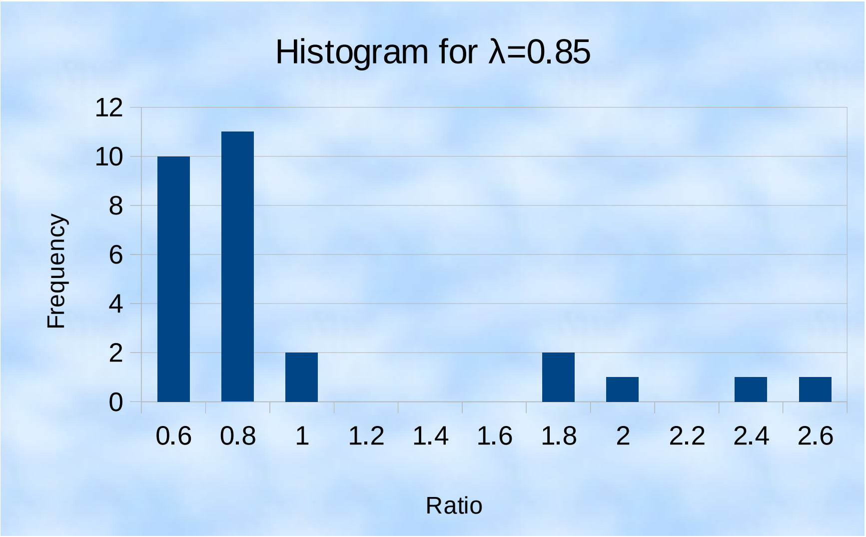
Conclusion: Adversarial is successful in decorrelating Δφ from the classifier output

The Stability of Adversarial Training

- ✗ Sometimes we find kinks in the *clf* output distribution



- ✗ Ratio fluctuates for the same λ



Summary

Achieved so far

- Constructed an Adversarial Network
- Studied Adversarial Training
- Successful decorrelation of $\Delta\varphi$ from the *clf* output at the expense of small efficiency loss

Next step

- Study the stability of the Adversarial Training
- Optimize the hyper-parameter space of both the classifier and the adversary
- Use the Adversarial Network on real data to predict bkg in signal region

Thank you

Danke

Backup

Results for different values of λ

λ	Ratio	AUC	Time	# of epochs
0 (normal classifier)	0.07944514	0.99	9 min	300
0.5	2.60919688	0.99	32 min	300
0.75	1.4703602	0.98	32 min	300
0.8	1.83401852	0.98	32 min	300
0.85	0.88716993	0.98	37 min	300
0.9	0.68016004	0.97	32 min	300
1	0.62123818	0.97	32 min	300
10	0.06472567	0.99	32 min	300

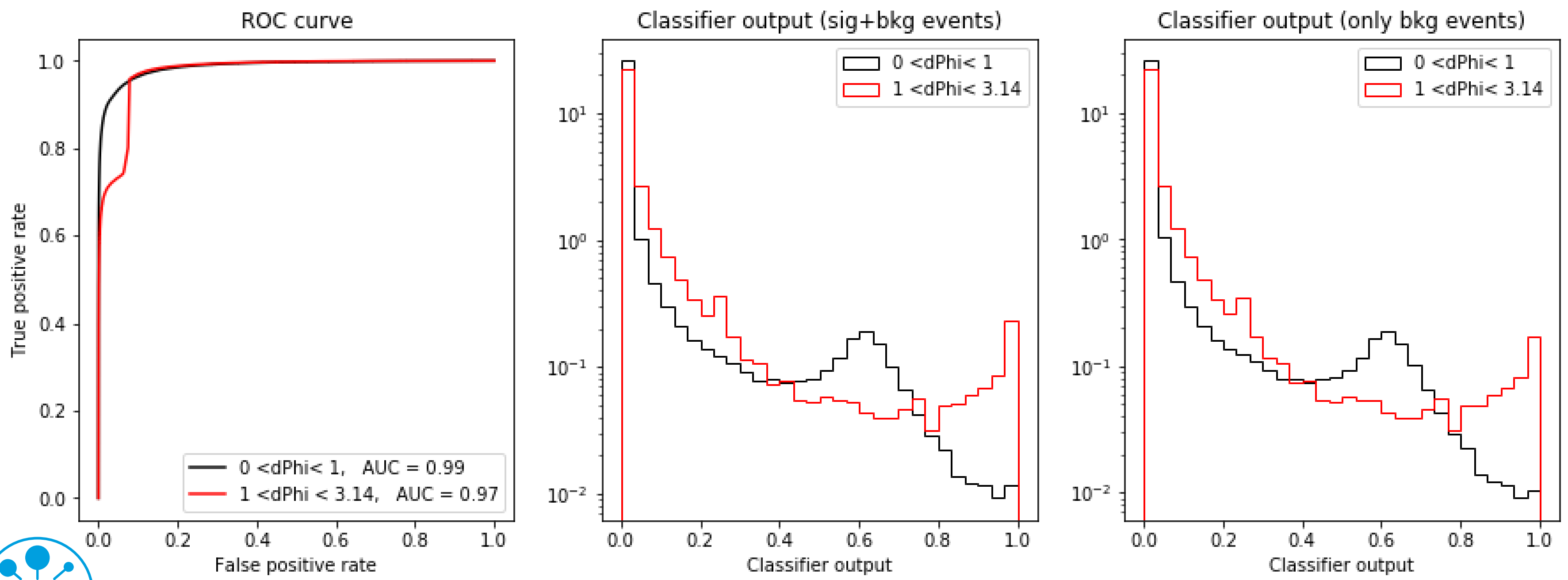
My Network's Architecture

$$loss = clf_{loss} - \lambda * adv_{loss}$$

Network	Classifier	Adversary
# of inputs	17	1
# of outputs	1	1
Hidden Layers (2)	125 neurons each	50 neurons each
Activation Function	reLu	
Output Activation Function	sigmoid	none
Loss	Binary Cross Entropy	Mean Squared Error
Learning Rate	0.001	0.005
Optimizer	Adam	

Training for λ=0.8

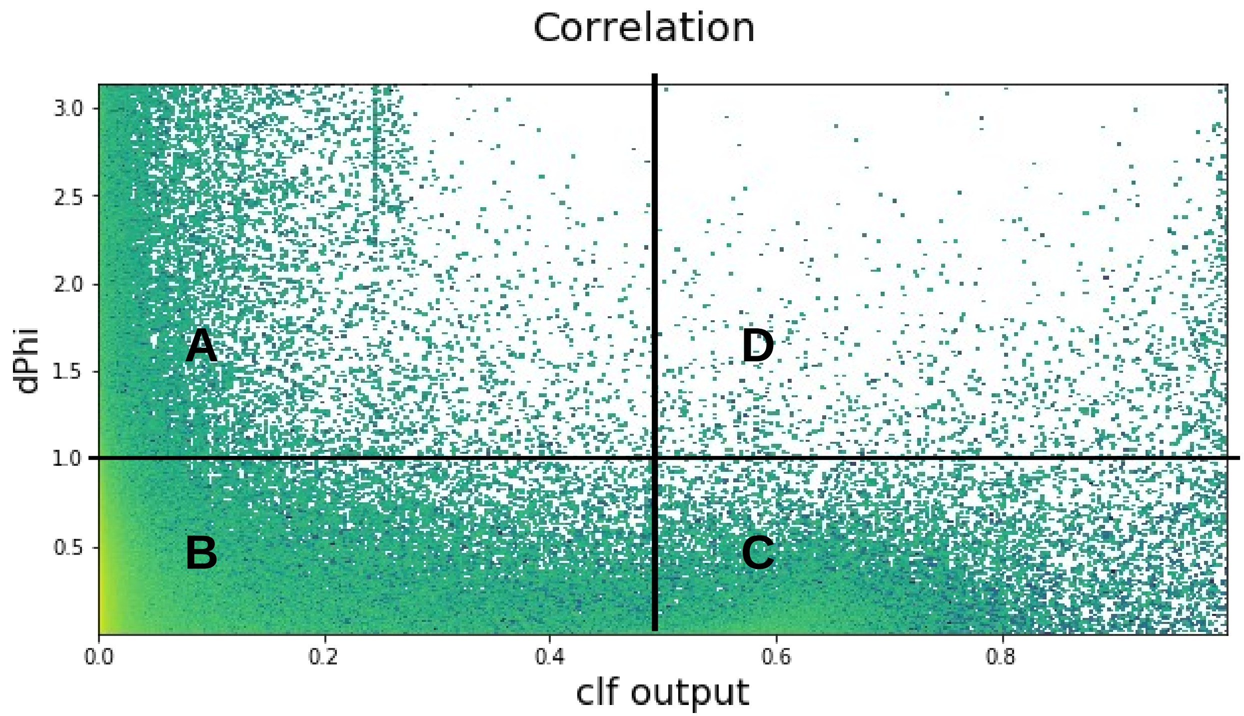
- ✔ Still: Large area under the ROC curve
- ✔ Still: Most events (especially small Δφ events) are classified as bkg
- ✔ Clf output is getting **more independent** of Δφ
- ✗ Small kink in ROC curve indicates small confusion



Training for λ=0.8

Ratio = $\dfrac{N_A/N_B}{N_D/N_C}$

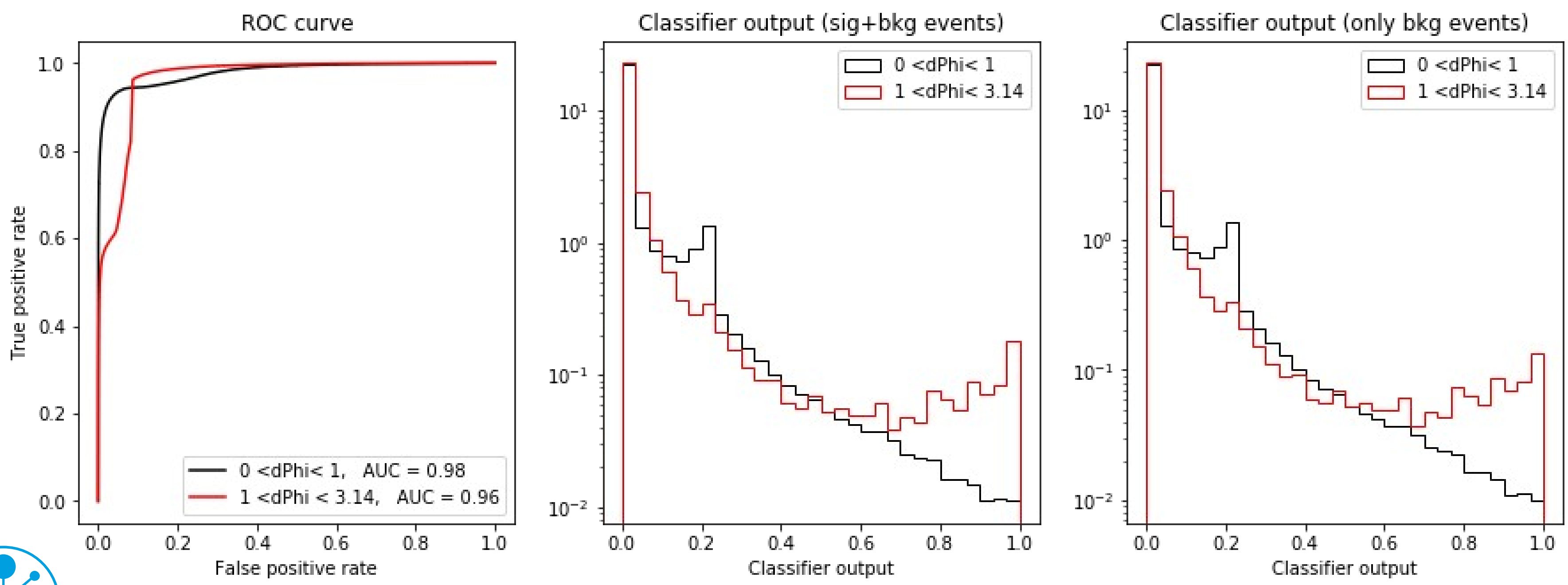
Ratio ≈ 1.834



✔ **Ratio is close to 1**

Training for λ=0.9

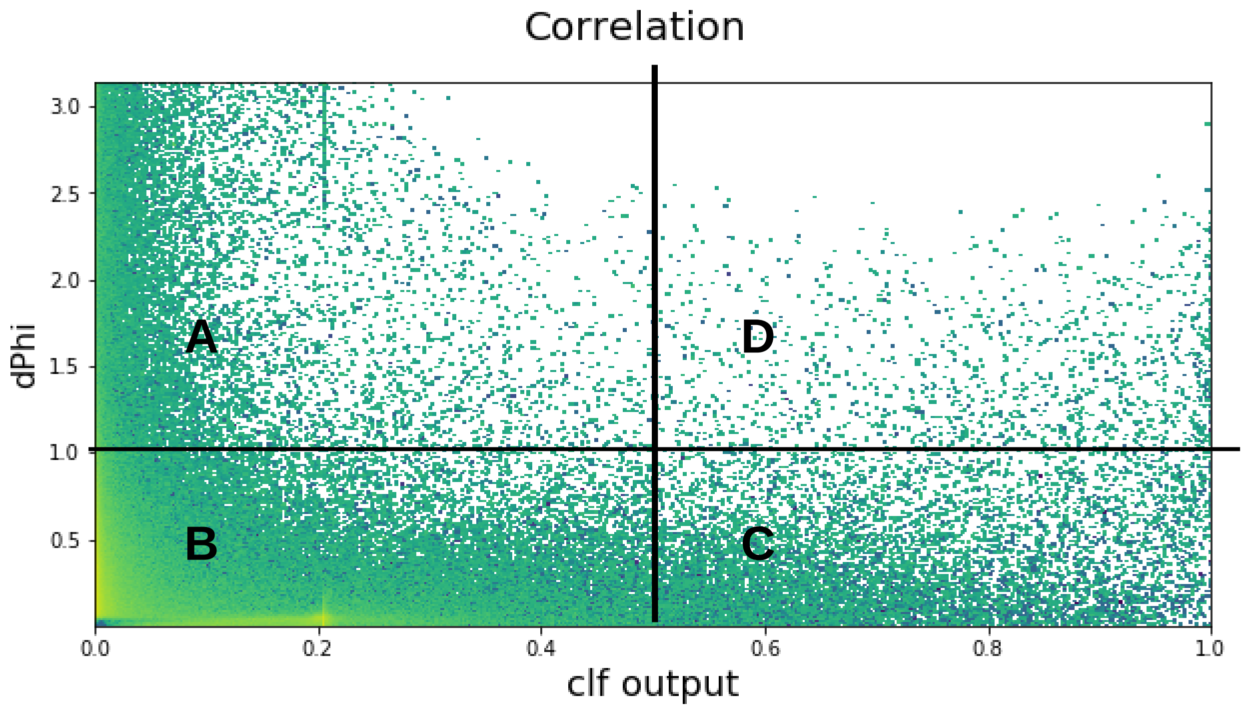
- ✔ Still: Large area under the ROC curve
- ✔ Still: Most events (especially small Δφ events) are classified as bkg
- ✔ Clf output is getting **more independent** of Δφ
- ✗ Small kink in ROC curve indicates small confusion



Training for λ=0.9

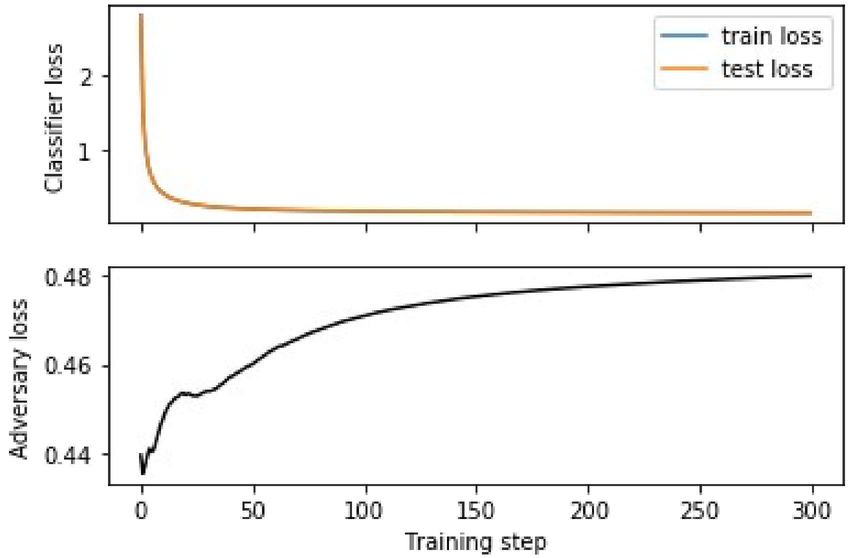
Ratio = $\frac{N_A/N_B}{N_D/N_C}$

Ratio ≈ 0.680

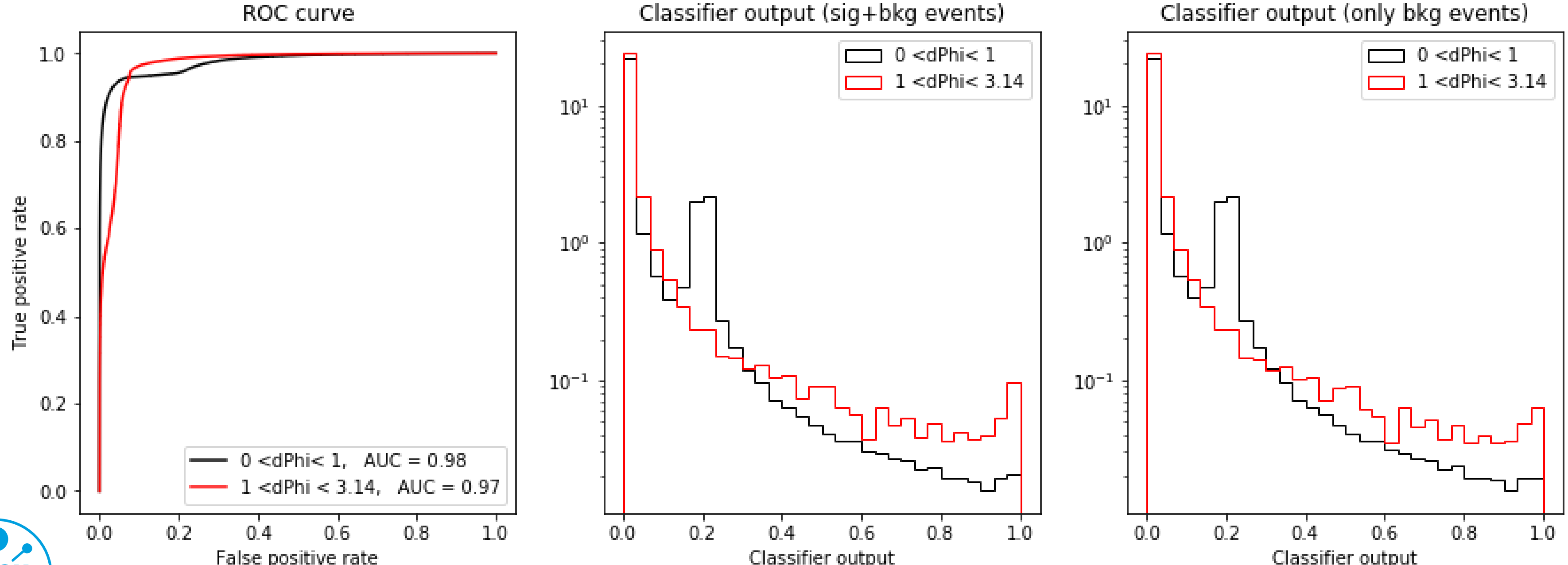


✔ **Ratio is close to 1**

Training for λ=0.85



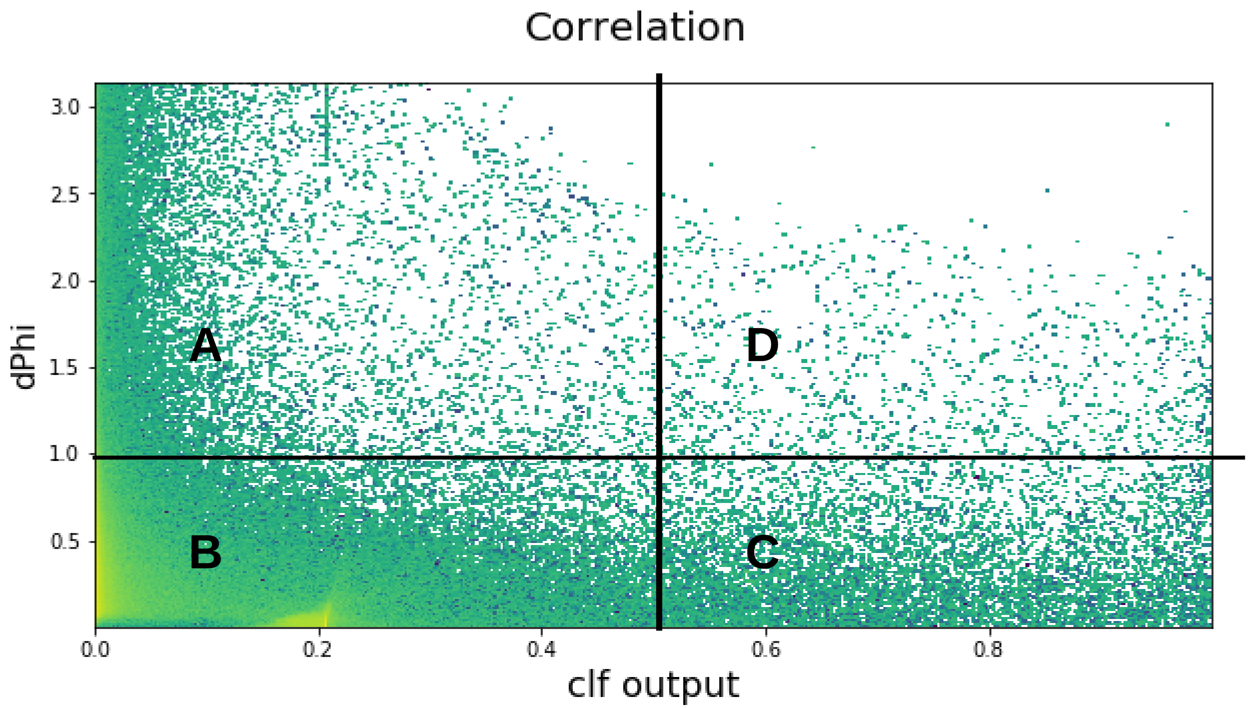
- ✔ Still: Large area under the ROC curve
- ✔ Still: Most events (especially small Δφ events) are classified as bkg
- ✔ Clf output is getting **more independent** of Δφ
- ✗ Small kinks in ROC curve and clf output indicate small confusion



Training for λ=0.85

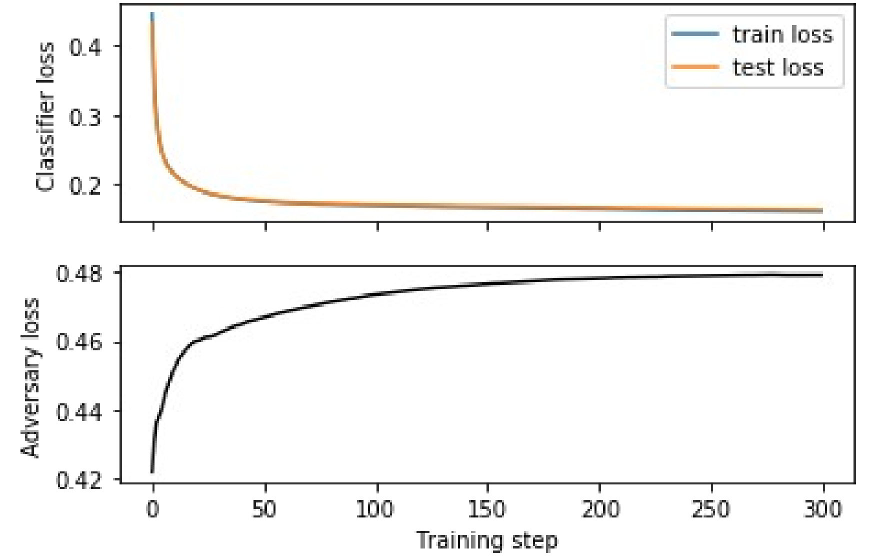
Ratio = $\frac{N_A/N_B}{N_D/N_C}$

Ratio ≈ 0.81

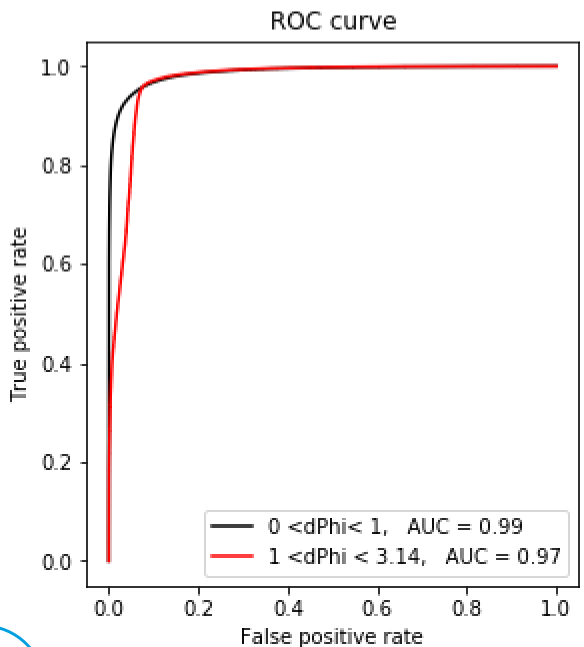


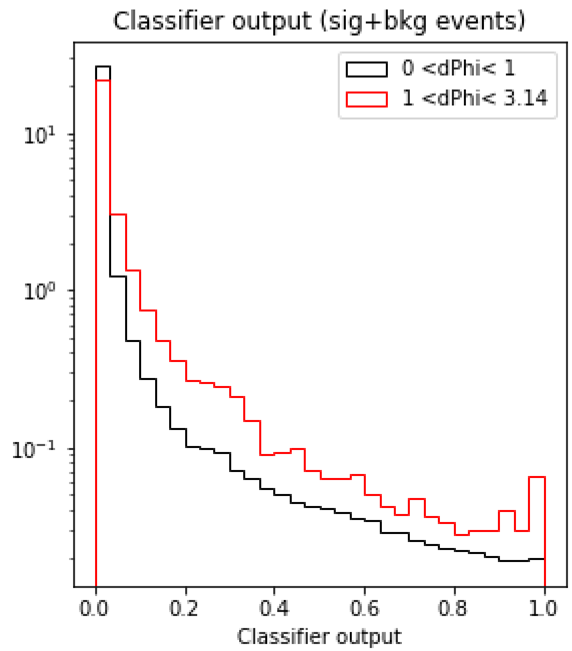
✔ **Ratio is close to 1**

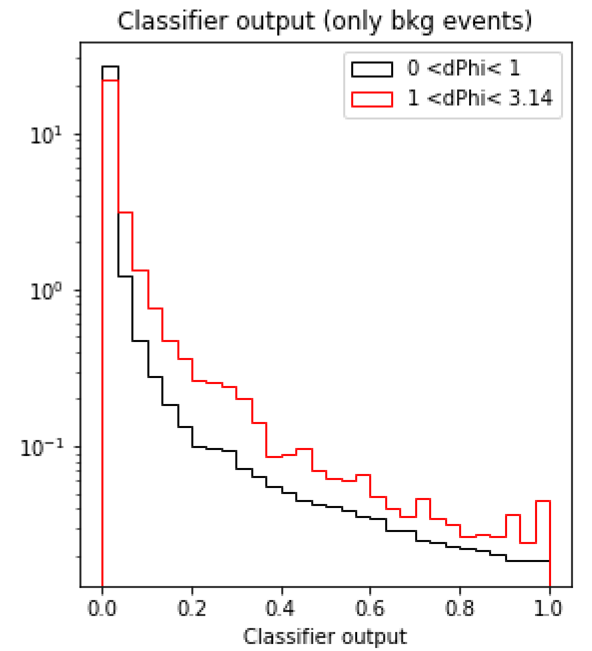
Training for λ=0.85



- ✔ Still: Large area under the ROC curve
- ✔ Still: Most events (especially small Δφ events) are classified as bkg
- ✔ Clf output is getting **more independent** of Δφ
- ✗ Small kink in ROC curve indicates small confusion



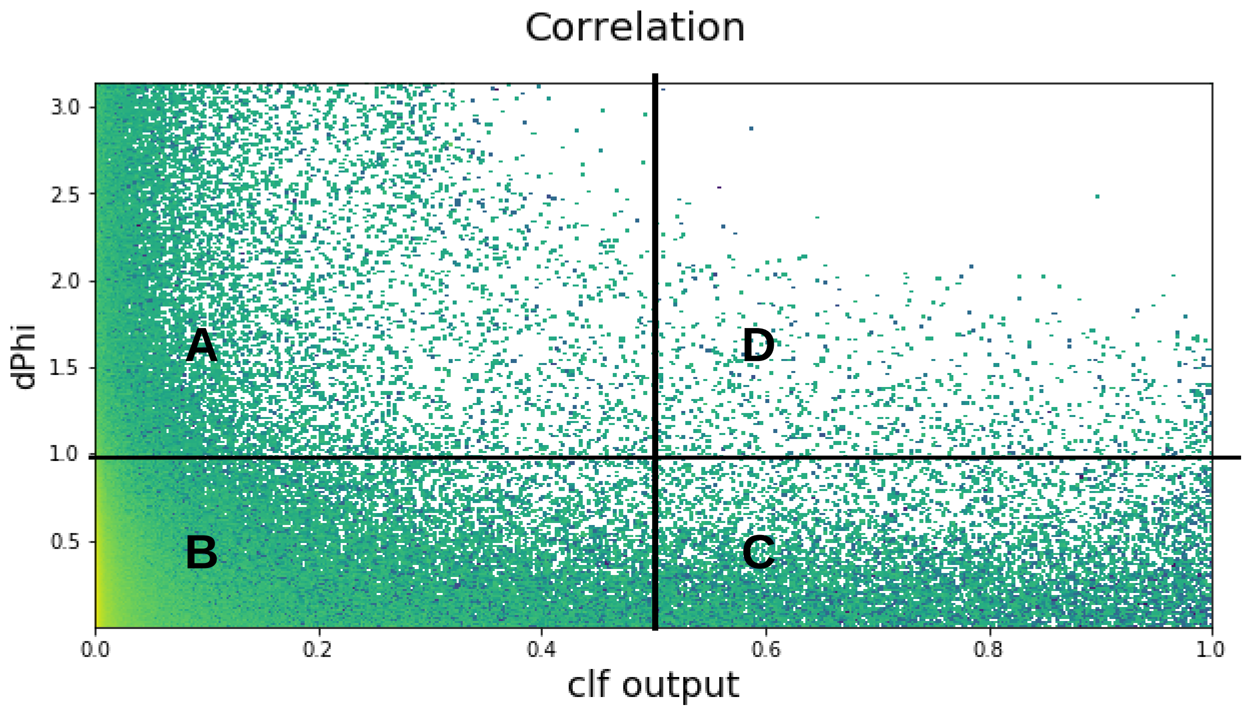




Training for λ=0.85

Ratio = $\frac{N_A/N_B}{N_D/N_C}$

Ratio ≈ 1.632



- ✔ **Ratio is close to 1**