Studies of the $t\bar{t} + \gamma$ production process

Pascal Herrmann

Supervisor: Prof. Dr. Arnulf Quadt

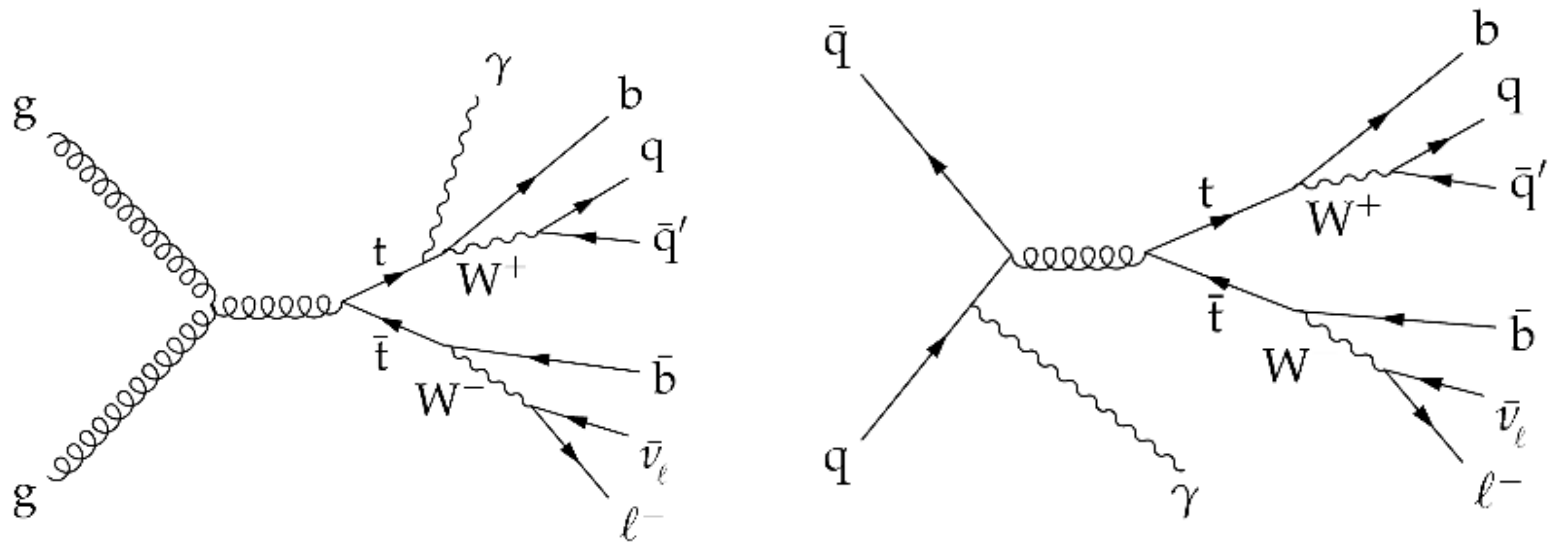
II. Physikalisches Institut

July 13, 2018

$t\bar{t} + \gamma$ production process

Try to distinguish photons which were radiated off

- top quark (FSR)
- quark in initial state (ISR)

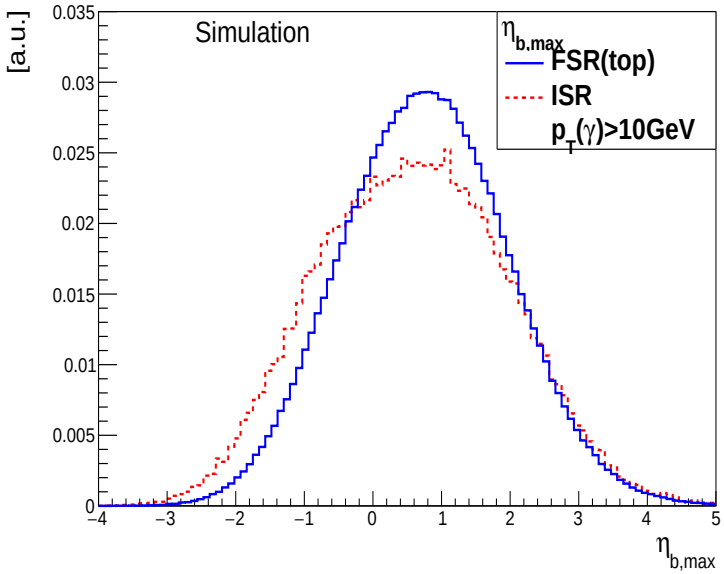


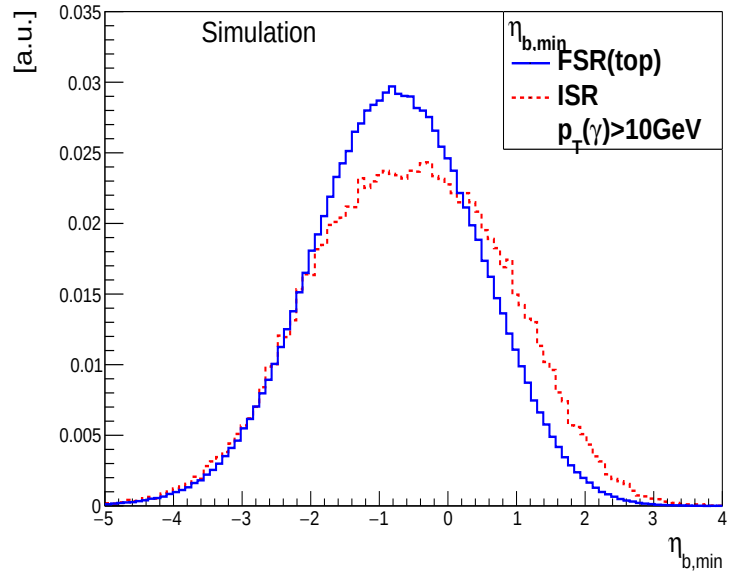
- Necessary for more precise cross section measurements

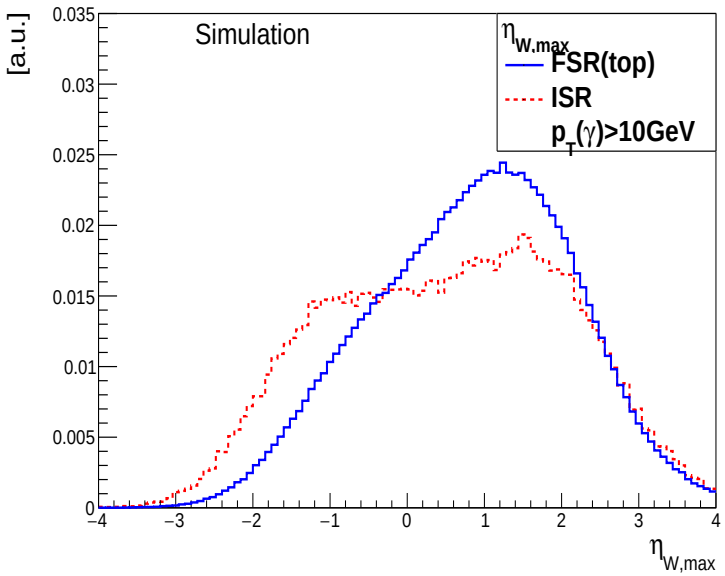
Investigation of discriminant variables

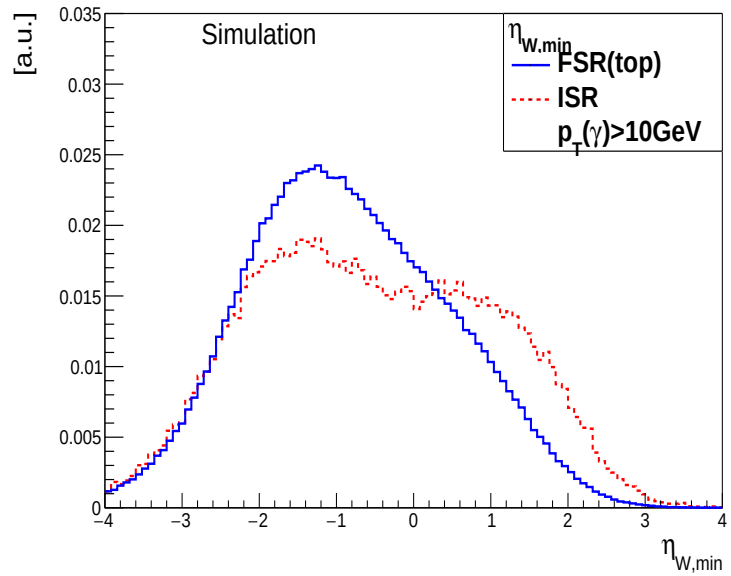
- Studies of discriminant variables
- Using modified version of MadGraph
  - MadGraph does not distinguish between FSR and ISR
  - Introduced new coupling for that

Investigation of discriminant variables

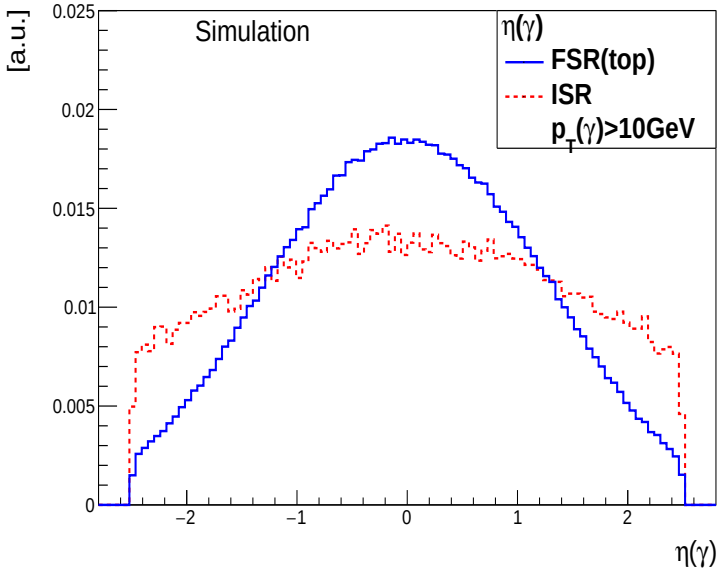


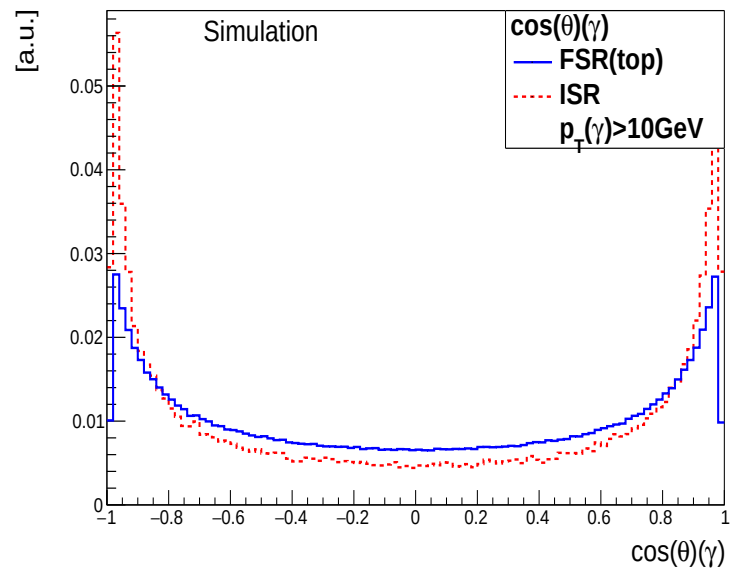


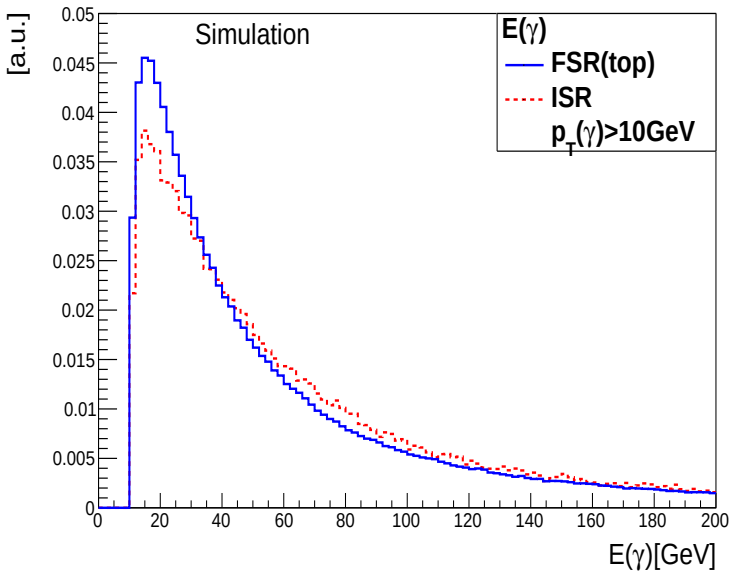


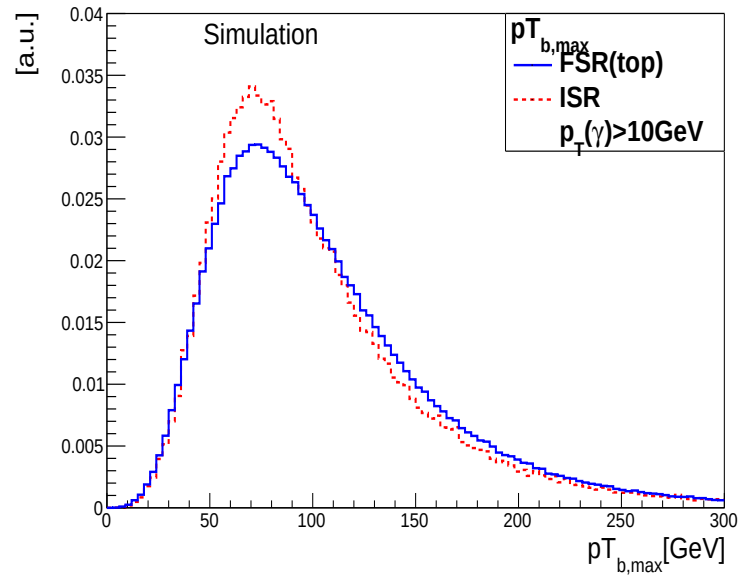


Investigation of discriminant variables

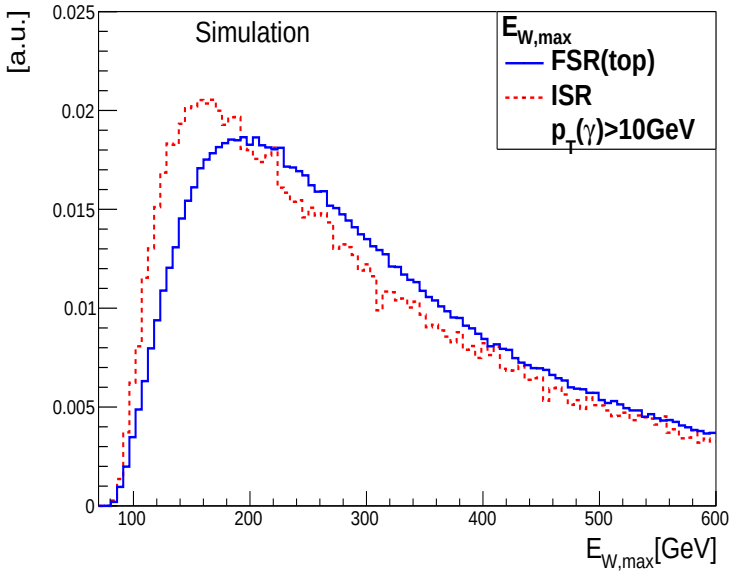


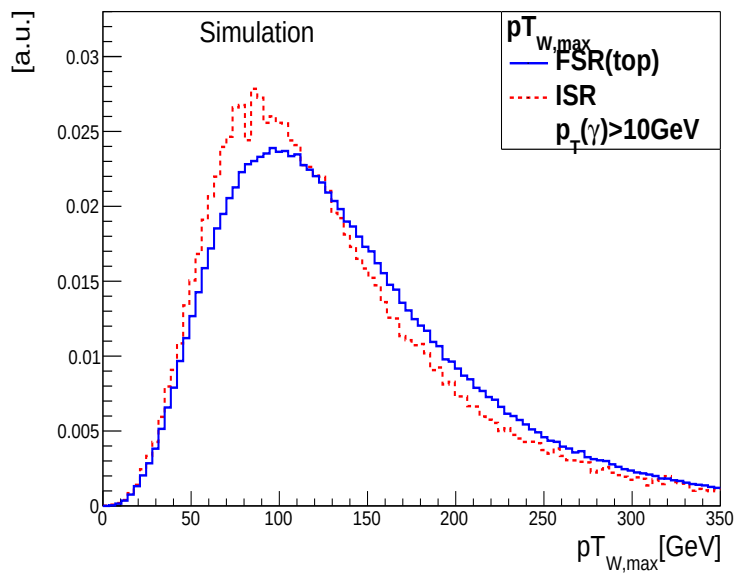


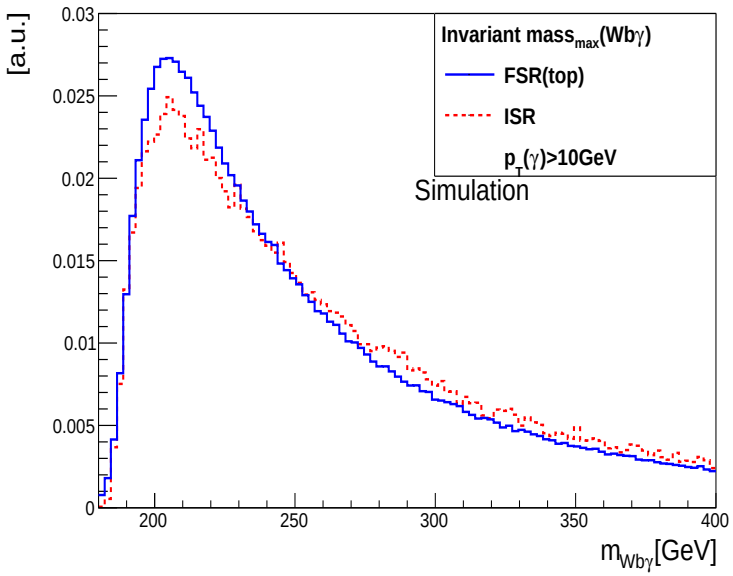


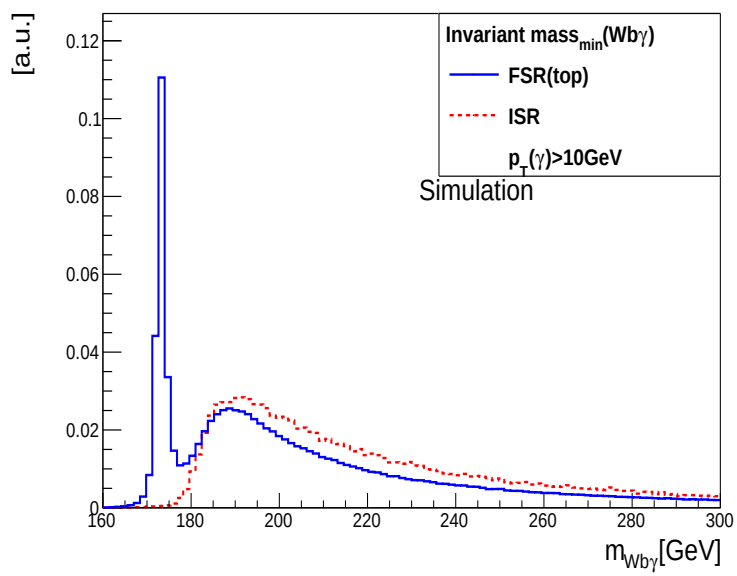


Investigation of discriminant variables

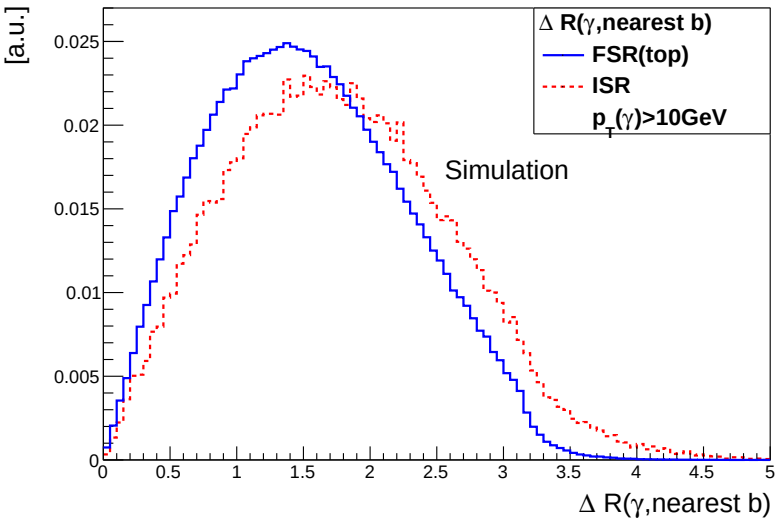


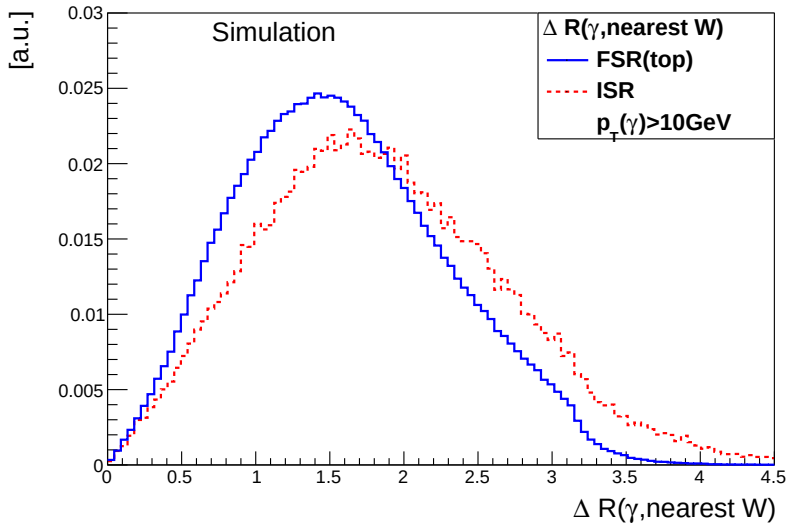






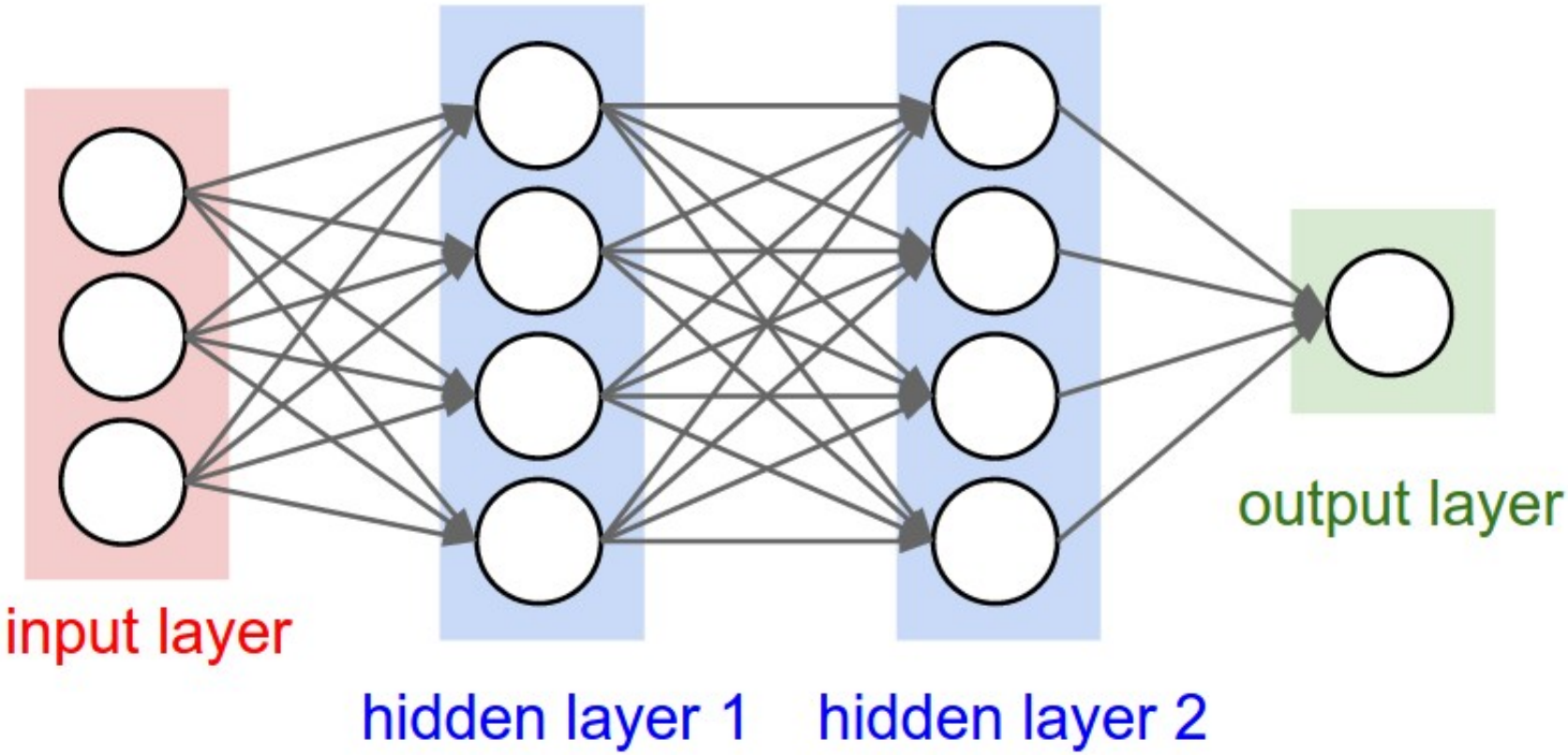
Investigation of discriminant variables





- Reason for differences
  - Energy differences through ISR photon
  - Angle differences (ISR $\rightarrow \eta \approx \infty$, FSR $\rightarrow \eta \approx 0$)

Neural networks

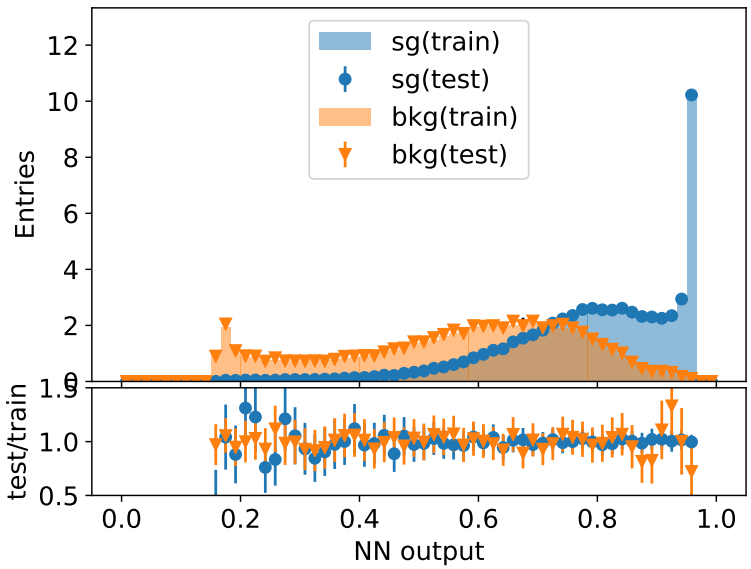


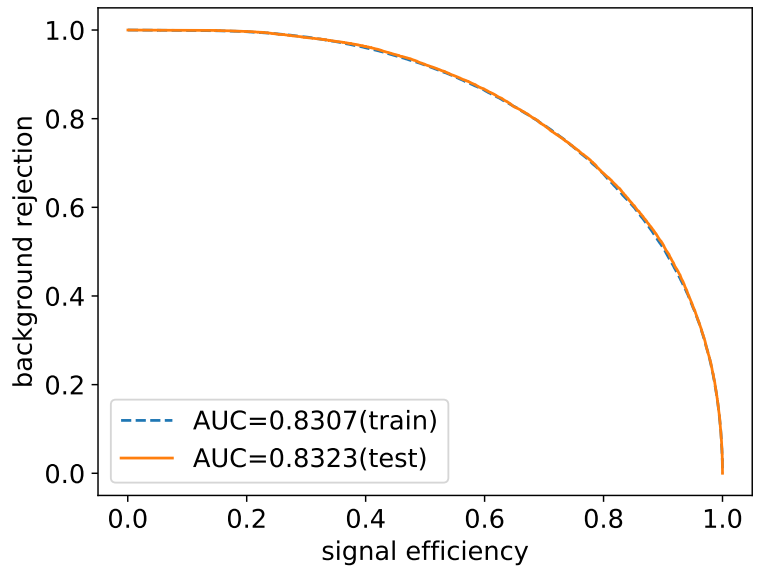
- Dense layers (weights, biases, activation functions)
- BatchNormalization layers

Training of a neural network

- 210 973 FSR and 75 618 ISR events
- 80% training data and 20% testing
- Batch size of 10 000
- 100 epochs
- ISR events weighted by a factor of 2
- ROC curves and AUC values for evaluation

Best Configuration





relu(19) → BN → softmax(95) → BN → softmax(29) → BN → softmax(11)

Conclusion

- Separation is possible (at parton level)
- AUC-Value of 0.8400 seems to be the maximal value (for my variables)
- Further research:
  - Check variables for correlation
  - Test NN on reco data
  - Search for new variables
  - More advanced NN algorithms

Backup slides

Activation functions

Softmax:

$$f_i(\vec{x}) = \frac{e^{x_i}}{\sum_{j=1}^{J} e^{e_j}}$$

Rectified linear unit(ReLU):

$$f(x) = \begin{cases} 0, & \text{if } x < 0 \\ x, & \text{otherwise} \end{cases}$$

Sigmoid:

$$sig(x) = \frac{1}{1 + e^{-x}}$$

Linear:

$$f(x) = x$$

tanh:

$$f(x) = tanh(x)$$

Activation functions

Softsign:

$$f(x) = \frac{x}{1 + |x|}$$

Exponential linear unit (ELU):

$$f(\alpha, x) = \begin{cases} \alpha(e^x - 1), & \text{if } x < 0 \\ x, & \text{otherwise} \end{cases}$$

Scaled exponential linear unit (SELU):

$$f(\alpha, x) = \lambda \begin{cases} \alpha(e^x - 1), & \text{if } x < 0 \\ x, & \text{otherwise} \end{cases}$$

with $\lambda = 1.0507$ and $\alpha = 1.67326$

Softplus:

$$f(x) = ln(1 + e^x)$$

Trained neural networks

Model	AUC(train)	AUC(test)
3 Layers(64,40,52(softmax)) with BN	0.8311	0.8258
3 Layers(64,40,52)	0.8053	0.8018
3 Layers(64,40,52) with BN	0.8306	0.8263
4 Layers(19,95,29,11) with BN	0.8119	0.8088
4 Layers(19,95,29,11(softmax)) with BN	0.8307	0.8323
1 Layers(8(softmax))	0.7355	0.7335
1 Layers(128)	0.7901	0.7871
1 Layers(8)	0.7358	0.7324

Trained neural networks

Model	AUC(train)	AUC(test)
2 Layers(24,64(softmax)) with BN	0.8269	0.8270
2 Layers(24,64) with BN	0.8273	0.8271
3 Layers(24,64,48(softmax)) with BN	0.8255	0.8252
3 Layers(24,64,48(softmax)) with BN	0.8247	0.8214
3 Layers(24,64,48(sigmoid)) with BN	0.8198	0.8198
4 Layers(64,152,60,50) with BN	0.8376	0.8209
2 Layers(24,64(softmax,sigmoid)) with BN	0.8191	0.8207
3 Layers(24,64,48(softmax,sigmoid)) with BN	0.8117	0.8141
3 Layers(24,64,48(sigmoid)) with BN	0.8174	0.8157

Trained neural networks

Model	AUC(train)	AUC(test)
4l(64relu,40relu,52relu,48soft) with BN	0.8315	0.8241
4l(64relu,40relu,52relu,82sig) with BN	0.8317	0.8266
4l(64relu,48soft,52relu,74sig) with BN	0.8182	0.8175
4l(64relu,48sig,52relu,74soft) with BN	0.8335	0.8292
5l(64relu,48sig,52relu,74soft,68relu) BN	0.8304	0.8159
4l(78relu,62sig,66relu,88soft) with BN	0.8354	0.8253
1l(64relu) with BN	0.7991	0.7987
1l(64sig) with BN	0.8137	0.8111
1l(64soft) with BN	0.8074	0.8083
1l(46relu) with BN	0.8010	0.8027

Trained neural networks

Model	AUC(train)	AUC(test)
1l(46sig) with BN	0.8145	0.8147
1l(46soft) with BN	0.8024	0.8002
1l(82relu) with BN	0.8064	0.8072
1l(82sig) with BN	0.8174	0.8129
1l(82soft) with BN	0.7730	0.7761
1l(128sig) with BN	0.8153	0.8173
2l(128sig,46soft) with BN	0.8310	0.8251
2l(128sig,46relu) with BN	0.8319	0.8231
2l(128sig,46sig) with BN	0.8258	0.8229
2l(128sig,64soft) with BN	0.8278	0.8248

Trained neural networks

Model	AUC(train)	AUC(test)
2l(128sig,64relu) with BN	0.8294	0.8230
2l(128sig,64sig) with BN	0.8203	0.8202
2l(128sig,82soft) with BN	0.8312	0.8250
2l(128sig,82relu) with BN	0.8280	0.8242
2l(128sig,82sig) with BN	0.8230	0.8193
3l(128sig,1sig,46soft) with BN	0.8129	0.8133
3l(128sig,1sig,128sig) with BN	0.8138	0.8080
2l(82sig,46relu) with BN	0.8301	0.8245
3l(128sig,82relu,46soft) with BN	0.8341	0.8204
1l(46elu) with BN	0.8174	0.8136

Trained neural networks

Model	AUC(train)	AUC(test)
1l(46lin) with BN	0.7357	0.7304
1l(46selu) with BN	0.8171	0.8128
1l(46softplus) with BN	0.8133	0.8154
1l(46softsign) with BN	0.8100	0.8103
1l(46tanh) with BN	0.8144	0.8102
1l(64elu) with BN	0.8143	0.8160
1l(64selu) with BN	0.8139	0.8122
1l(64softplus) with BN	0.8173	0.8158
1l(64softsign) with BN	0.8040	0.8034
1l(64tanh) with BN	0.8182	0.8152

Trained neural networks

Model	AUC(train)	AUC(test)
1l(82elu) with BN	0.8155	0.8160
1l(82selu) with BN	0.8172	0.8137
1l(82softplus) with BN	0.8149	0.8159
1l(82softsign) with BN	0.8038	0.8027
1l(82tanh) with BN	0.8165	0.8147
2l(64softplus,46relu) with BN	0.8307	0.8260
2l(64softplus,46sig) with BN	0.8111	0.8133
2l(64softplus,46elu) with BN	0.8223	0.8225
2l(64softplus,46selu) with BN	0.8293	0.8243
2l(64softplus,46soft) with BN	0.8290	0.8282

Trained neural networks

Model	AUC(train)	AUC(test)
2l(64softplus,46softplus) with BN	0.8091	0.8058
2l(64softplus,46softsign) with BN	0.8282	0.8254
2l(64softplus,46tanh) with BN	0.8149	0.8136
2l(64softplus,64relu) with BN	0.8302	0.8242
2l(64softplus,64sig) with BN	0.8087	0.8077
2l(64softplus,64soft) with BN	0.8289	0.8253
2l(64softplus,64elu) with BN	0.8229	0.8233
2l(64softplus,64selu) with BN	0.8282	0.8286
2l(64softplus,64softsign) with BN	0.8259	0.8228
2l(64softplus,64tanh) with BN	0.8250	0.8214

Trained neural networks

Model	AUC(train)	AUC(test)
2l(64softplus,64softplus) with BN	0.8166	0.8172
2l(64softplus,82relu) with BN	0.8292	0.8263
2l(64softplus,82sig) with BN	0.8182	0.8205
2l(64softplus,82soft) with BN	0.8321	0.8280
2l(64softplus,82elu) with BN	0.8173	0.8182
2l(64softplus,82selu) with BN	0.8181	0.8160
2l(64softplus,82softplus) with BN	0.8144	0.8122
2l(64softplus,82softsign) with BN	0.8279	0.8253
2l(64softplus,82tanh) with BN	0.8184	0.8191
3l(64softplus,82soft,46relu) with BN	0.7997	0.7955