

UNIVERSITY OF
COPENHAGEN



A PHYSICIST'S APPROACH TO MACHINE LEARNING
Understanding The Basic Bricks

CHRISTIAN MICHELSSEN
Master's Thesis (Cand. Scient.)
January 3rd, 2020

Supervised by
Troels Petersen



UNIVERSITY OF COPENHAGEN

Copyright © 2020
Christian Michelsen

[HTTPS://GITHUB.COM/CHRISTIANMICHelsen](https://github.com/CHRISTIANMICHelsen)

This thesis was inspired by the works of Edward R. Tufte and is based on the Tufte-L^AT_EX package.

First printing, January 2020

Abstract

Here will be a decent abstract at some pointTM.

Contents

<i>Abstract</i>	iii
<i>Table of Contents</i>	v
<i>Foreword</i>	ix
1 <i>Introduction</i>	1
<i>Part I</i>	3
2 <i>Machine Learning Theory</i>	5
2.1 <i>Statistical Learning Theory</i>	5
2.2 <i>Supervised Learning</i>	6
2.3 <i>Generalization Bound</i>	7
2.3.1 <i>Generalization Bound for Infinite Hypotheses</i>	9
2.4 <i>Avoiding overfitting</i>	10
2.4.1 <i>Model Regularization</i>	10
2.4.2 <i>Cross Validation</i>	12
2.4.3 <i>Early Stopping</i>	13
2.5 <i>Loss functions</i>	14
2.5.1 <i>Evaluation Function</i>	16
2.6 <i>Decision Trees</i>	16
2.6.1 <i>Ensembles of Decision Trees</i>	17
2.7 <i>Hyperparameter Optimization</i>	19
2.7.1 <i>Grid Search</i>	20
2.7.2 <i>Random Search</i>	20
2.7.3 <i>Bayesian Optimization</i>	21
2.8 <i>Feature Importance</i>	22

3	<i>Danish Housing Prices</i>	27
	3.1 <i>Data Preparation and Exploratory Data Analysis</i>	28
	3.1.1 <i>Correlations</i>	30
	3.1.2 <i>Validity of Input Variables</i>	31
	3.1.3 <i>Cuts</i>	33
	3.2 <i>Feature Augmentation</i>	33
	3.2.1 <i>Time-Dependent Price Index</i>	34
	3.3 <i>Evaluation Function</i>	35
	3.4 <i>Initial Hyperparameter Optimization</i>	36
	3.5 <i>Hyperparameter Optimization</i>	38
	3.6 <i>Results</i>	40
	3.7 <i>Model Inspection</i>	43
	3.8 <i>Multiple Models</i>	45
	3.9 <i>Discussion</i>	47
	3.10 <i>Conclusion</i>	49
	<i>Part II</i>	51
4	<i>Particle Physics and LEP</i>	53
	4.1 <i>The Standard Model</i>	53
	4.2 <i>Quark Hadronization</i>	54
	4.3 <i>The ALEPH Detector and LEP</i>	56
	4.4 <i>Jet Clustering</i>	58
	4.5 <i>The Variables</i>	58
5	<i>Quark Gluon Analysis</i>	63
	5.1 <i>Data Preprocessing</i>	63
	5.2 <i>Exploratory Data Analysis</i>	64
	5.2.1 <i>Dimensionality Reduction</i>	66
	5.2.2 <i>Correlations</i>	67
	5.3 <i>Loss and Evaluation Function</i>	67
	5.4 <i>b</i> - <i>Tagging Analysis</i>	68
	5.4.1 <i>b</i> - <i>Tagging Hyperparameter Optimization</i>	68
	5.4.2 <i>b</i> - <i>Tagging Results</i>	70
	5.4.3 <i>b</i> - <i>Tagging Model Inspection</i>	71
	5.5 <i>b</i> - <i>Tagging Efficiency</i>	72
	5.6 <i>g</i> - <i>Tagging Analysis</i>	74
	5.6.1 <i>Permutation Invariance</i>	75
	5.6.2 <i>Truncated Uniform PDF</i>	75

5.6.3	<i>g</i> -Tagging Hyperparameter Optimization	76
5.6.4	<i>PermNet</i>	77
5.6.5	<i>1D Comparison of LGB and PermNet</i>	78
5.6.6	<i>g</i> -Tagging Results	78
5.7	<i>g</i> -Tagging Efficiency	81
5.8	Generalized Angularities in 3-jet events	82
5.9	Gluon splitting	85
5.9.1	<i>Variables</i>	85
5.9.2	<i>Efficiencies</i>	87
5.9.3	<i>Closure Test</i>	87
5.9.4	<i>Results</i>	90
5.10	<i>Discussion</i>	92
5.11	<i>Conclusion</i>	94
A	<i>Housing Prices Appendix</i>	95
A.1	<i>Input Distributions</i>	95
A.2	<i>Correlations</i>	110
A.3	<i>Validity Correlations</i>	111
A.4	<i>Energy Rating Mapping</i>	112
A.5	<i>Facebook Prophet for Houses</i>	112
A.6	<i>Final Variable List</i>	113
A.7	<i>Initial Hyperparameter Results Tables</i>	115
A.8	<i>Initial Hyperparameter Results Parallel Coordinates</i> .	118
A.9	<i>Hyperparameter Results Parallel Coordinates</i>	118
A.10	<i>Hyperparameter Optimization as a Function of Iterations</i>	119
A.11	<i>Histogram of z Values for Apartments with the Tight Cut</i>	121
A.12	<i>Performance Metrics for the Tight Selection</i>	121
A.13	<i>Multiple Models for Houses</i>	121
B	<i>Quarks vs. Gluons Appendix</i>	123
B.1	<i>Fraction of Different Types of Jets</i>	123
B.2	<i>Random Search PDFs for XGB</i>	123
B.3	<i>UMAP Parameters Comparison</i>	124
B.4	<i>t-SNE Parameters Comparison</i>	124
B.5	<i>HPO results for 3-jet b-Tagging</i>	125
B.6	<i>HPO results for 4-jet b-Tagging</i>	125
B.7	<i>b-Tag Scores in 3-Jet Events</i>	126
B.8	<i>b-Tag Scores in 4-Jet Events</i>	126

<i>B.9 ROC curve for 3-jet b-tagging</i>	127
<i>B.10 ROC curve for 4-jet b-tagging</i>	127
<i>B.11 Distribution of b-Tags in 3-Jet Events</i>	128
<i>B.12 Distribution of b-Tags in 4-Jet Events</i>	128
<i>B.13 Global Feature Importances for the LGB b-Tagging Algorithm on 3-Jet Events</i>	129
<i>B.14 Global Feature Importances for the LGB b-Tagging Algorithm on 4-Jet Events</i>	129
<i>B.15 Parallel Coordinate Plot of HPO Results for 3-Jet g-Tagging for Energy Ordered Jets</i>	130
<i>B.16 Parallel Coordinate Plot of HPO Results for 3-Jet g-Tagging for Shuffled Jets</i>	130
<i>B.17 Parallel Coordinate Plot of HPO Results for 4-Jet g-Tagging for Energy Ordered Jets</i>	131
<i>B.18 Parallel Coordinate Plot of HPO Results for 4-Jet g-Tagging for Shuffled Jets</i>	131
<i>B.19 PermNet Architecture</i>	132
<i>List of Figures</i>	159
<i>List of Tables</i>	162
<i>Bibliography</i>	163

Part I

Part I of this thesis covers the introductory theory of machine learning in [chapter 2](#) along with some extra technical aspects of it. In [chapter 3](#) machine learning is applied to estimate Danish housing prices as precisely and accurately as possible.

Part II

Part II of this thesis deals with particle physics and the discriminatory power of machine learning for quark-gluon identification and subsequent analysis. In [chapter 4](#) the theory of the Standard Model is introduced together with a description of the ALEPH detector. Theory is applied in [chapter 5](#) where the types of jets and events in each collision is analysed using machine learning to improve the understanding of how gluon jets hadronizes and splits: simply said how they look and behave.

B. Quarks vs. Gluons Appendix

B.1 Fraction of Different Types of Jets

	<i>b</i>	<i>c</i>	<i>uds</i>	<i>g</i>	non- <i>q</i> -matched
2	37.2 %	12.9 %	29.1 %	0.0 %	20.7 %
3	22.6 %	8.9 %	19.7 %	31.2 %	17.5 %
4	14.6 %	7.0 %	15.0 %	45.1 %	18.3 %
5	10.0 %	5.7 %	12.2 %	52.5 %	19.6 %
6	7.1 %	4.4 %	8.8 %	54.4 %	25.2 %

Table B.1: Number of different types of jets for MC and MCb written in relative numbers such that each row sum to 100 %. See also Table 5.3.

B.2 Random Search PDFs for XGB

Hyperparameter	Range
<code>subsample</code>	$\mathcal{U}(0.4, 1)$
<code>colsample_bytree</code>	$\mathcal{U}_{\text{trunc}}(0.4, 1, 2)$
<code>max_depth</code>	$\mathcal{U}_{\text{int}}(1, 20)$
<code>min_child_weight</code>	$\mathcal{U}_{\text{int}}(0, 10)$

Table B.2: Probability Density Functions for the random search hyperparameter optimization process for the XGBoost model.

B.3 UMAP Parameters Comparison

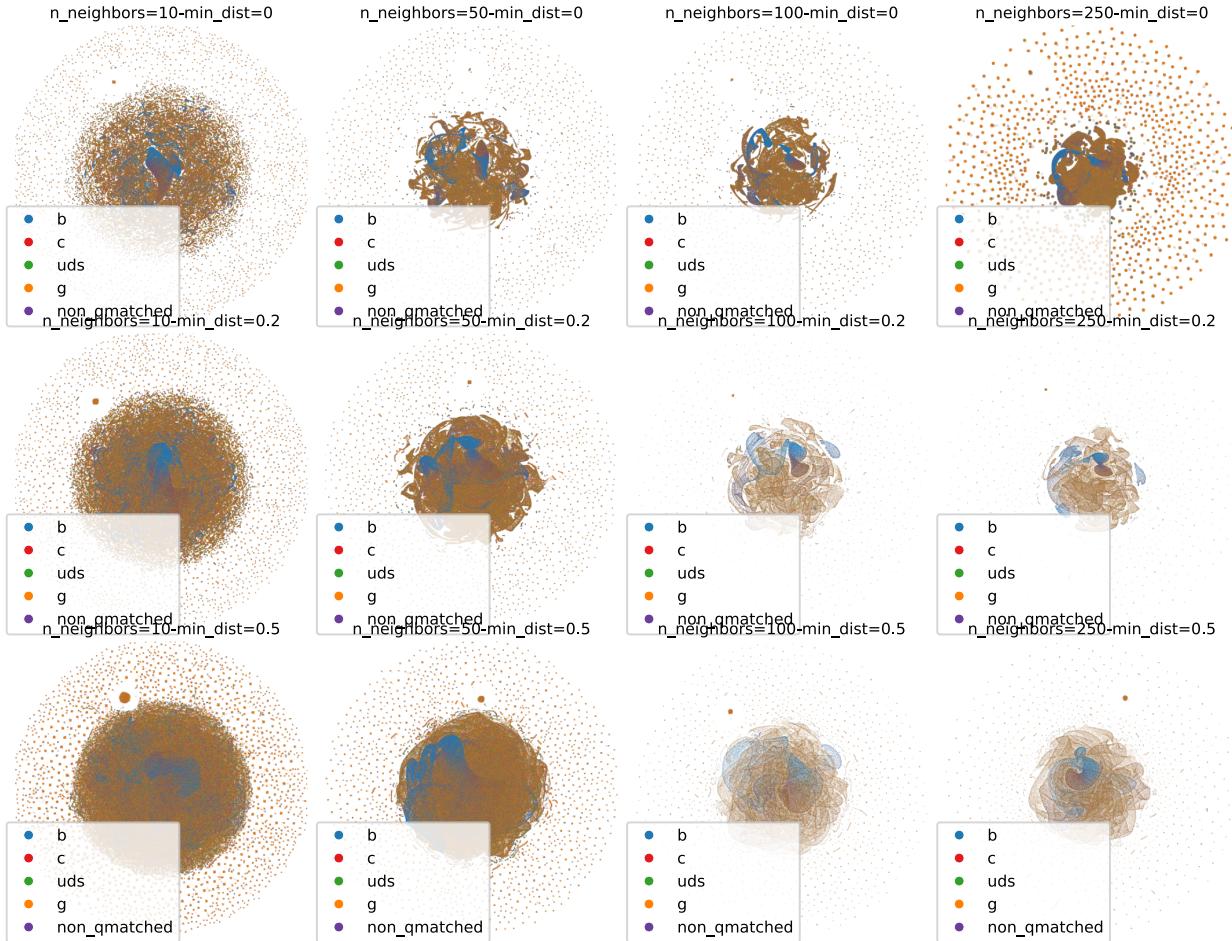


Figure B.1: Grid search of the two parameters `n_neighbors` and `min_dist` for the UMAP algorithm run on 4-jet events. For an explanation of these, see section 5.2.

B.4 t-SNE Parameters Comparison

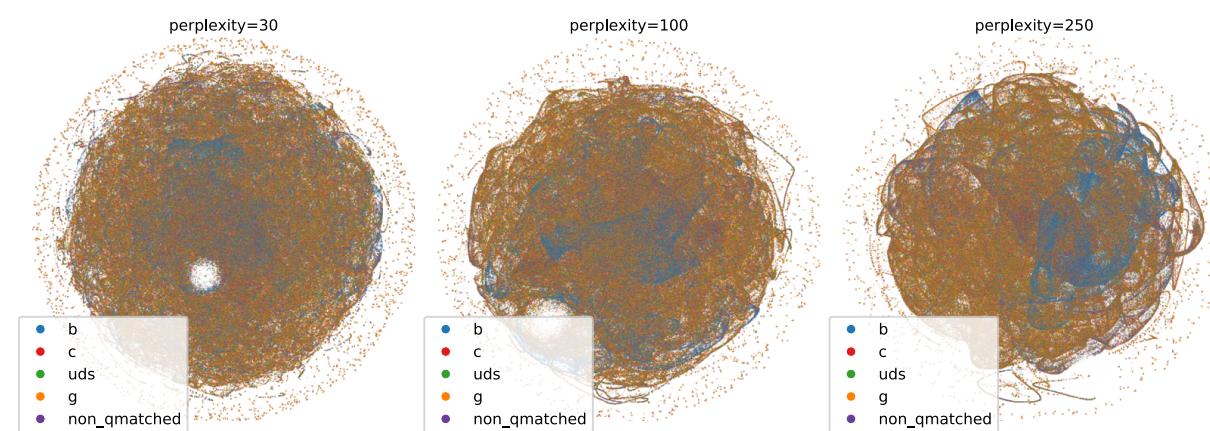


Figure B.2: Visualization of the t-SNE algorithm as a function of the `perplexity` parameters for 4-jet events.

B.5 HPO results for 3-jet b-Tagging

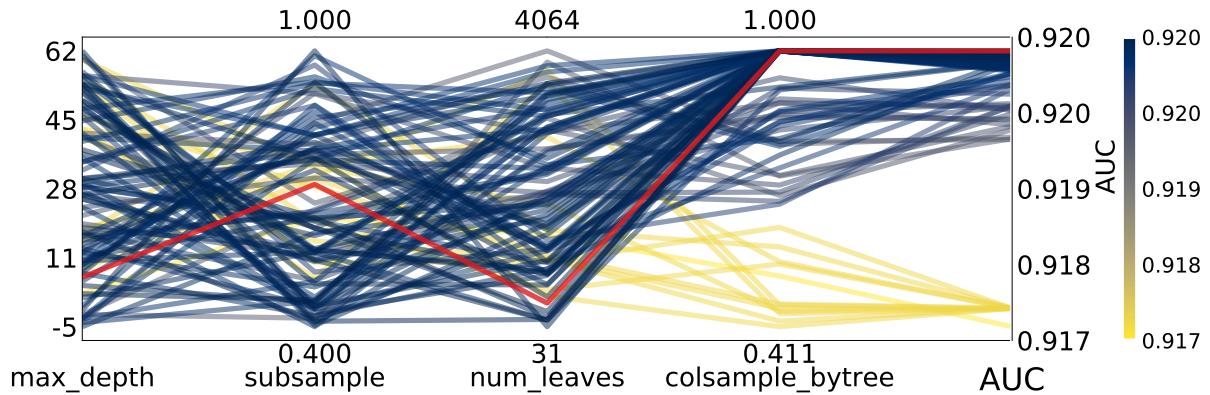


Figure B.3: Hyperparameter optimization results of b -tagging for 3-jet events. The results are shown as parallel coordinates with each hyperparameter along the x -axis and the value of that parameter on the y -axis. Each line is an event in the 4-dimensional space colored according to the performance of that hyperparameter as measured by AUC from highest AUC in dark blue to lowest AUC in yellow. The **single best hyperparameter** is shown in red.

B.6 HPO results for 4-jet b-Tagging

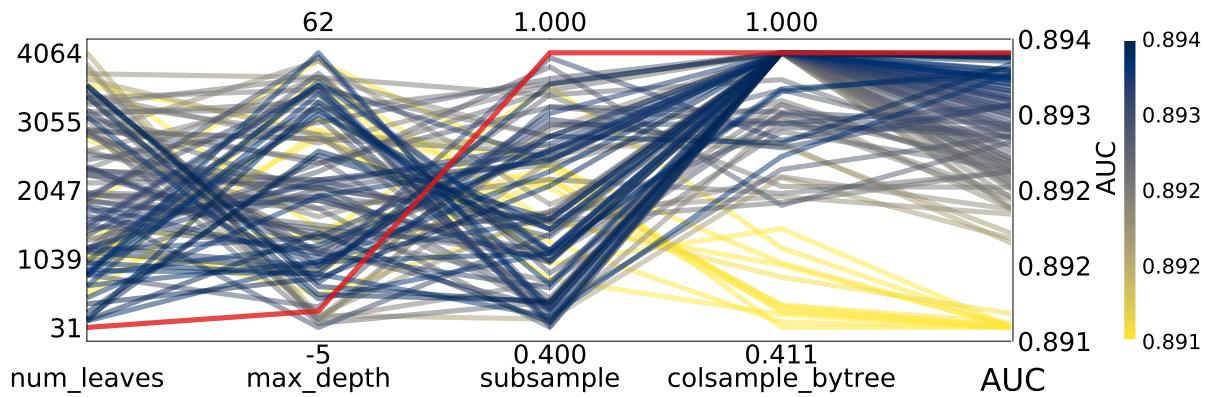
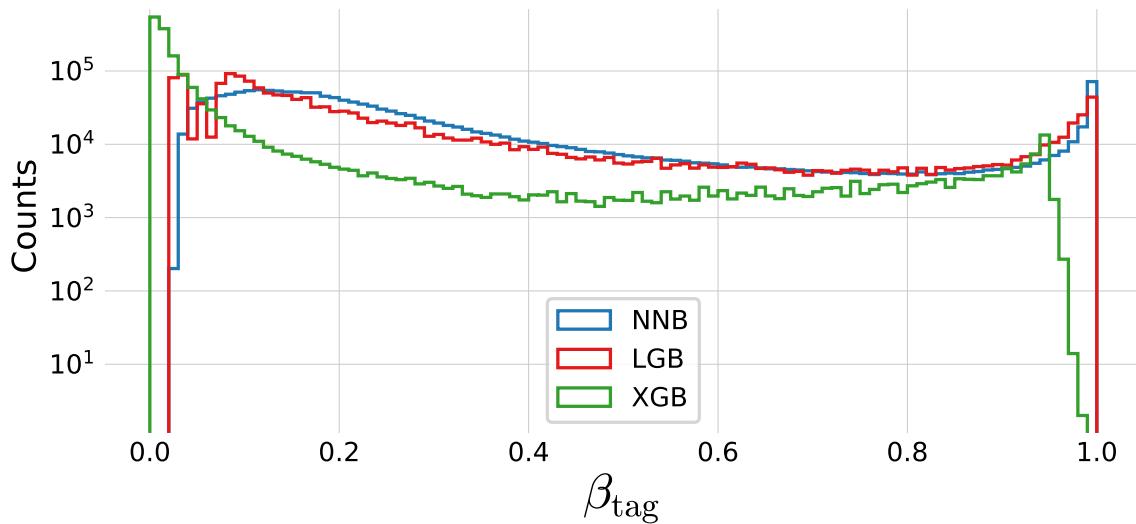


Figure B.4: Hyperparameter optimization results of b -tagging for 4-jet events. The results are shown as parallel coordinates with each hyperparameter along the x -axis and the value of that parameter on the y -axis. Each line is an event in the 4-dimensional space colored according to the performance of that hyperparameter as measured by AUC from highest AUC in dark blue to lowest AUC in yellow. The **single best hyperparameter** is shown in red.

B.7 *b*-Tag Scores in 3-Jet Events



B.8 *b*-Tag Scores in 4-Jet Events

Figure B.5: Histogram of b -tag scores β_{tag} in 3-jet events for **NNB** (the neural network pre-trained by ALEPH, also called `nnbjet`) in blue, **LGB** in red, and **XGB** in green.

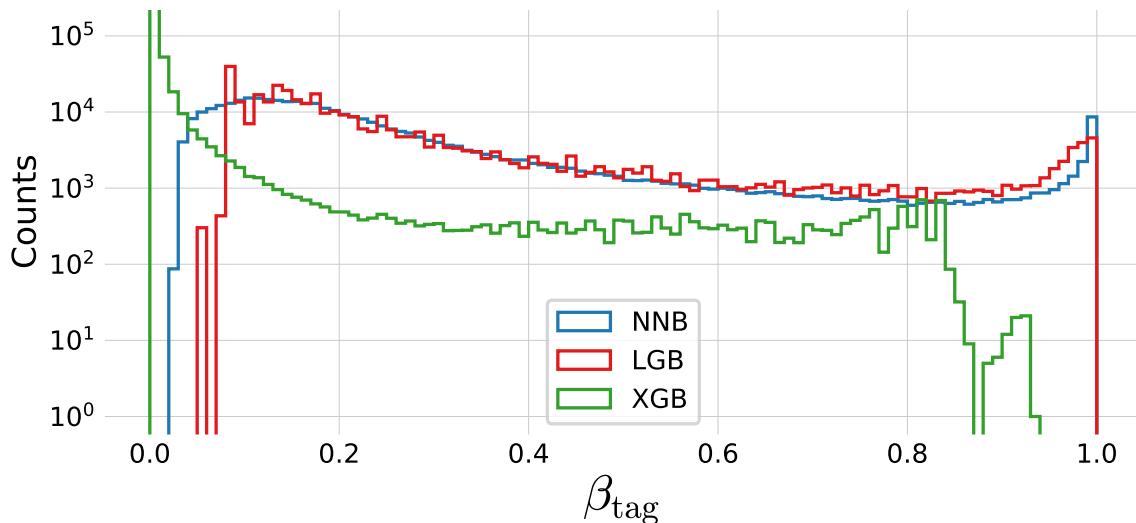
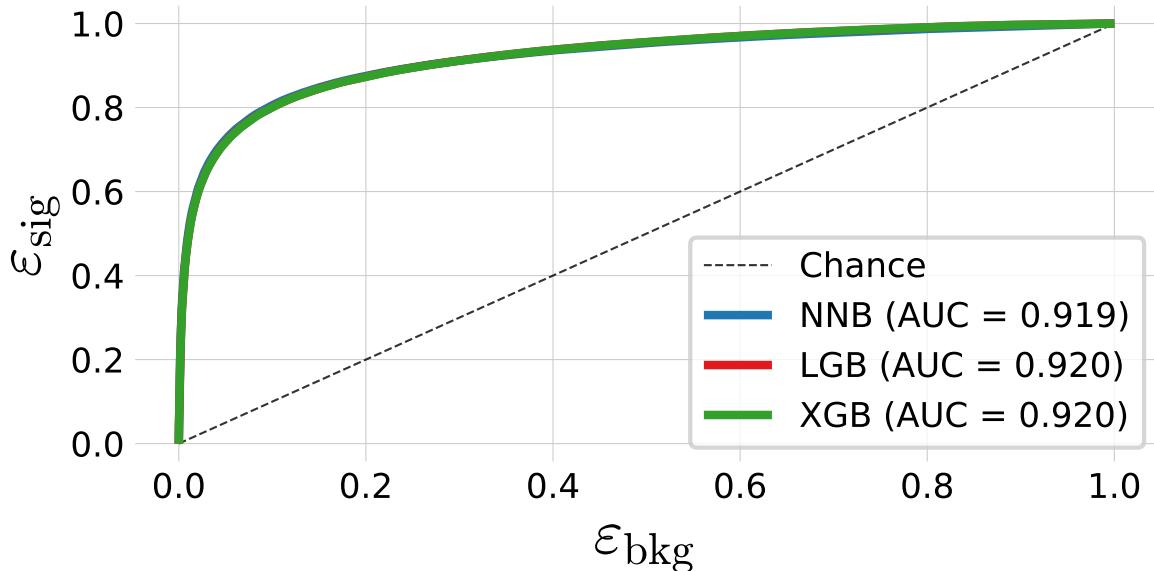


Figure B.6: Histogram of b -tag scores β_{tag} in 4-jet events for **NNB** (the neural network pre-trained by ALEPH, also called `nnbjet`) in blue, **LGB** in red, and **XGB** in green.

B.9 ROC curve for 3-jet b -tagging



B.10 ROC curve for 4-jet b -tagging

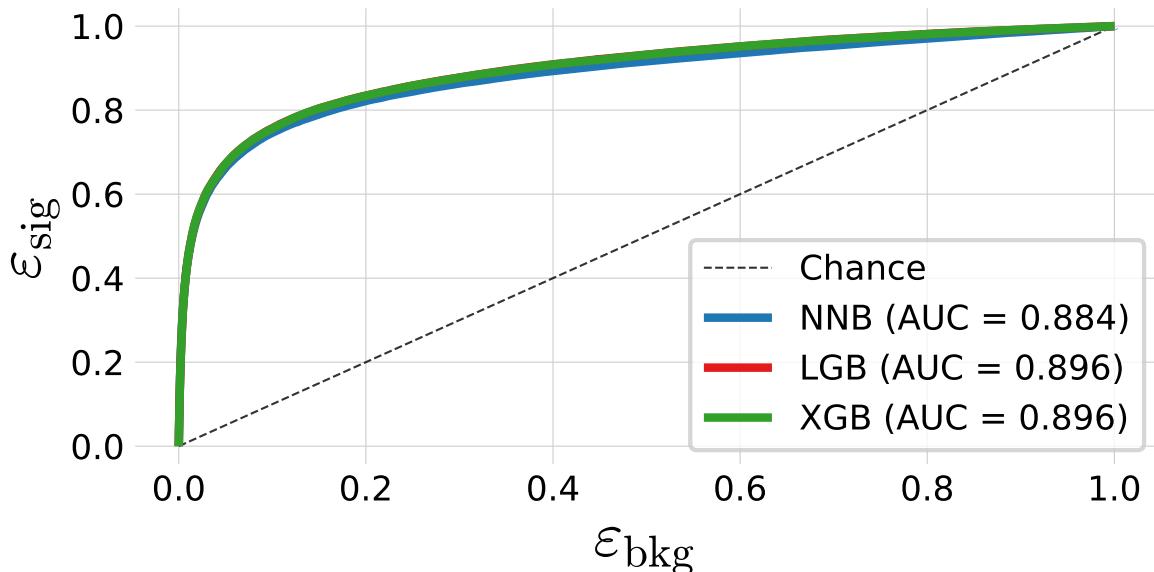


Figure B.7: ROC curve of the three b -tag models in 3-jet events for **NNB** (the pre-trained neural network trained by ALEPH, also called `nnbjet`) in blue, **LGB** in red, and **XGB** in green. In the legend the area under curve (AUC) is also shown. Notice that the LGB and XGB models share performance and it is thus due to overplotting that only the green line for XGB can be seen.

Figure B.8: ROC curve of the three b -tag models in 4-jet events for **NNB** (the pre-trained neural network trained by ALEPH, also called `nnbjet`) in blue, **LGB** in red, and **XGB** in green. In the legend the area under curve (AUC) is also shown. Notice that the LGB and XGB models share performance and it is thus due to overplotting that only the green line for XGB can be seen.

B.11 Distribution of b -Tags in 3-Jet Events

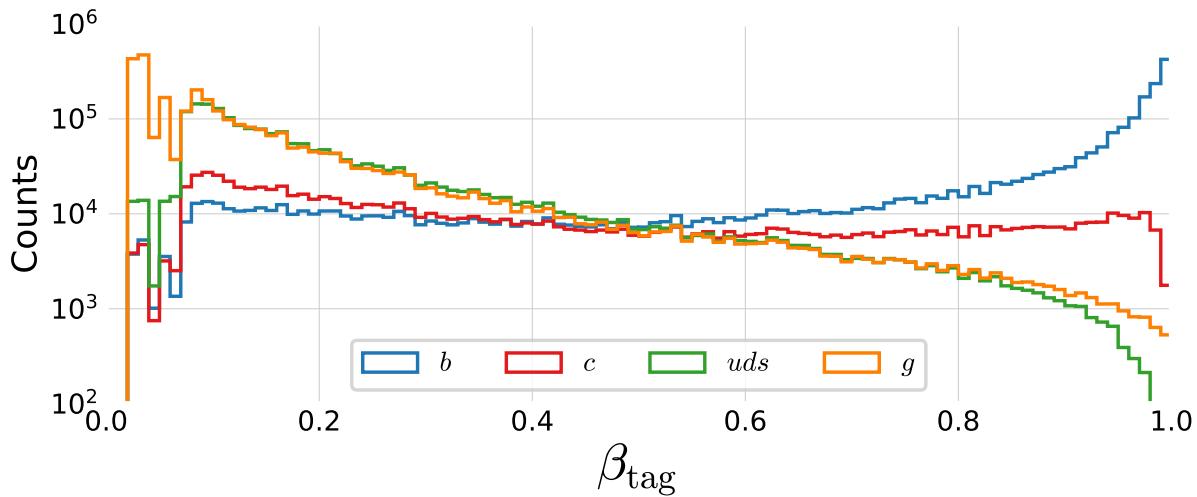


Figure B.9: Distribution of b -tags in 3-jet events for b -jets in blue, c -jets in red, uds in green and g in orange.

B.12 Distribution of b -Tags in 4-Jet Events

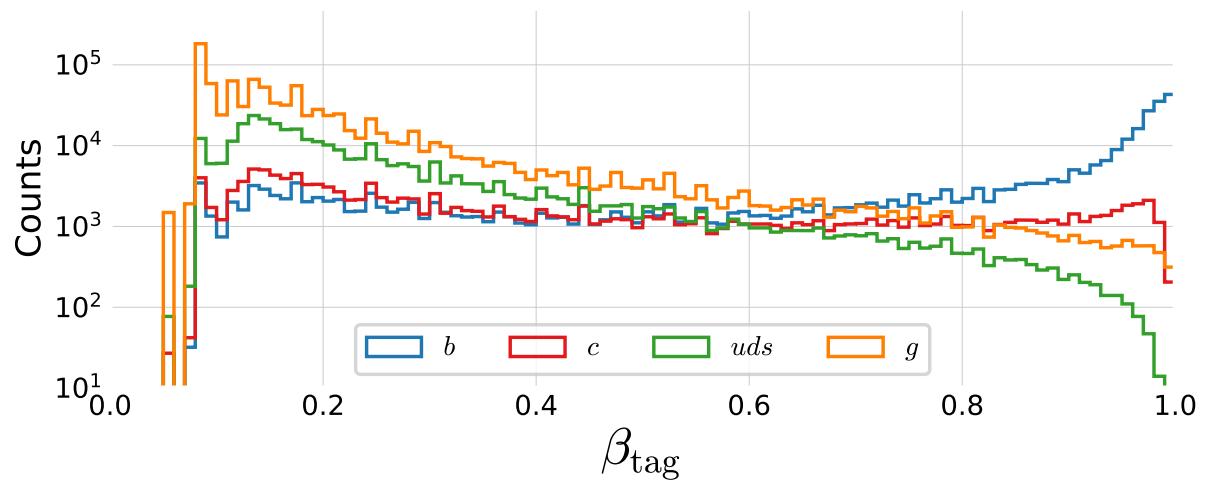


Figure B.10: Distribution of b -tags in 4-jet events for b -jets in blue, c -jets in red, uds in green and g in orange.

B.13 Global Feature Importances for the LGB b-Tagging Algorithm on 3-Jet Events

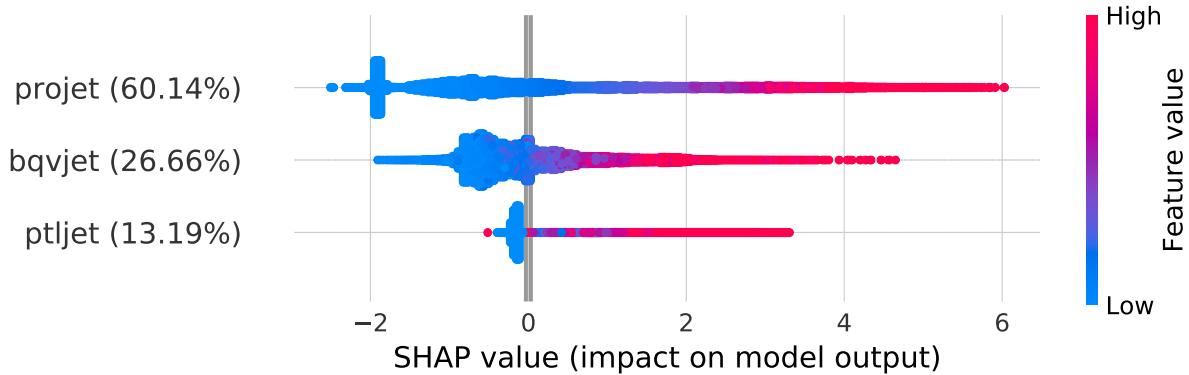


Figure B.11: Global feature importances for the LGB b -tagging algorithm on 3-jet events. The normalized feature importance is shown in the parenthesis and the each dot is an observation showing the dependance between the SHAP value and the feature's value.

B.14 Global Feature Importances for the LGB b-Tagging Algorithm on 4-Jet Events

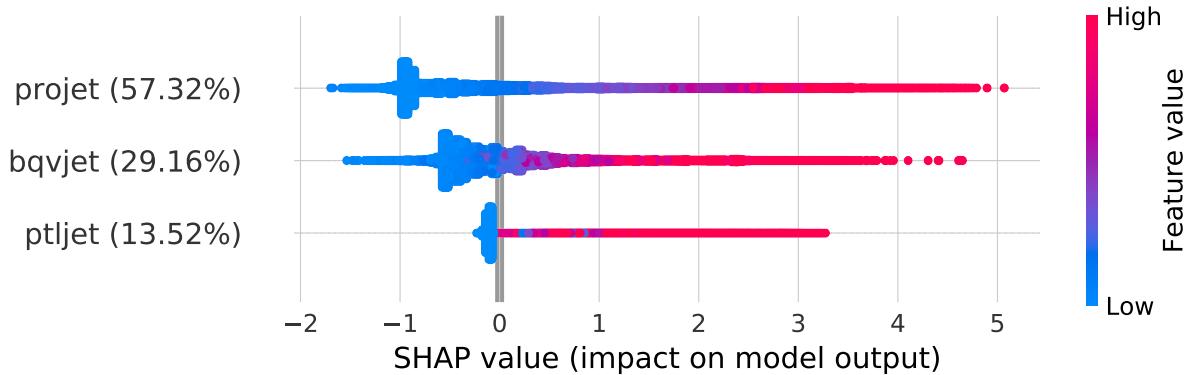


Figure B.12: Global feature importances for the LGB b -tagging algorithm on 4-jet events. The normalized feature importance is shown in the parenthesis and the each dot is an observation showing the dependance between the SHAP value and the feature's value.

B.15 Parallel Coordinate Plot of HPO Results for 3-Jet g-Tagging for Energy Ordered Jets

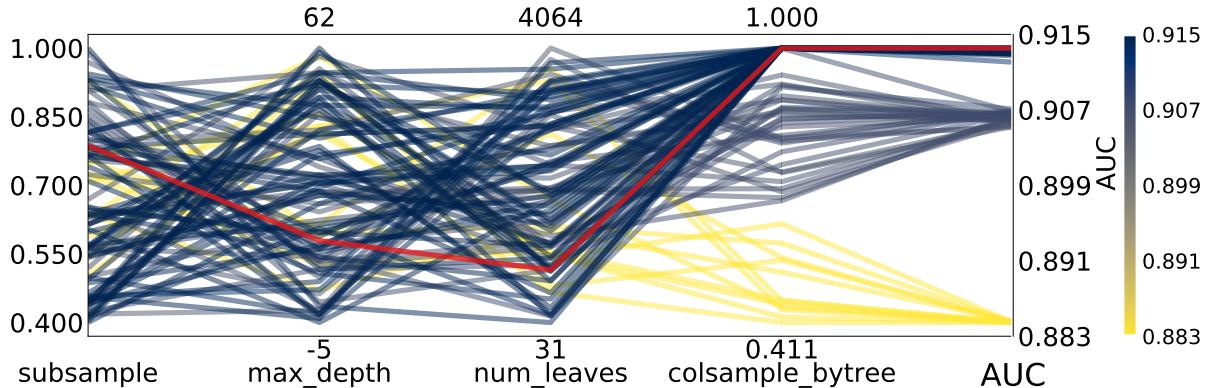


Figure B.13: Hyperparameter optimization results of g -tagging for 3-jet events for energy ordered jets shown as parallel coordinates.

B.16 Parallel Coordinate Plot of HPO Results for 3-Jet g-Tagging for Shuffled Jets

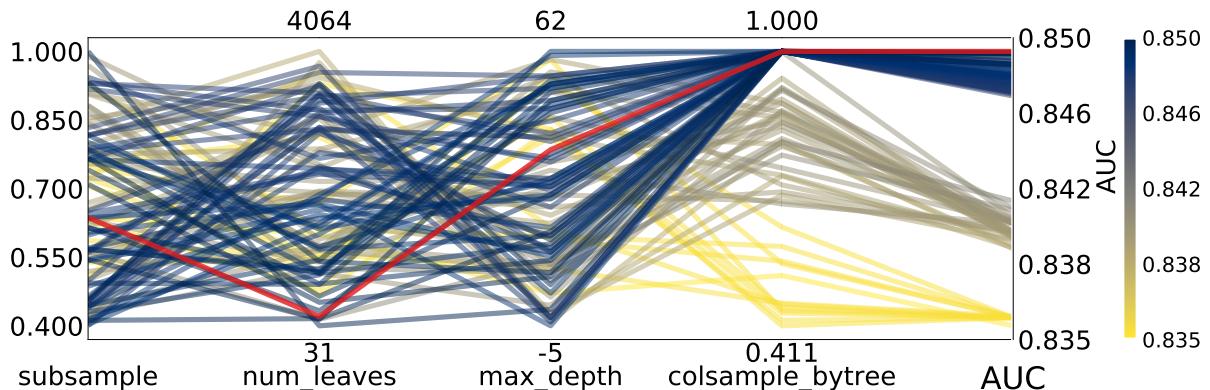


Figure B.14: Hyperparameter optimization results of g -tagging for 3-jet events for (row) shuffled jets shown as parallel coordinates.

B.17 Parallel Coordinate Plot of HPO Results for 4-Jet g -Tagging for Energy Ordered Jets

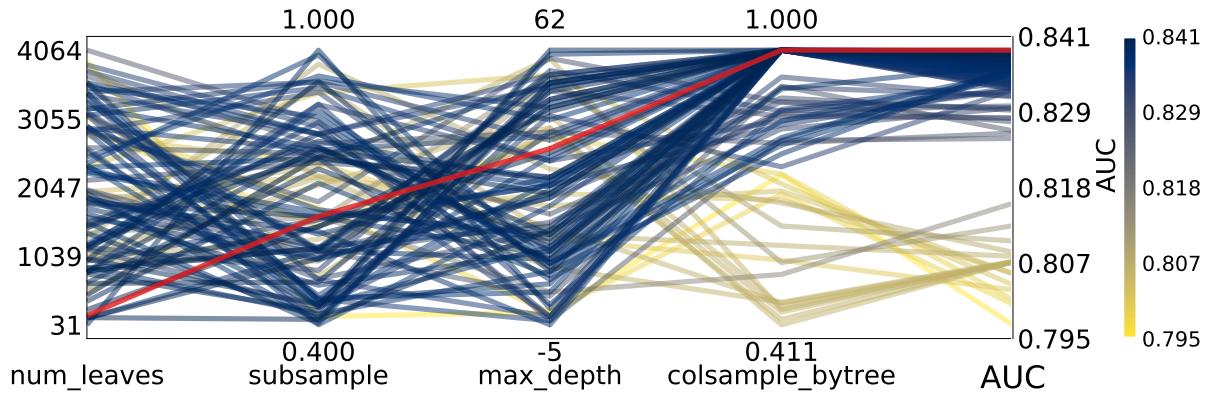


Figure B.15: Hyperparameter optimization results of g -tagging for 4-jet events for energy ordered jets shown as parallel coordinates.

B.18 Parallel Coordinate Plot of HPO Results for 4-Jet g -Tagging for Shuffled Jets

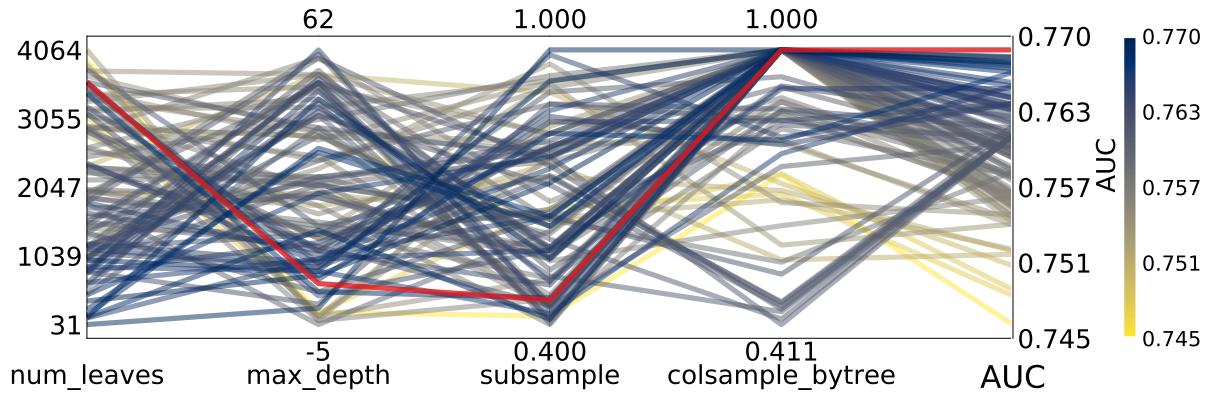


Figure B.16: Hyperparameter optimization results of g -tagging for 4-jet events for (row) shuffled jets shown as parallel coordinates.

B.19 PermNet Architecture

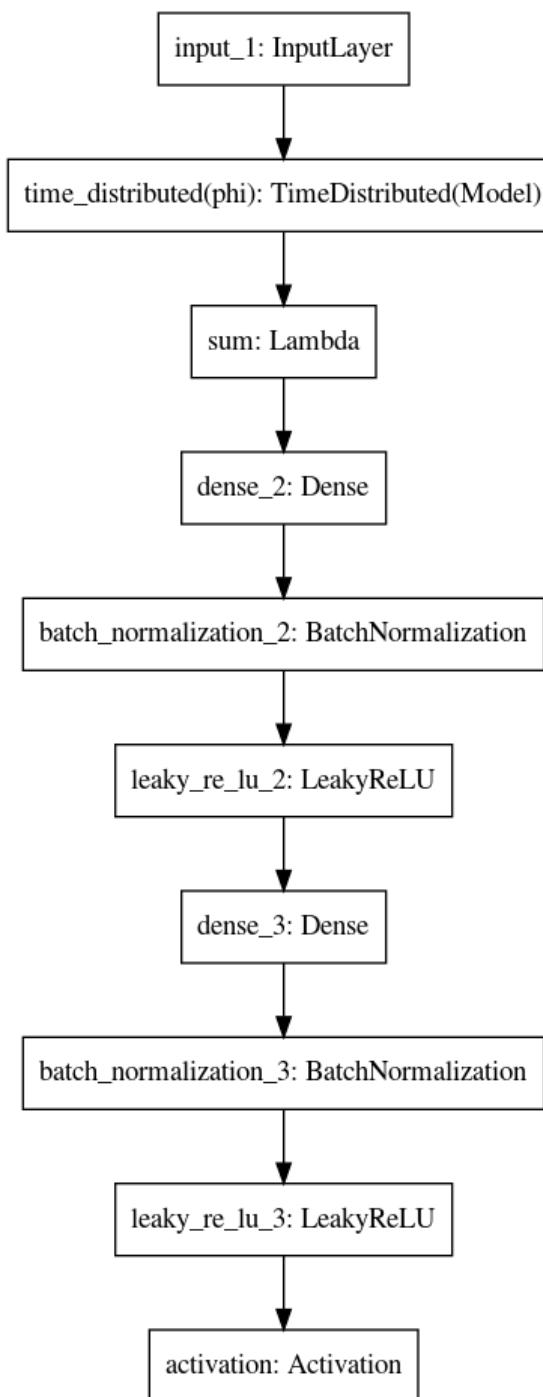


Figure B.17: Architecture of the PermNet neural network.

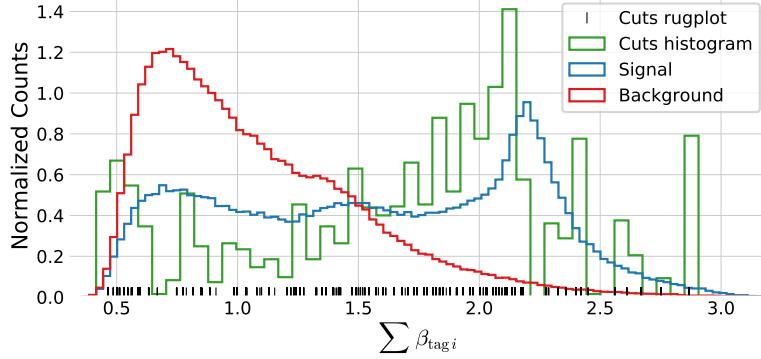


Figure B.18: Histogram of the distribution of [signal](#) in blue and [background](#) in red for the 1-dimensional sum of b -tags for 4-jet events. A histogram of the [cut values](#) from the LGB model trained on this data is shown in green together with a rug plot of the cut values in black. Notice how most of the cuts match up with the signal peak at around a $\sum \beta_i \sim 2.1$.

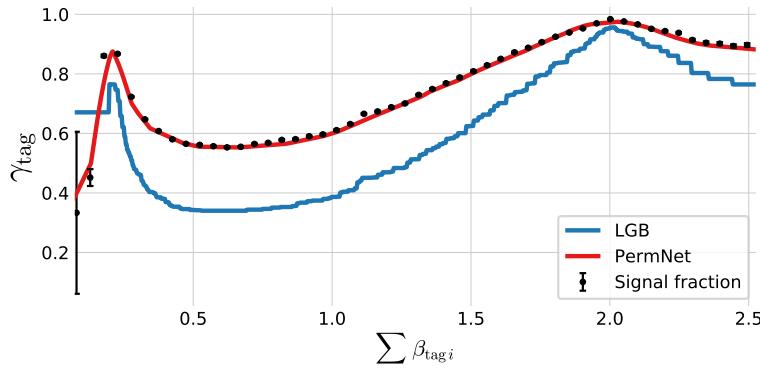


Figure B.19: Plot of the (1D) [g](#)-tag scores for 3-jet events as a function of $\sum \beta_i$ for the [LGB](#) model in blue and the [PermNet](#) model in red. The signal fraction (based on the signal and background histograms in Figure B.20) is plotted as black error bars where the size of the error bars is based on the propagated uncertainties of the signal and background histogram assuming Poissonian statistics.

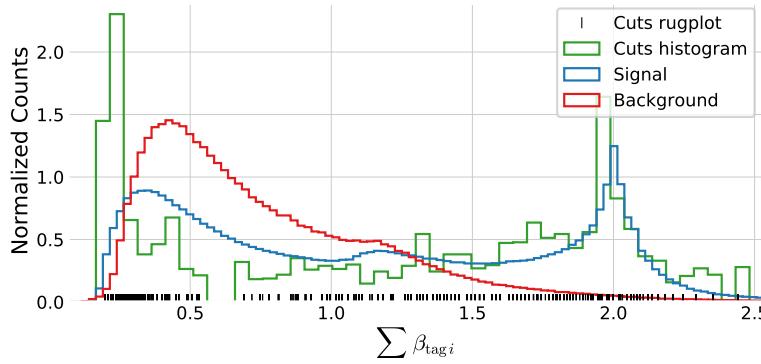


Figure B.20: Histogram of the distribution of [signal](#) in blue and [background](#) in red for the 1-dimensional sum of b -tags for 3-jet events. A histogram of the [cut values](#) from the LGB model trained on this data is shown in green together with a rug plot of the cut values in black. Notice how most of the cuts match up with the signal peak at around a $\sum \beta_i \sim 2.1$.

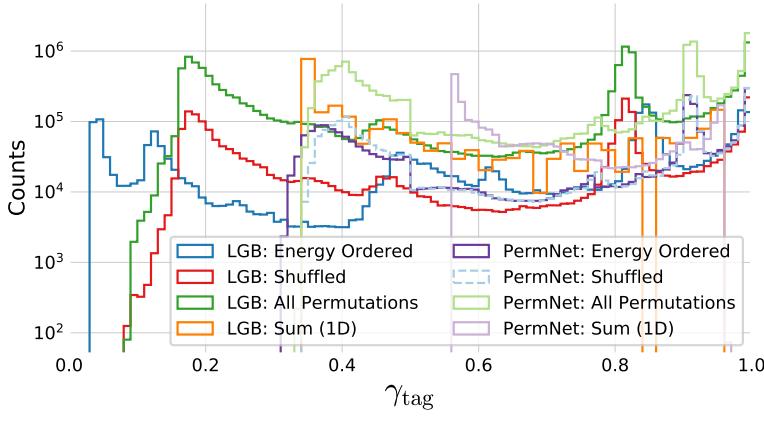


Figure B.21: Distribution of g -tag scores in 3-jet events shown with a logarithmic y -scale for LGB: Energy Ordered in blue, LGB: Shuffled in red, LGB: All Permutations in green, LGB: Sum 1D in orange, PermNet: Energy Ordered in purple, PermNet: Shuffled in light-blue, PermNet: All Permutations in light-green, PermNet: Sum 1D in light-purple. Here LGB and PermNet are the two different type of models and “Energy Ordered”, “Shuffled”, “All Permutations”, and “Sum 1D” are the different methods used for making the input data permutation invariant (except energy ordered).

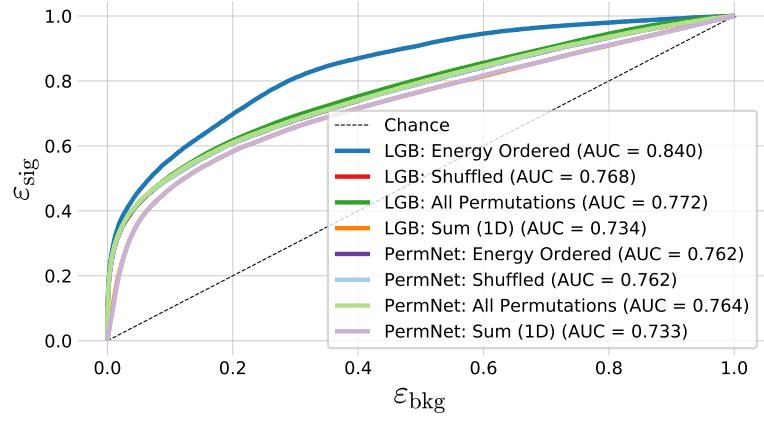


Figure B.22: ROC curve of the eight g -tag models in 4-jet events. First one in dashed black is the ROC curve that you get by random chance. The colors are the same as in Figure 5.20 and in the legend also the Area Under the ROC curve (AUC) is shown. Notice that the XGB model which uses the energy ordered data produced the best model, however, this model is not permutation invariant. Of the permutation invariant models (the rest), the XGB model trained on all permutations of the b -tags performs highest. The lowest performing models are the two models trained only on the 1-dimensional sum of b -tags, as expected, however, still with a better performance than expected by the author.

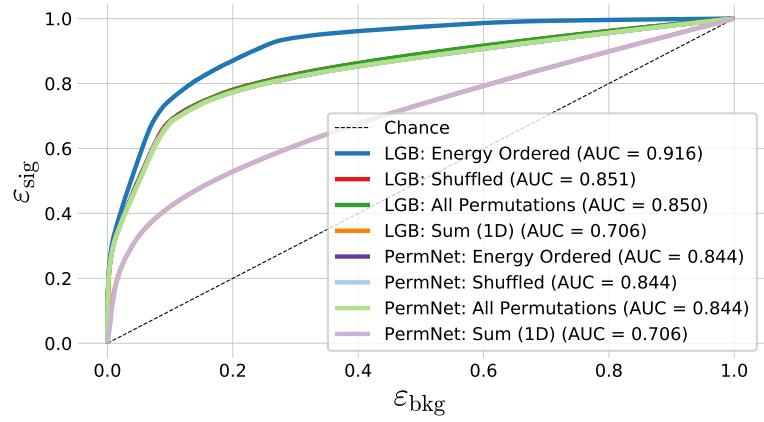


Figure B.23: ROC curve of the eight g -tag models in 3-jet. First one in dashed black is the ROC curve that you get by random chance. The colors are the same as in Figure 5.20 and in the legend also the Area Under the ROC curve (AUC) is shown.

β_{tag_i}	Energy Ordered	Shuffled	All Permutations
1	0.827 ± 0.006	0.924 ± 0.006	0.923 ± 0.006
2	0.749 ± 0.006	0.909 ± 0.006	0.918 ± 0.005
3	1.198 ± 0.006	0.878 ± 0.005	0.906 ± 0.005

Table B.3: Global SHAP feature importances $\phi_{\beta_i}^{\text{tot}}$ for the three g -Tagging Models in 3-Jet Events. Each $\phi_{\beta_i}^{\text{tot}}$ is shown for the three methods in the columns and the three b -tags as variables in the rows.

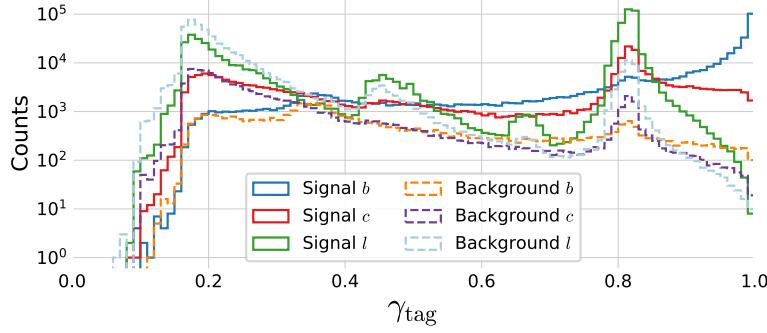


Figure B.24: Histogram of g -tag scores from the LGB-model in 3-jet events for b signal in blue, c signal in red, l (uds) signal in green, b background in orange, c background in purple, l (uds) background in light-blue.

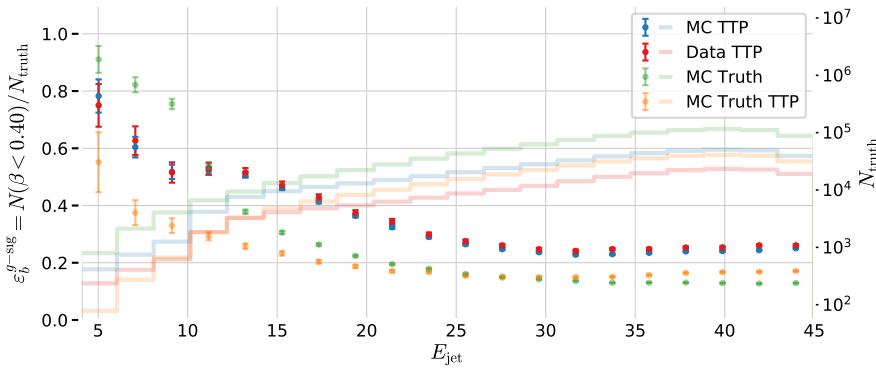


Figure B.25: b -tag efficiency for b -jets in the g -signal region for 3-jet events, $\varepsilon_b^{g-\text{sig}}$, as a function of jet energy E_{jet} . In the plot the efficiencies are shown for MC TTP in blue, Data TTP in red, MC Truth TTP in green, and MC Truth in orange. The efficiencies (the errorbars) can be read off on the left y -axis and the counts (histograms) on the right y -axis.

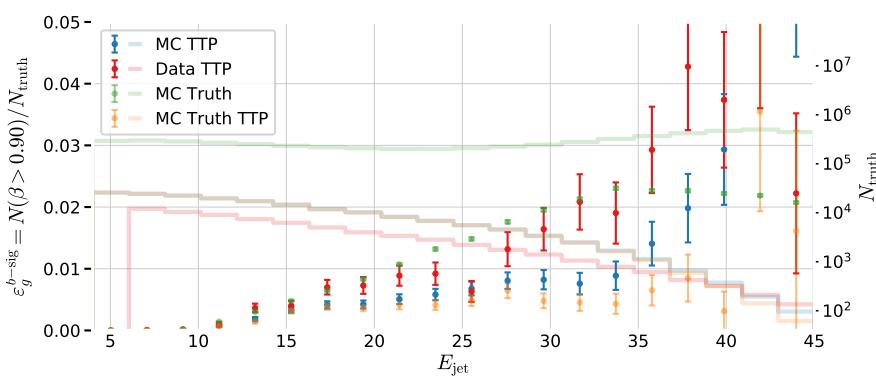


Figure B.26: b -tag efficiency for b -jets in the b -signal region for 3-jet events, $\varepsilon_g^{b-\text{sig}}$, as a function of jet energy E_{jet} . In the plot the efficiencies are shown for MC TTP in blue, Data TTP in red, MC Truth TTP in green, and MC Truth in orange. The efficiencies (the errorbars) can be read off on the left y -axis and the counts (histograms) on the right y -axis.

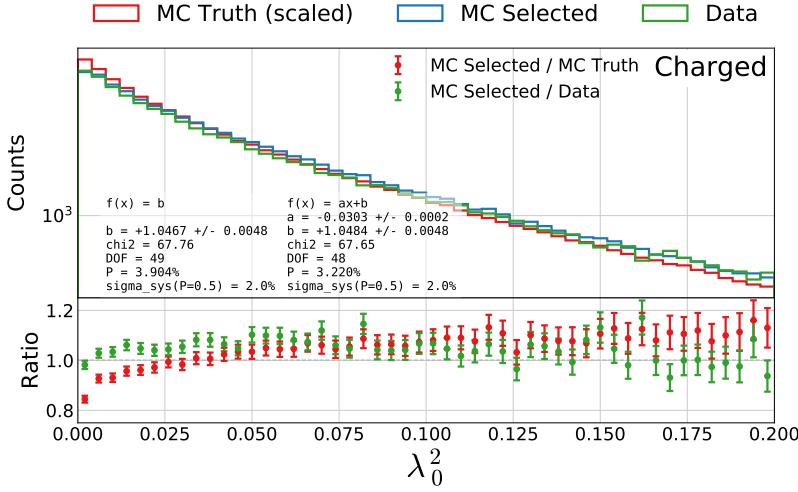


Figure B.27: Distribution of the generalized angularity λ_0^2 for charged gluons jets in 3-jet events. The distributions for **MC Truth** is shown in red, **MC Truth** in blue, and **Data** in green in the top plot and in the bottom plot the ratio between **MC Selected** and **MC Truth** is shown in red and between **MC Selected** and **Data** in green.

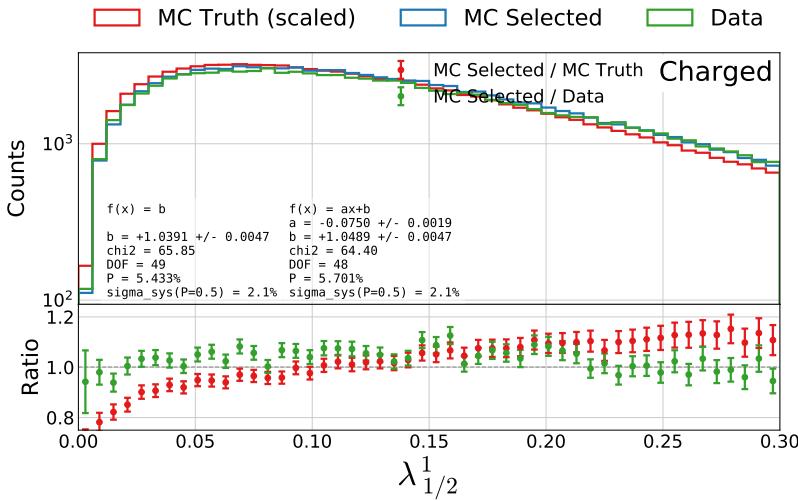


Figure B.28: Distribution of the generalized angularity $\lambda_{1/2}^2$ for charged gluons jets in 3-jet events. The distributions for **MC Truth** is shown in red, **MC Truth** in blue, and **Data** in green in the top plot and in the bottom plot the ratio between **MC Selected** and **MC Truth** is shown in red and between **MC Selected** and **Data** in green.

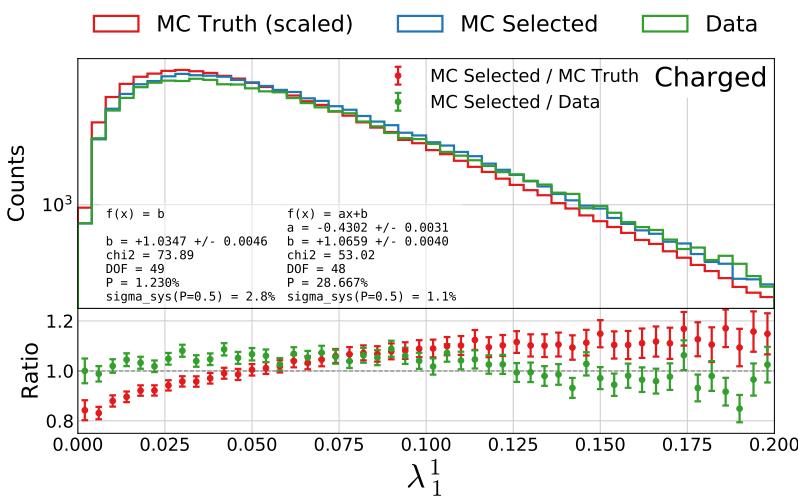
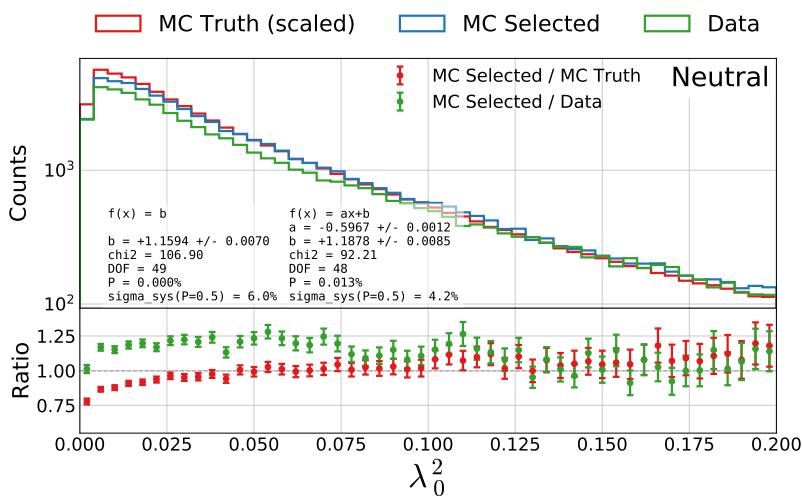
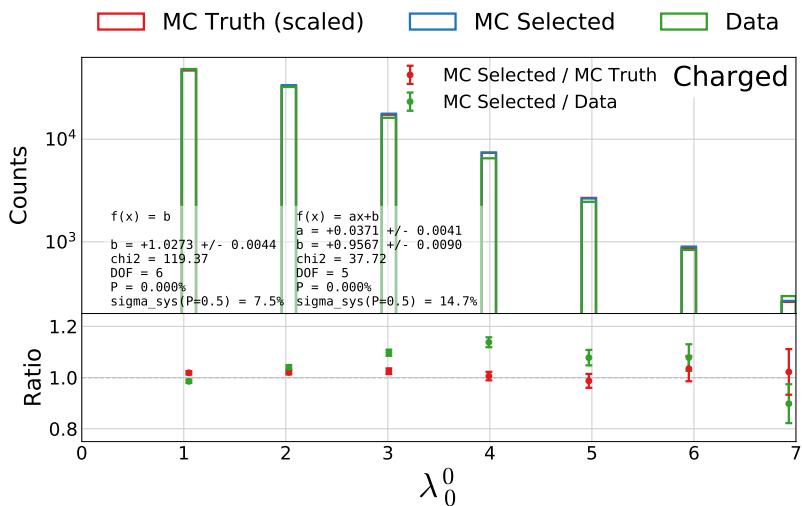
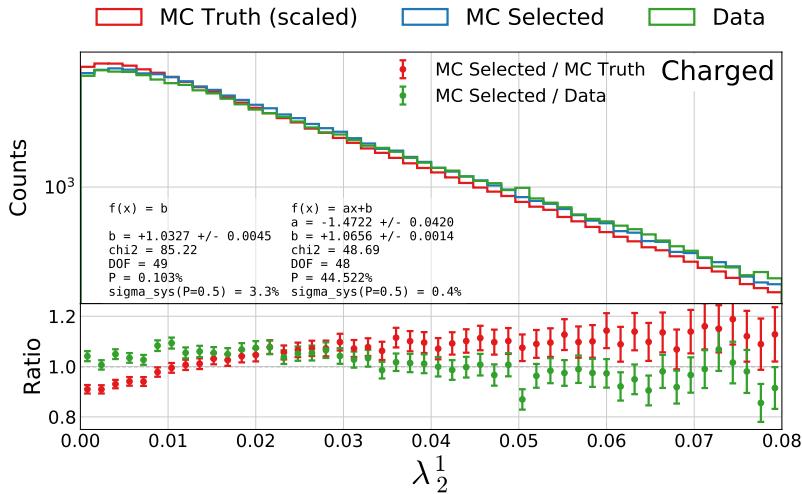
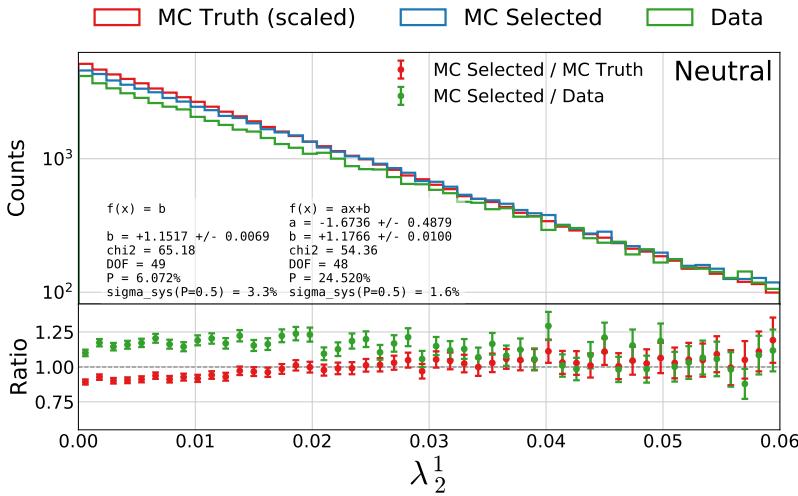
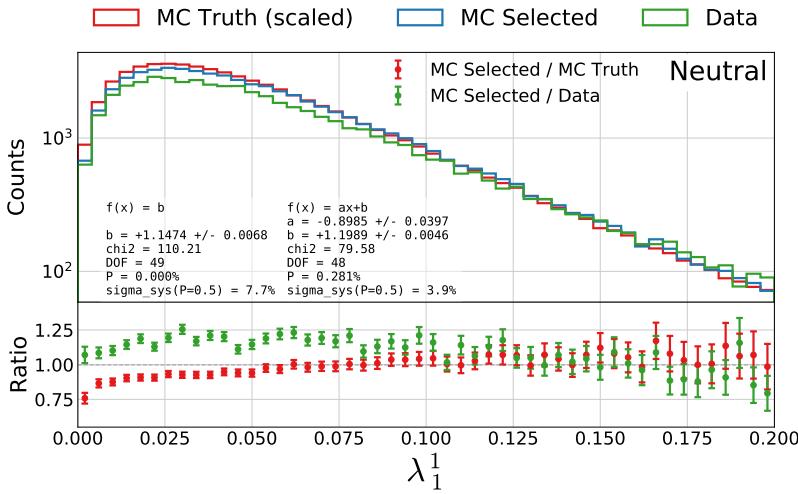
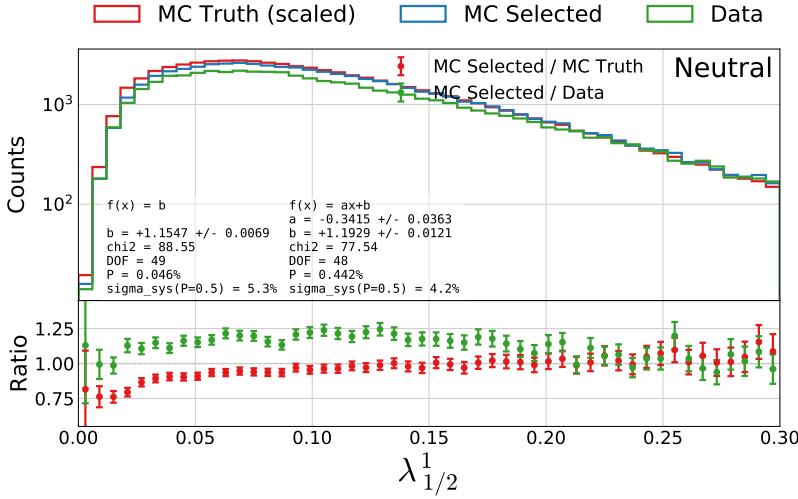


Figure B.29: Distribution of the generalized angularity λ_1^1 for charged gluons jets in 3-jet events. The distributions for **MC Truth** is shown in red, **MC Truth** in blue, and **Data** in green in the top plot and in the bottom plot the ratio between **MC Selected** and **MC Truth** is shown in red and between **MC Selected** and **Data** in green.





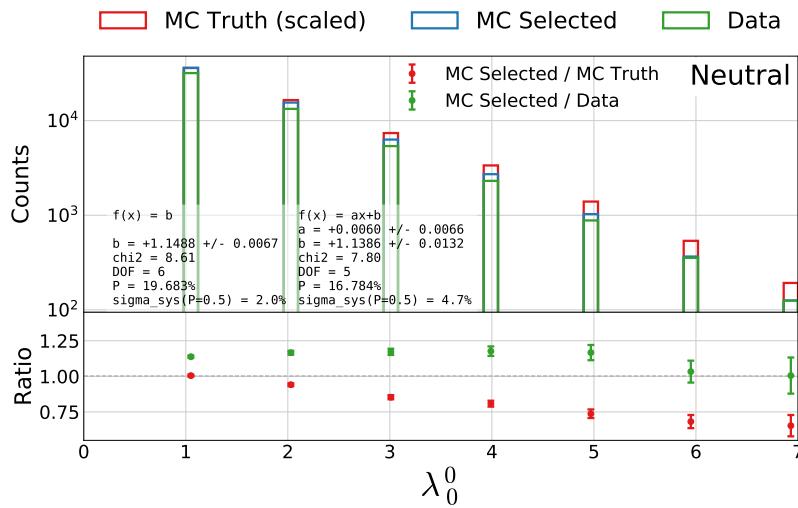


Figure B.36: Distribution of the generalized angularity λ_0^0 for neutral gluons clusters in 3-jet events. The distributions for **MC Truth** is shown in red, **MC Selected** in blue, and **Data** in green in the top plot and in the bottom plot the ratio between **MC Selected and MC Truth** is shown in red and between **MC Selected and Data** in green.

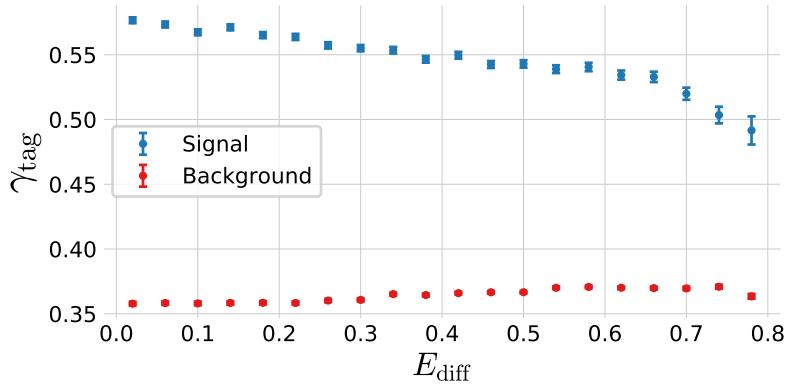


Figure B.37: Relationship between the g -tag value γ_{tag} and the gluon splitting variable E_{diff} . The **signal events** (according to MC Truth) are plotted in blue and **background events** (according to MC Truth) in red.

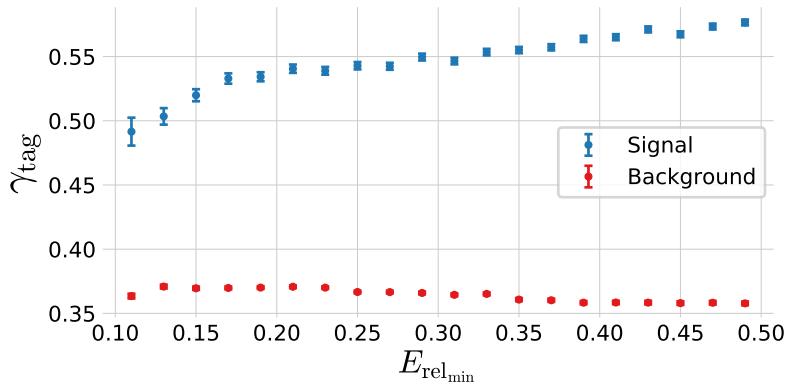


Figure B.38: Relationship between the g -tag value γ_{tag} and the gluon splitting variable $E_{\text{rel,min}}$. The **signal events** (according to MC Truth) are plotted in blue and **background events** (according to MC Truth) in red.

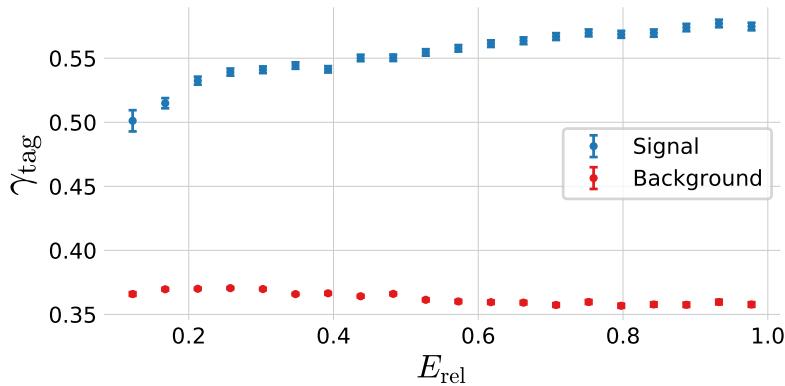


Figure B.39: Relationship between the g -tag value γ_{tag} and the gluon splitting variable E_{rel} . The **signal events** (according to MC Truth) are plotted in blue and **background events** (according to MC Truth) in red.

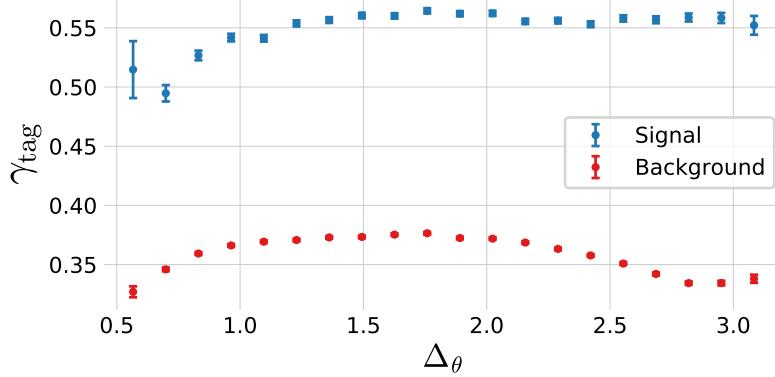


Figure B.40: Relationship between the g -tag value γ_{tag} and the gluon splitting variable Δ_θ . The **signal events** (according to MC Truth) are plotted in blue and **background events** (according to MC Truth) in red.

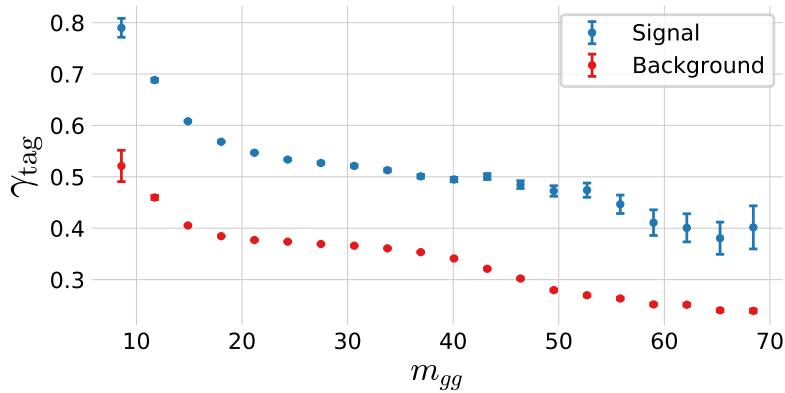


Figure B.41: Relationship between the g -tag value γ_{tag} and the gluon splitting variable m_{gg} . The **signal events** (according to MC Truth) are plotted in blue and **background events** (according to MC Truth) in red.

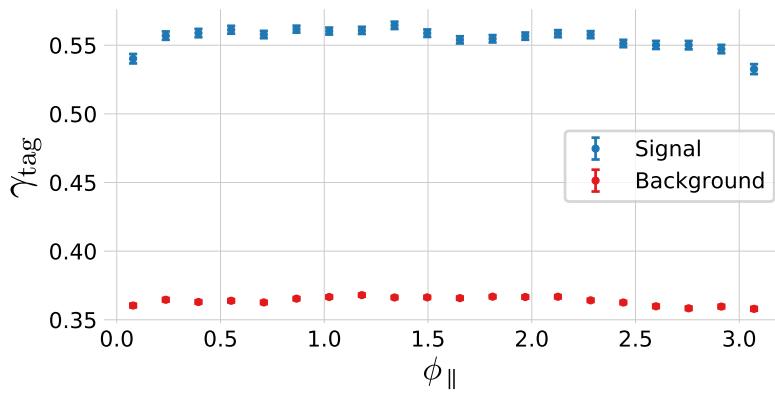


Figure B.42: Relationship between the g -tag value γ_{tag} and the gluon splitting variable ϕ_{\parallel} . The **signal events** (according to MC Truth) are plotted in blue and **background events** (according to MC Truth) in red.

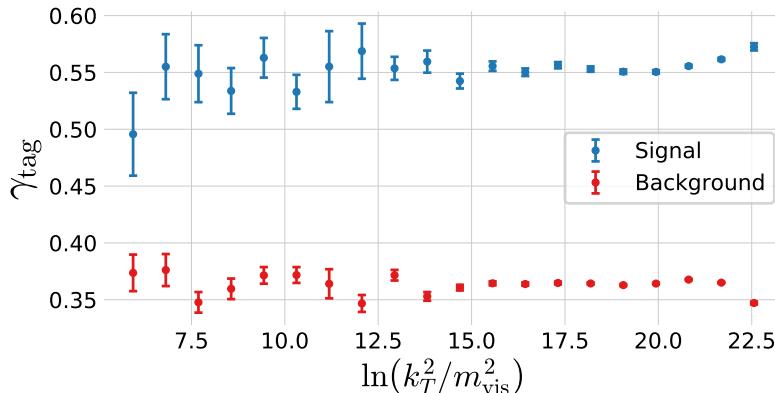


Figure B.43: Relationship between the g -tag value γ_{tag} and the gluon splitting variable $\ln(k_T^2/m_{\text{vis}}^2)$. The **signal events** (according to MC Truth) are plotted in blue and **background events** (according to MC Truth) in red.

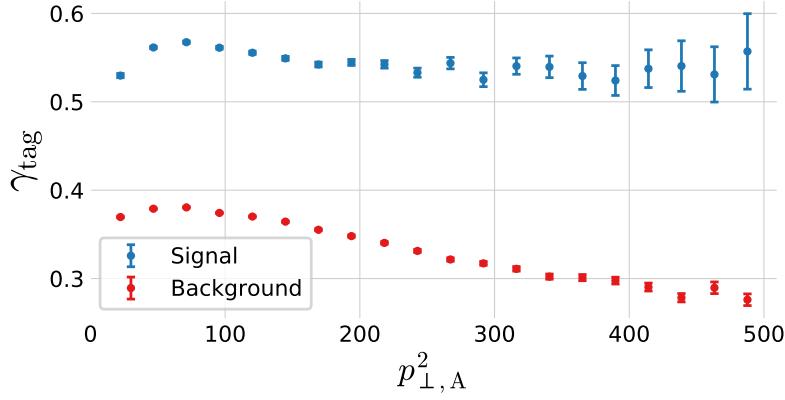


Figure B.44: Relationship between the g -tag value γ_{tag} and the gluon splitting variable $p_{\perp A}^2$. The **signal events** (according to MC Truth) are plotted in blue and **background events** (according to MC Truth) in red.

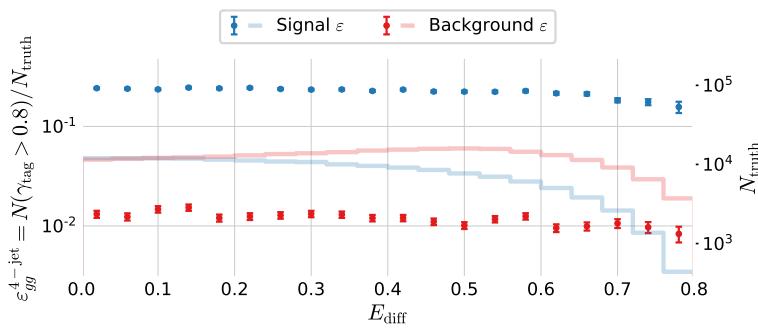


Figure B.45: Efficiency of the g -tagging algorithm for 4-jet events as a function of normalized gluon-gluon jet energy difference (asymmetry) E_{diff} in MC. The efficiency is measured as the number of events with a g -tag higher than 0.8 ($\gamma > 0.8$) out of the total number. The efficiency is plotted for **signal events** according to MC Truth in blue and **background events** according to MC Truth in red.

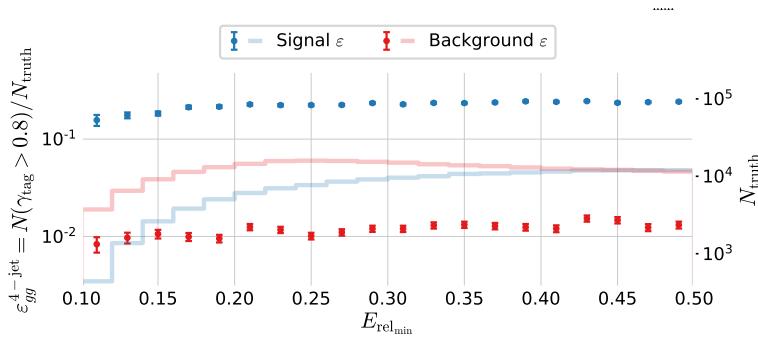
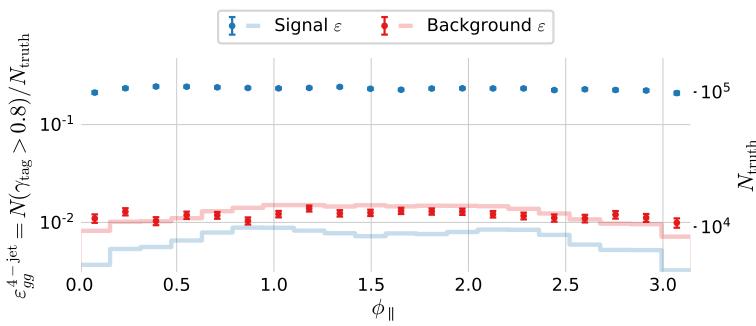
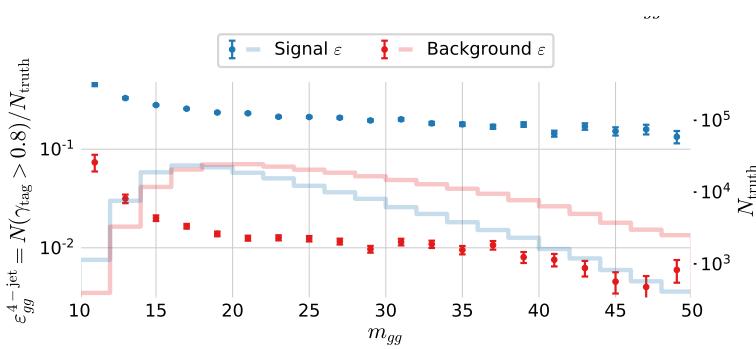
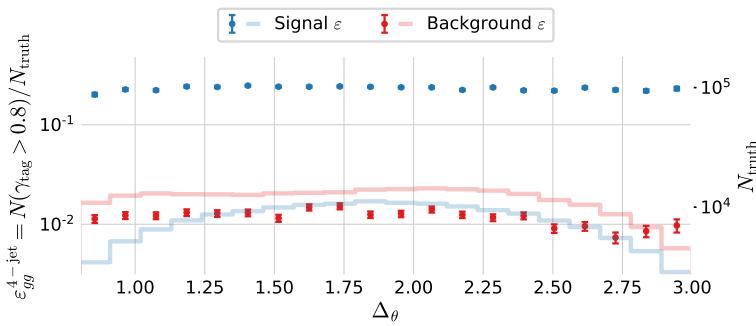
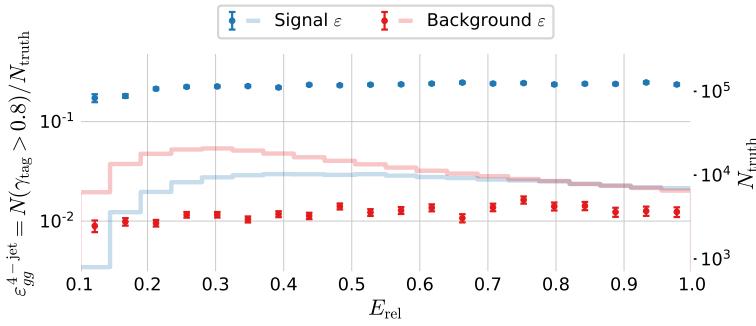


Figure B.46: Efficiency of the g -tagging algorithm for 4-jet events as a function of $E_{\text{rel,min}}$ in MC. The efficiency is measured as the number of events with a g -tag higher than 0.8 ($\gamma > 0.8$) out of the total number. The efficiency is plotted for **signal events** according to MC Truth in blue and **background events** according to MC Truth in red.



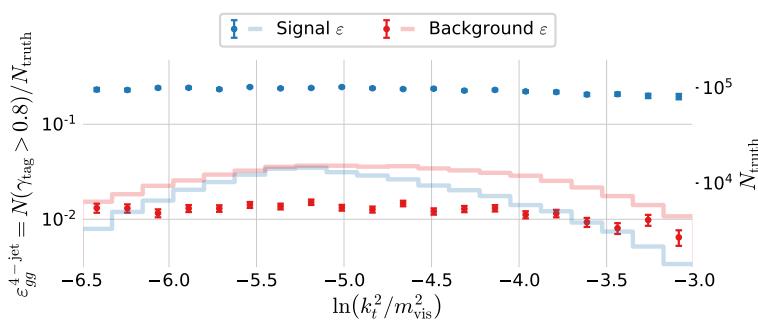


Figure B.51: Efficiency of the g -tagging algorithm for 4-jet events as a function of $\ln(k_t^2/m_{\text{vis}}^2)$ in MC. The efficiency is measured as the number of events with a g -tag higher than 0.8 ($\gamma > 0.8$) out of the total number. The efficiency is plotted for [signal events](#) according to MC Truth in blue and [background events](#) according to MC Truth in red.

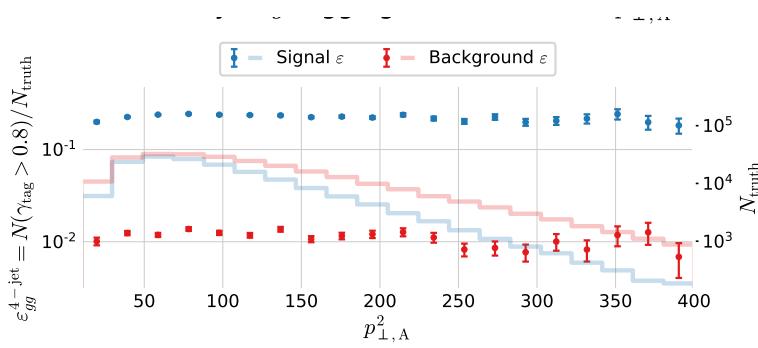


Figure B.52: Efficiency of the g -tagging algorithm for 4-jet events as a function of $p_{\perp,A}^2$ in MC. The efficiency is measured as the number of events with a g -tag higher than 0.8 ($\gamma > 0.8$) out of the total number. The efficiency is plotted for [signal events](#) according to MC Truth in blue and [background events](#) according to MC Truth in red.

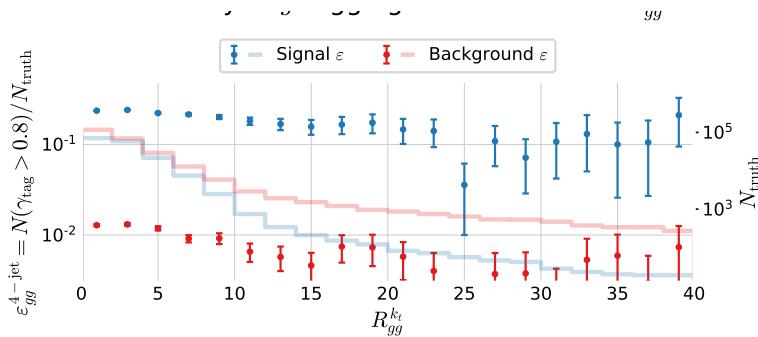


Figure B.53: Efficiency of the g -tagging algorithm for 4-jet events as a function of $R_{gg}^{k_t}$ in MC. The efficiency is measured as the number of events with a g -tag higher than 0.8 ($\gamma > 0.8$) out of the total number. The efficiency is plotted for [signal events](#) according to MC Truth in blue and [background events](#) according to MC Truth in red.

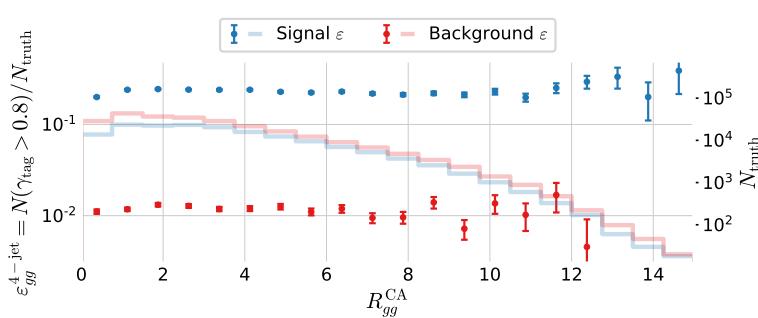
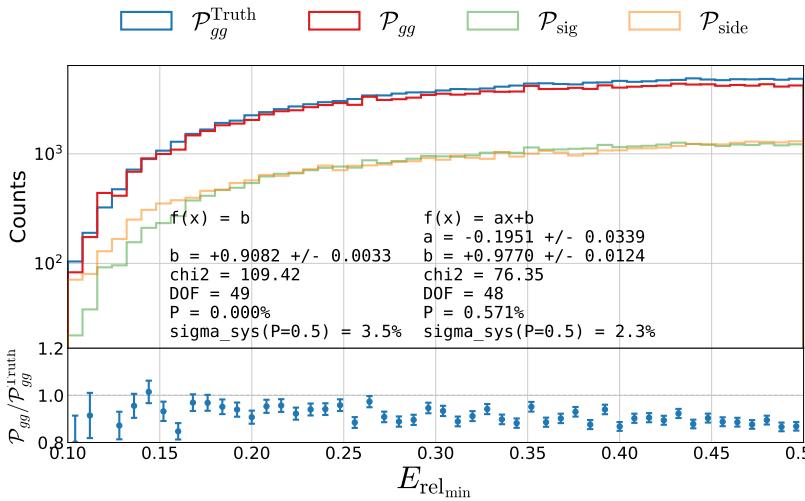
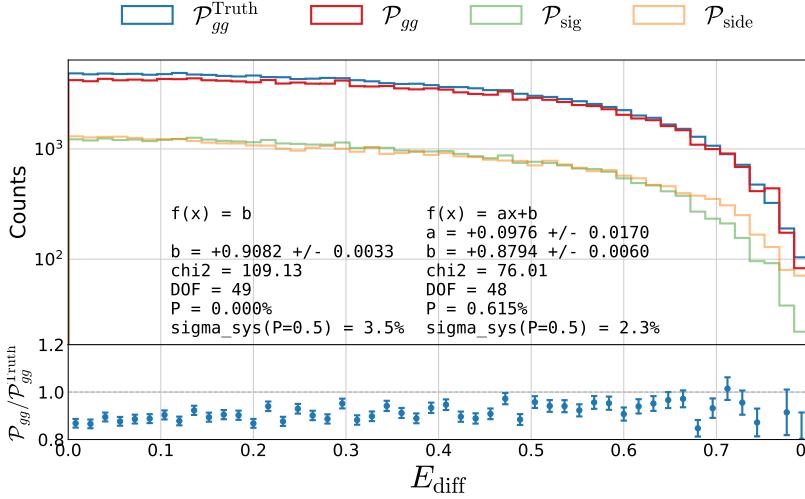


Figure B.54: Efficiency of the g -tagging algorithm for 4-jet events as a function of E in MC. The efficiency is measured as the number of events with a g -tag higher than 0.8 ($\gamma > 0.8$) out of the total number. The efficiency is plotted for [signal events](#) according to MC Truth in blue and [background events](#) according to MC Truth in red.



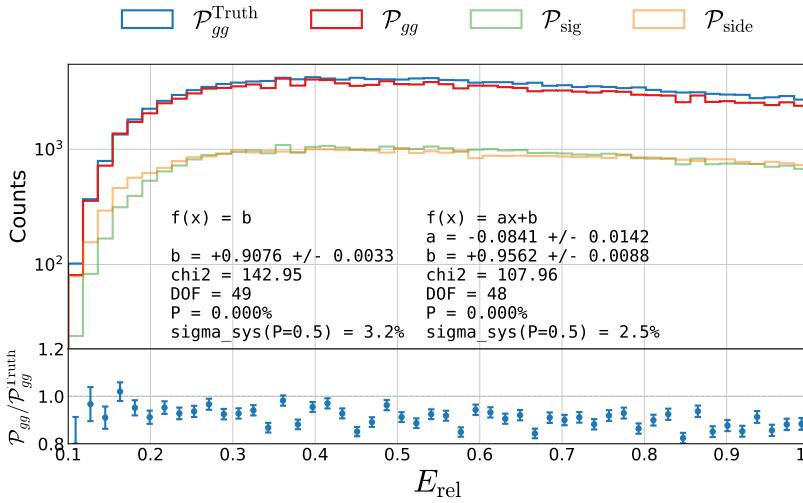


Figure B.57: Closure plot comparing MC Truth and the efficiency corrected g -tagging model in 4-jet events for E_{rel} . In the top part of the plot $\mathcal{P}_{gg}^{\text{Truth}}$ based on MC Truth is shown in blue, the \mathcal{P}_{gg} based on MC but without Truth in red, the distribution in the signal region \mathcal{P}_{sig} in light green and the distribution in the sideband region $\mathcal{P}_{\text{side}}$ in light orange. In the bottom part of the plot the ratio between \mathcal{P}_{gg} and $\mathcal{P}_{gg}^{\text{Truth}}$ is shown.

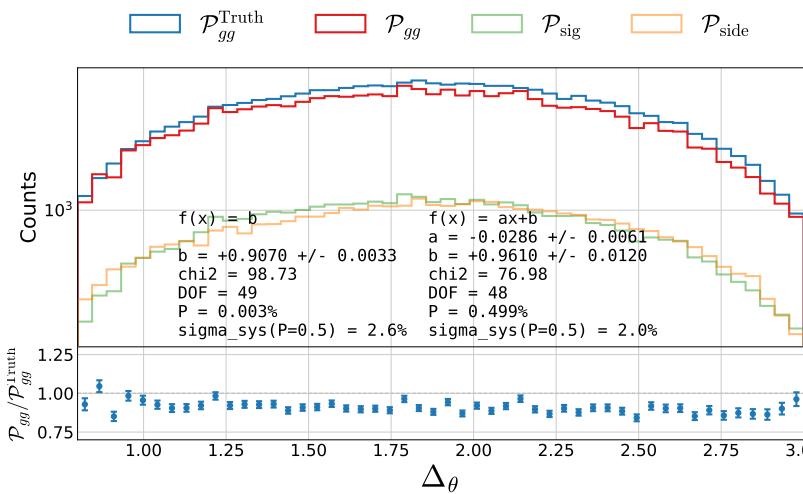


Figure B.58: Closure plot comparing MC Truth and the efficiency corrected g -tagging model in 4-jet events for Δ_θ . In the top part of the plot $\mathcal{P}_{gg}^{\text{Truth}}$ based on MC Truth is shown in blue, the \mathcal{P}_{gg} based on MC but without Truth in red, the distribution in the signal region \mathcal{P}_{sig} in light green and the distribution in the sideband region $\mathcal{P}_{\text{side}}$ in light orange. In the bottom part of the plot the ratio between \mathcal{P}_{gg} and $\mathcal{P}_{gg}^{\text{Truth}}$ is shown.

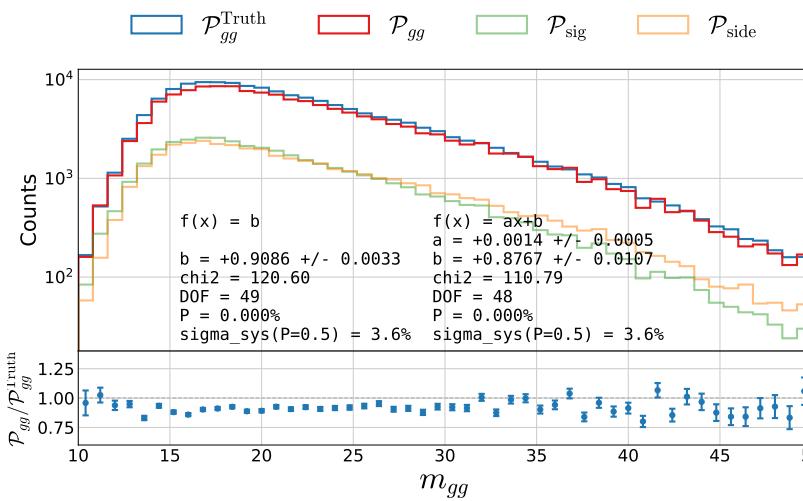


Figure B.59: Closure plot comparing MC Truth and the efficiency corrected g -tagging model in 4-jet events for m_{gg} . In the top part of the plot $\mathcal{P}_{gg}^{\text{Truth}}$ based on MC Truth is shown in blue, the \mathcal{P}_{gg} based on MC but without Truth in red, the distribution in the signal region \mathcal{P}_{sig} in light green and the distribution in the sideband region $\mathcal{P}_{\text{side}}$ in light orange. In the bottom part of the plot the ratio between \mathcal{P}_{gg} and $\mathcal{P}_{gg}^{\text{Truth}}$ is shown.

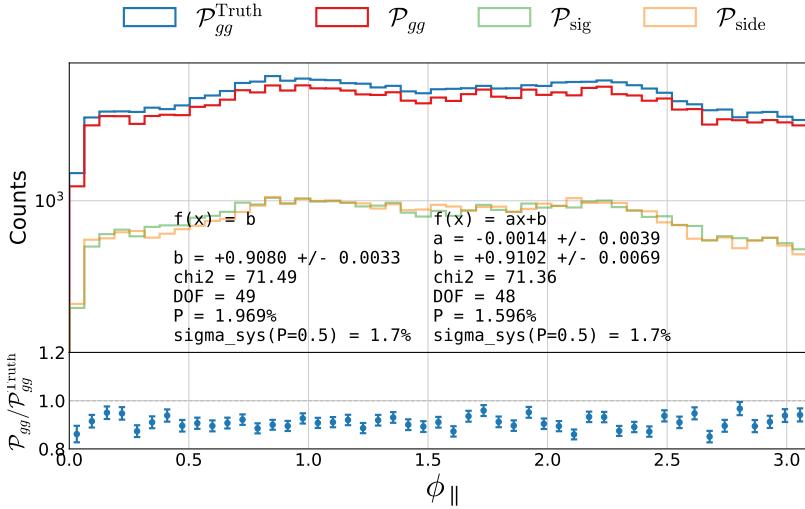


Figure B.60: Closure plot comparing MC Truth and the efficiency corrected g -tagging model in 4-jet events for ϕ_{\parallel} . In the top part of the plot $\mathcal{P}_{gg}^{\text{Truth}}$ based on MC Truth is shown in blue, the \mathcal{P}_{gg} based on MC but without Truth in red, the distribution in the signal region \mathcal{P}_{sig} in light green and the distribution in the sideband region $\mathcal{P}_{\text{side}}$ in light orange. In the bottom part of the plot the ratio between \mathcal{P}_{gg} and $\mathcal{P}_{gg}^{\text{Truth}}$ is shown.

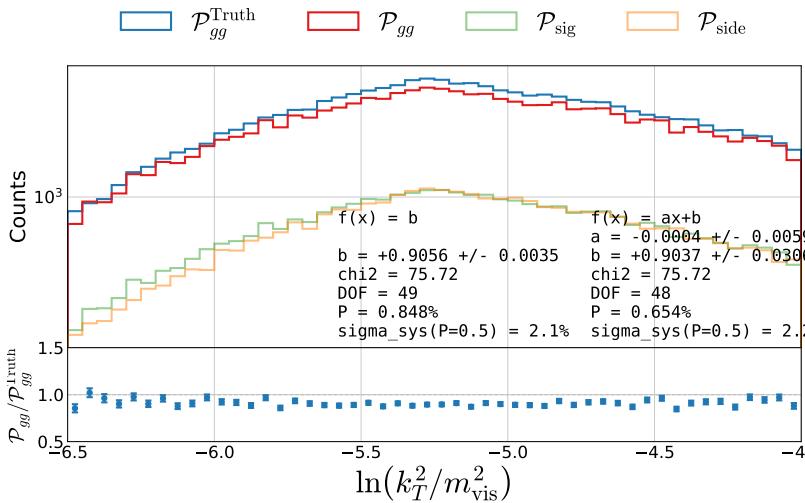


Figure B.61: Closure plot comparing MC Truth and the efficiency corrected g -tagging model in 4-jet events for $\ln(k_T^2/m_{\text{vis}}^2)$. In the top part of the plot $\mathcal{P}_{gg}^{\text{Truth}}$ based on MC Truth is shown in blue, the \mathcal{P}_{gg} based on MC but without Truth in red, the distribution in the signal region \mathcal{P}_{sig} in light green and the distribution in the sideband region $\mathcal{P}_{\text{side}}$ in light orange. In the bottom part of the plot the ratio between \mathcal{P}_{gg} and $\mathcal{P}_{gg}^{\text{Truth}}$ is shown.

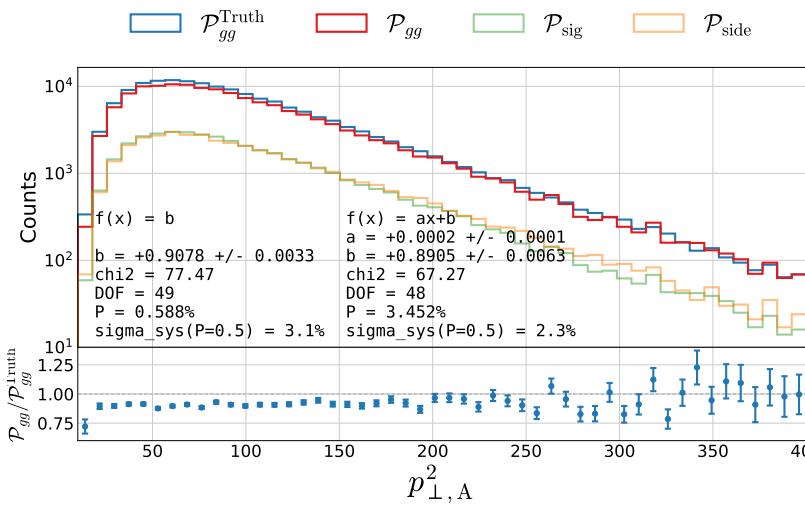


Figure B.62: Closure plot comparing MC Truth and the efficiency corrected g -tagging model in 4-jet events for $p_{\perp,A}^2$. In the top part of the plot $\mathcal{P}_{gg}^{\text{Truth}}$ based on MC Truth is shown in blue, the \mathcal{P}_{gg} based on MC but without Truth in red, the distribution in the signal region \mathcal{P}_{sig} in light green and the distribution in the sideband region $\mathcal{P}_{\text{side}}$ in light orange. In the bottom part of the plot the ratio between \mathcal{P}_{gg} and $\mathcal{P}_{gg}^{\text{Truth}}$ is shown.

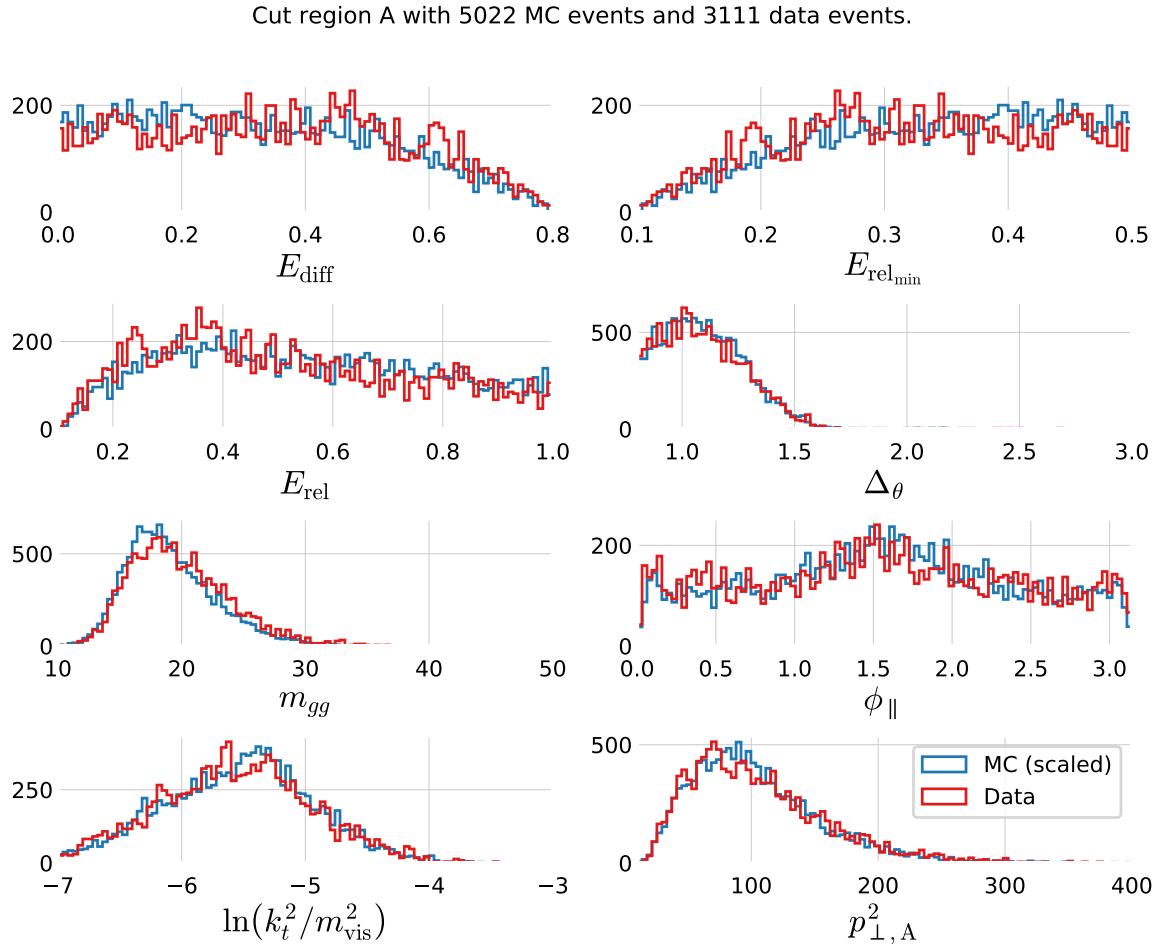


Figure B.63: Comparison of the gluon splitting distributions in MC and Data for $R_{gg}^{k_t}$ - R_{gg}^{CA} Phase Space Area A, see Table ???. The distribution for MC (scaled to Data) is shown in blue and for Data in red. These eight distributions are for the $R_{gg}^{k_t}$ - R_{gg}^{CA} Phase Space Area A which has 5022 events in the MC sample and 3111 in the Data sample.

Cut region B with 7382 MC events and 4035 data events.

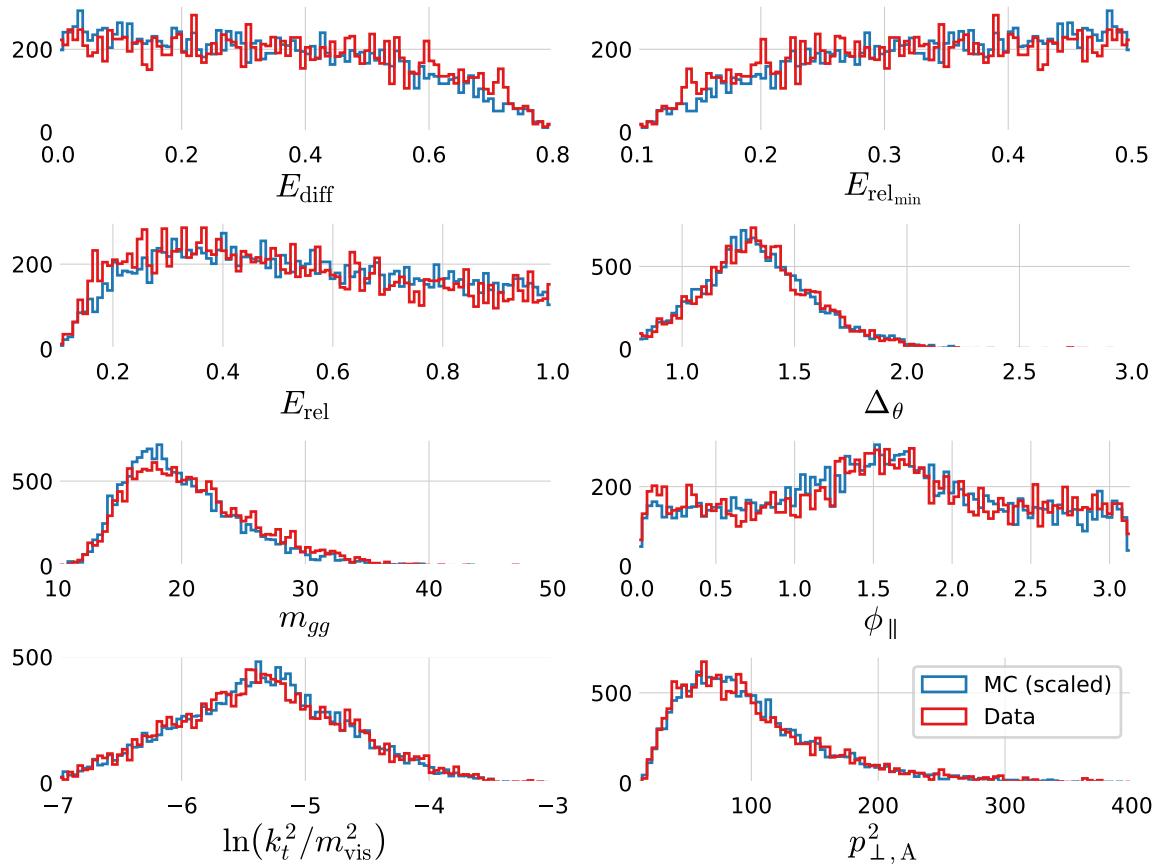


Figure B.64: Comparison of the gluon splitting distributions in MC and Data for $R_{gg}^k - R_{gg}^{\text{CA}}$ Phase Space Area B, see Table ???. The distribution for MC (scaled to Data) is shown in blue and for Data in red. These eight distributions are for the $R_{gg}^k - R_{gg}^{\text{CA}}$ Phase Space Area B which has 7382 events in the MC sample and 4035 in the Data sample.

Cut region C with 9417 MC events and 5344 data events.

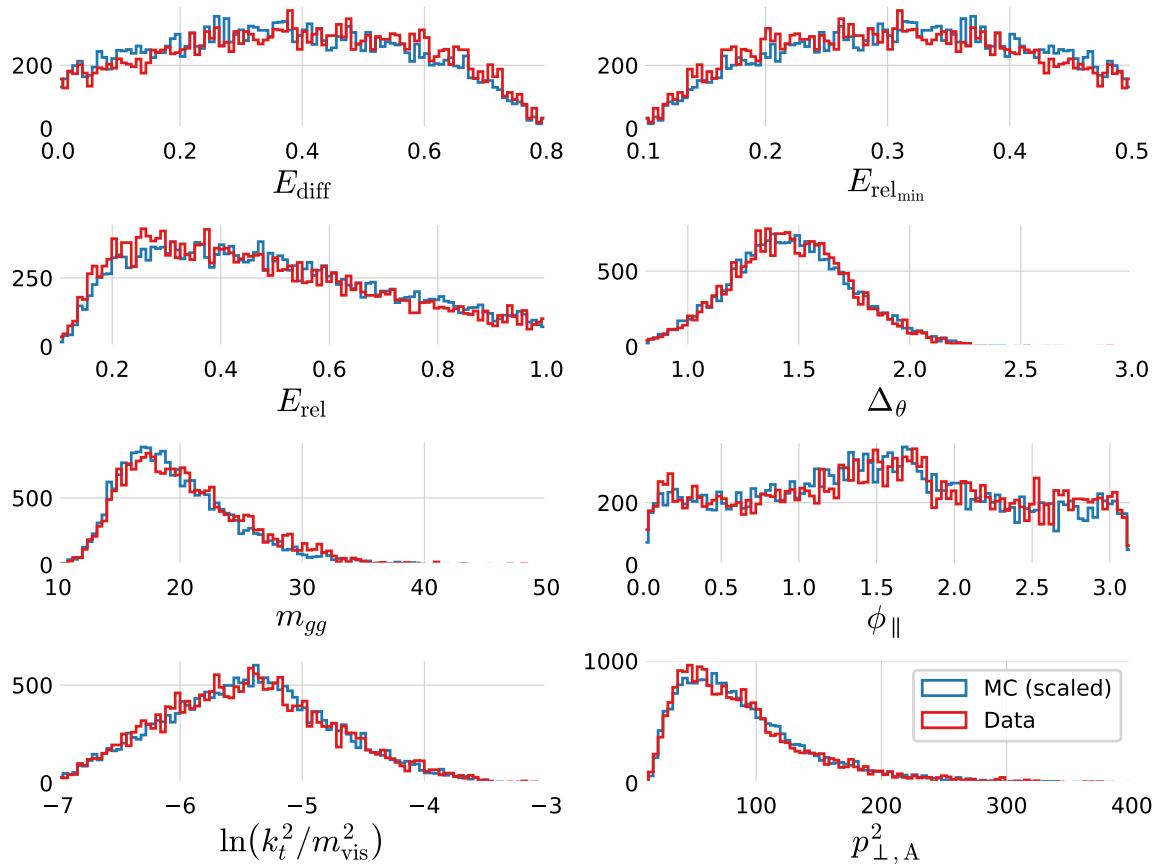


Figure B.65: Comparison of the gluon splitting distributions in MC and Data for $R_{gg}^{k_t}$ - R_{gg}^{CA} Phase Space Area C, see Table ???. The distribution for MC (scaled to Data) is shown in blue and for Data in red. These eight distributions are for the $R_{gg}^{k_t}$ - R_{gg}^{CA} Phase Space Area C which has 9417 events in the MC sample and 5344 in the Data sample.

Cut region D with 26366 MC events and 13780 data events.

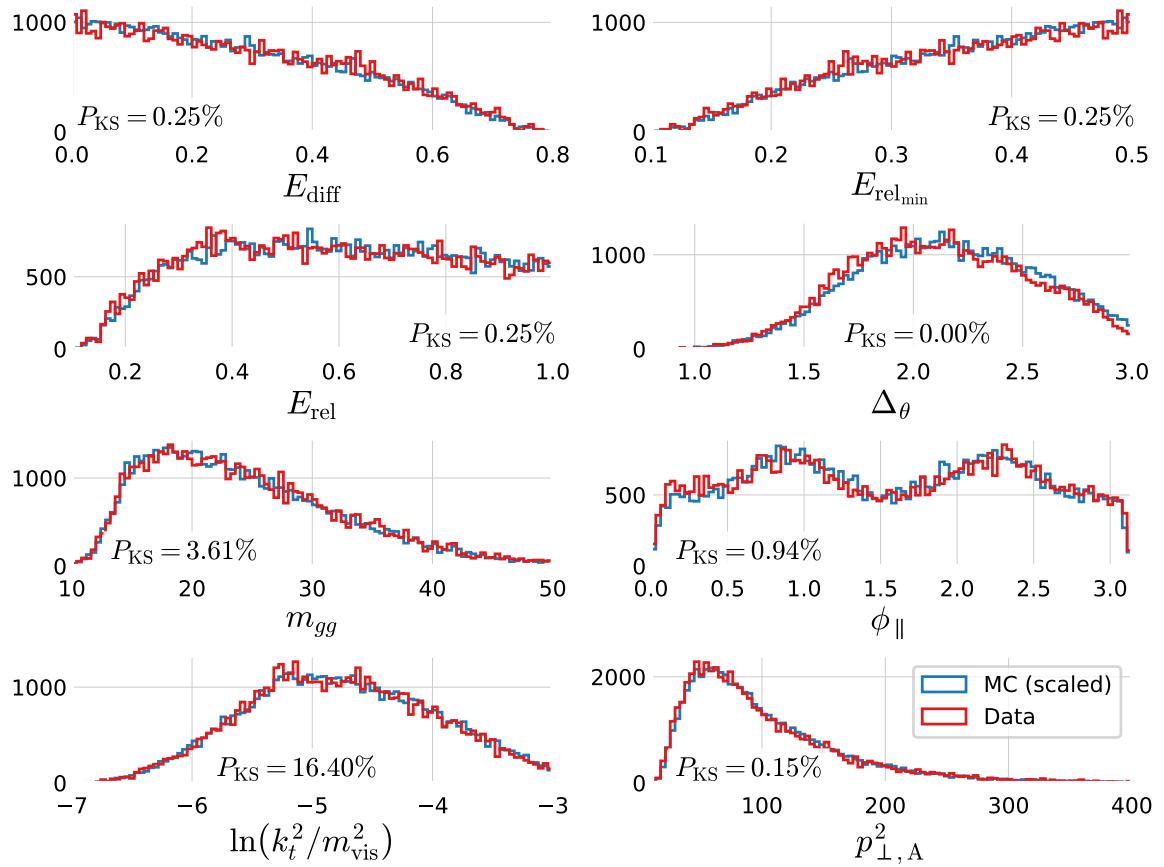


Figure B.66: Comparison of the gluon splitting distributions in MC and Data for $R_{gg}^{k_t}$ - R_{gg}^{CA} Phase Space Area D, see Table ???. The distribution for MC (scaled to Data) is shown in blue and for Data in red. These eight distributions are for the $R_{gg}^{k_t}$ - R_{gg}^{CA} Phase Space Area D which has 26366 events in the MC sample and 13780 in the Data sample.

List of Figures

2.1	The Learning Problem	6
2.2	Approximation-Estimation Tradeoff	10
2.3	Regularization Strength	11
2.4	Regularization Effect of L_2	12
2.5	Regularization Effect of L_1	12
2.6	k -Fold Cross Validation	13
2.7	k -Fold Cross Validation for Time Series Data	13
2.8	Objective Functions Zoom In	15
2.9	Objective Functions	15
2.10	Decision Tree Cuts In Feature Space	16
2.11	Decision Tree	16
2.12	Grid Search	20
2.13	Random Search	20
2.14	Bayesian Optimization	22
3.1	Danish Housing Price Index	27
3.2	Distributions of the Variables in the Housing Prices	28
3.3	Geographic Distribution of the Sold Residences	29
3.4	Histogram of Prices of Houses and Apartments Sold in Denmark	30
3.5	Linear Correlation Between Variables and Price	31
3.6	Comparison of the Linear Correlation ρ and the Nonlinear MIC	31
3.7	Nonlinear Correlation Between Variables and Price	32
3.8	Validity of Input Features	32
3.9	Validity Dendrogram	33
3.10	Prophet Forecast for Apartments	34
3.11	Prophet Trends	35
3.12	Sample Weight as a Function of Time for Different Half-Lives.	36
3.13	Parallel Coordinate Plot of the Initial Hyperparameter Optimization for Apartments	37
3.14	Initial HPO Results for the Weight Half-life $T_{\frac{1}{2}}$	38
3.15	Initial HPO Results for the Loss Function	38
3.16	Parallel Coordinate Plot of the Random Search Hyperparameter Optimization Results of XGBoost for Apartments	39
3.17	Hyperparameter Optimization: Random Search Results	40
3.18	Early Stopping Results	40
3.19	Performance of XGB-model for Apartments	41
3.20	Standard Deviation and MAD of the Static and Dynamic XGB Forecasts	41
3.21	Market Index based on the Static and Dynamic XGB Forecasts	42
3.22	SHAP Prediction Explanation for Apartments	44
3.23	Feature importance of Apartments Prices	44

3.24	Feature Importance Interaction Plot for Apartments	45
3.25	Performance Comparison of Multiple Models	46
3.26	Feature Importance of Villas With Descriptions	49
4.1	The Standard Model	54
4.2	Feynman Diagram for the Jet Production at LEP	55
4.3	Quark Splitting	55
4.4	Hadronization Process	56
4.5	The ALEPH Detector	57
4.6	Polar Angle	57
4.7	Azimuthal Angle	57
4.8	Track Significance	59
5.1	Histograms of the Vertex Variables	65
5.2	UMAP Visualization of the Vertex Variables for 4-Jet Events	66
5.3	UMAP Visualization of the Vertex Variables for 3-Jet Events	66
5.4	UMAP Visualization of the Vertex Variables for 2-Jet Events	66
5.5	Correlation of the Vertex Variables	67
5.6	Plot of the Log-Loss ℓ_{\log}	68
5.7	Hyperparameter Optimization of b -Tagging	69
5.8	Parallel Plot of HPO Results for 4-Jet b -Tagging	69
5.9	b -Tag Scores in 4-Jet Events	70
5.10	ROC Curve for 4-Jet b -Tagging	70
5.11	Distribution of b -Tags in 4-Jet Events	71
5.12	Global Feature Importances for the LGB b -Tagging Algorithm on 4-Jet Events	71
5.13	The Expit Function	71
5.14	The Logit Function	72
5.15	SHAP 3-Jet Model Explanation for b -Like Jet	72
5.16	b -Tagging Efficiency $\varepsilon_b^{b\text{-sig}}$ as a Function of Jet Energy	74
5.17	b -Tagging Efficiency $\varepsilon_g^{g\text{-sig}}$ as a Function of Jet Energy	74
5.18	Hyperparameter Optimization of g -Tagging	77
5.19	1D Sum Models Predictions and Signal Fraction for 4-jets events	78
5.20	g -Tag Scores in 4-Jet Events	79
5.21	ROC Curve for g -Tag in 4-Jet Events	80
5.22	Distribution of g -Tag Scores in 4-Jet Events for Signal and Background	80
5.23	Distribution of b -Tag Scores in 3-Jet l -Quark Events for Low and High g -Tag Values	80
5.24	3D Scatter Plot of β_{tag} -Values for High and Low γ_{tag} l -Quark Events	81
5.25	g -Tagging Pseudo Efficiency for $b\bar{b}g$ -Events as a Function of g -Tag	82
5.26	g -Tagging Pseudo Efficiency for $b\bar{b}g$ -Events as a Function of The Mean Invariant Mass	82
5.27	Generalized Angularities	83
5.28	Generalized Angularities for Charged Gluons Jets in 3-Jet Events: λ_0^2	84
5.29	Soft Wide Angle Gluons in 4-Jet Events	86
5.30	Soft Collinear Gluons in 4-Jet Events	87
5.31	Hard Non $g \rightarrow gg$ Gluons in 4-Jet Events	87
5.32	g -Tagging Efficiency for 4-Jet Events in MC as a Function of the Normalized Gluon-Gluon Jet Energy Difference Asymmetry E_{diff}	87
5.33	Closure Plot Comparing MC Truth and the Efficiency Corrected g -Tagging Model in 4-Jet Events for the Normalized Gluon Gluon Jet Energy Asymmetry	89
5.34	Overview of the Four Regions in the $R_{gg}^{k_l}$ - R_{gg}^{CA} Phase Space	90

5.35 Gluon Splitting Distribution Comparison in MC and Data for $R_{gg}^{k_t}$ - R_{gg}^{CA} Phase Space Region A	91
5.36 Ratio Plot of E_{diff} in MC and Data for $R_{gg}^{k_t}$ - R_{gg}^{CA} for Region A	92
5.37 Ratio Plot of E_{diff} in MC and Data for $R_{gg}^{k_t}$ - R_{gg}^{CA} for Region D	92
A.1 Distributions of the Input Variables – Part 1	96
A.2 Distributions of the Input Variables – Part 2	97
A.3 Distributions of the Input Variables – Part 3	98
A.4 Distributions of the Input Variables – Part 4	99
A.5 Distributions of the Input Variables – Part 5	100
A.6 Distributions of the Input Variables – Part 6	101
A.7 Distributions of the Input Variables – Part 7	102
A.8 Distributions of the Input Variables – Part 8	103
A.9 Distributions of the Input Variables – Part 9	104
A.10 Distributions of the Input Variables – Part 10	105
A.11 Distributions of the Input Variables – Part 11	106
A.12 Distributions of the Input Variables – Part 12	107
A.13 Distributions of the Input Variables – Part 13	108
A.14 Distributions of the Input Variables – Part 14	109
A.15 Linear Correlations	110
A.16 Comparison of the Linear Correlation ρ and the Nonlinear MIC for Different Levels of Noise	111
A.17 Validity Heatmap	111
A.18 Prophet Forecast for apartments	112
A.19 Prophet Trends	113
A.20 Parallel Coordinate Plot of the Initial Hyperparameter Optimization for Apartments	118
A.21 Parallel Coordinate Plot of the Initial Hyperparameter Optimization for Houses	118
A.22 Parallel Coordinate Plot of the Hyperparameter Optimization for Apartments Using Random Search	118
A.23 Parallel Coordinate Plot of the Hyperparameter Optimization for Apartments Using Bayesian Optimization	119
A.24 Parallel Coordinate Plot of the Hyperparameter Optimization for Houses Using Random Search	119
A.25 Parallel Coordinate Plot of the Hyperparameter Optimization for Houses Using Bayesian Optimization	119
A.26 Random Search Results as a Function of Iteration for Apartments	120
A.27 Bayesian Optimization Results as a Function of Iteration for Apartments	120
A.28 Random Search Results as a Function of Iteration for Houses	120
A.29 Bayesian Optimization Results as a Function of Iteration for Houses	120
A.30 Performance of XGB-model on apartment prices	121
A.31 Performance Comparison of Multiple Models for Houses	122
B.1 UMAP Parameter Grid Search	124
B.2 Visualization of the t-SNE algorithm	124
B.3 Parallel Coordinate Plot of HPO results for 3-Jet b -Tagging	125
B.4 Parallel Coordinate Plot of HPO results for 4-Jet b -Tagging	125
B.5 b -Tag Scores in 3-Jet Events	126
B.6 b -Tag Scores in 4-Jet Events	126
B.7 ROC curve for 3-jet b -tagging	127
B.8 ROC curve for 4-jet b -tagging	127
B.9 Distribution of b -Tags in 3-Jet Events	128

B.10	Distribution of b -Tags in 4-Jet Events	128
B.11	Global Feature Importances for the LGB b -Tagging Algorithm on 3-Jet Events	129
B.12	Global Feature Importances for the LGB b -Tagging Algorithm on 4-Jet Events	129
B.13	Parallel Coordinate Plot of HPO Results for 3-Jet g -Tagging for Energy Ordered Jets	130
B.14	Parallel Coordinate Plot of HPO Results for 3-Jet g -Tagging for Shuffled Jets	130
B.15	Parallel Coordinate Plot of HPO Results for 4-Jet g -Tagging for Energy Ordered Jets	131
B.16	Parallel Coordinate Plot of HPO Results for 4-Jet g -Tagging for Shuffled Jets	131
B.17	PermNet Architecture	132
B.18	1D LGB Model Cuts for 4-jets events	133
B.19	1D Sum Models Predictions and Signal Fraction for 3-jets events	133
B.20	1D LGB Model Cuts for 3-jets events	133
B.21	g -Tag Scores in 3-Jet Events	134
B.22	ROC curve for g -tag in 4-jet events	134
B.23	ROC Curve for g -Tag in 3-Jet Events	134
B.24	Distribution of g -Tag Scores in 3-Jet Events for Signal and Background	135
B.25	b -Tagging Efficiency $\varepsilon_b^{g\text{-sig}}$ as a Function of Jet Energy	135
B.26	b -Tagging Efficiency $\varepsilon_g^{b\text{-sig}}$ as a Function of Jet Energy	135
B.27	Generalized Angularities for Charged Gluons Jets: λ_0^2	136
B.28	Generalized Angularities for Charged Gluons Jets: $\lambda_{\frac{1}{2}}^1$	136
B.29	Generalized Angularities for Charged Gluons Jets: λ_1^1	136
B.30	Generalized Angularities for Charged Gluons Jets: λ_1^2	137
B.31	Generalized Angularities for Charged Gluons Jets: λ_0^0	137
B.32	Generalized Angularities for Neutral Gluons Jets: λ_0^2	137
B.33	Generalized Angularities for Neutral Gluons Jets: $\lambda_{\frac{1}{2}}^1$	138
B.34	Generalized Angularities for Neutral Gluons Jets: λ_1^1	138
B.35	Generalized Angularities for Neutral Gluons Jets: λ_1^2	138
B.36	Generalized Angularities for Neutral Gluons Jets: λ_0^0	139
B.37	Relationship Between the g -Tag Value γ_{tag} and the Gluon Splitting Variable E_{diff}	140
B.38	Relationship Between the g -Tag Value γ_{tag} and the Gluon Splitting Variable $E_{\text{rel,min}}$	140
B.39	Relationship Between the g -Tag Value γ_{tag} and the Gluon Splitting Variable E_{rel}	140
B.40	Relationship Between the g -Tag Value γ_{tag} and the Gluon Splitting Variable Δ_θ	141
B.41	Relationship Between the g -Tag Value γ_{tag} and the Gluon Splitting Variable m_{gg}	141
B.42	Relationship Between the g -Tag Value γ_{tag} and the Gluon Splitting Variable ϕ_{\parallel}	141
B.43	Relationship Between the g -Tag Value γ_{tag} and the Gluon Splitting Variable $\ln(k_t^2/m_{\text{vis}}^2)$	141
B.44	Relationship Between the g -Tag Value γ_{tag} and the Gluon Splitting Variable $p_{\perp,A}^2$	142
B.45	g -Tagging Efficiency for 4-Jet Events in MC as a Function of E_{diff}	142
B.46	g -Tagging Efficiency for 4-Jet Events in MC as a Function of $E_{\text{rel,min}}$	142
B.47	g -Tagging Efficiency for 4-Jet Events in MC as a Function of E_{rel}	143
B.48	g -Tagging Efficiency for 4-Jet Events in MC as a Function of Δ_θ	143
B.49	g -Tagging Efficiency for 4-Jet Events in MC as a Function of m_{gg}	143
B.50	g -Tagging Efficiency for 4-Jet Events in MC as a Function of ϕ_{\parallel}	143
B.51	g -Tagging Efficiency for 4-Jet Events in MC as a Function of $\ln(k_t^2/m_{\text{vis}}^2)$	144
B.52	g -Tagging Efficiency for 4-Jet Events in MC as a Function of $p_{\perp,A}^2$	144
B.53	g -Tagging Efficiency for 4-Jet Events in MC as a Function of R_{gg}^{kt}	144
B.54	g -Tagging Efficiency for 4-Jet Events in MC as a Function of R_{gg}^{CA}	144
B.55	Closure Plot Comparing MC Truth and the Efficiency Corrected g -Tagging Model in 4-Jet Events for E_{diff}	145

B.56 Closure Plot Comparing MC Truth and the Efficiency Corrected g -Tagging Model in 4-Jet Events for E_{relmin}	145
B.57 Closure Plot Comparing MC Truth and the Efficiency Corrected g -Tagging Model in 4-Jet Events for E_{rel}	146
B.58 Closure Plot Comparing MC Truth and the Efficiency Corrected g -Tagging Model in 4-Jet Events for Δ_θ	146
B.59 Closure Plot Comparing MC Truth and the Efficiency Corrected g -Tagging Model in 4-Jet Events for m_{gg}	146
B.60 Closure Plot Comparing MC Truth and the Efficiency Corrected g -Tagging Model in 4-Jet Events for ϕ_{\parallel}	147
B.61 Closure Plot Comparing MC Truth and the Efficiency Corrected g -Tagging Model in 4-Jet Events for $\ln(k_t^2/m_{\text{vis}}^2)$	147
B.62 Closure Plot Comparing MC Truth and the Efficiency Corrected g -Tagging Model in 4-Jet Events for $p_{\perp,A}^2$	147
B.63 Gluon Splitting Distribution Comparison in MC and Data for $R_{gg}^{k_t}$ - R_{gg}^{CA} Phase Space Area A	148
B.64 Gluon Splitting Distribution Comparison in MC and Data for $R_{gg}^{k_t}$ - R_{gg}^{CA} Phase Space Area B	149
B.65 Gluon Splitting Distribution Comparison in MC and Data for $R_{gg}^{k_t}$ - R_{gg}^{CA} Phase Space Area C	150
B.66 Gluon Splitting Distribution Comparison in MC and Data for $R_{gg}^{k_t}$ - R_{gg}^{CA} Phase Space Area D	151

List of Tables

3.1	Mapping between the Code in <code>SagTypeNr</code> and the Type of Residence	29
3.2	Basic Cuts	33
3.3	Side Door Mapping.	33
3.4	Street Mapping	33
3.5	Number of Observations in the Housing Dataset	36
3.6	Number of Observations in the Housing Dataset for the Tight Selection	36
3.7	Results from the Initial Hyperparameter Optimization for Apartments	37
3.8	Results from the Initial Hyperparameter Optimization for Houses	37
3.9	PDFs Used in the Random Search	39
3.10	Realtors' MAD	41
3.11	Performance Metrics for Apartments	43
3.12	Performance Metrics for Houses	43
5.1	Dimensions of the Dataset for Data	64
5.2	Dimensions of the Dataset for MC and MC _b	64
5.3	Number of Different Types of Jets for MC and MC _b for n -Jet Events	65
5.4	Random Search PDFs for LGB	69
5.5	Global SHAP Feature Importances for the g -Tagging Models in 4-Jet Events	77
5.6	Generalized Angularities Systematic Errors, Charged Tracks	84
5.7	Generalized Angularities Systematic Errors, Neutral Clusters	84
5.8	Gluon Splitting Systematic Errors	89
5.9	Rgion Definition in the $R_{gg}^{k_t}$ - R_{gg}^{CA} Phase Space	90
A.1	Energy Rating Mapping	112
A.2	Final Variable List	114
A.3	Initial Hyperparameter Optimization Results for Apartments – SE Loss Function	115
A.4	Initial Hyperparameter Optimization Results for Apartments – Log-Cosh Loss Function	115
A.5	Initial Hyperparameter Optimization Results for Apartments – Cauchy Loss Function	115
A.6	Initial Hyperparameter Optimization Results for Apartments – Welsch Loss Function	116
A.7	Initial Hyperparameter Optimization Results for Apartments – Fair Loss Function	116
A.8	Initial Hyperparameter Optimization Results for Houses – SE Loss Function	116
A.9	Initial Hyperparameter Optimization Results for Houses – Log-Cosh Loss Function	117
A.10	Initial Hyperparameter Optimization Results for Houses – Cauchy Loss Function	117
A.11	Initial Hyperparameter Optimization Results for Houses – Welsch Loss Function	117
A.12	Initial Hyperparameter Optimization Results for Houses – Fair Loss Function	117
A.13	Performance Metrics for Apartments with the Tight Selection	121
A.14	Performance Metrics for Houses with the Tight Selection	121

B.1	Number of different types of jets for MC and MC _b written in relative numbers such that each row sum to 100 %. See also Table 5.3.	123
B.2	Random Search PDFs for XGB	123
B.3	Global SHAP Feature Importances for the <i>g</i> -Tagging Models in 3-Jet Events	134

Bibliography

- [1] Advanced Topics in Machine Learning (ATML). URL <https://kurser.ku.dk/course/ndak15014u>.
- [2] Allstate Claims Severity - Fair Loss. URL <https://kaggle.com/c/allstate-claims-severity>.
- [3] Dmlc/xgboost. URL <https://github.com/dmlc/xgboost>.
- [4] HEP meets ML award | The Higgs Machine Learning Challenge. URL <https://higgsml.lal.in2p3.fr/prizes-and-award/award/>.
- [5] The Large Electron-Positron Collider | CERN. URL <https://home.cern/science/accelerators/large-electron-positron-collider>.
- [6] Microsoft/LightGBM. URL https://github.com/microsoft/LightGBM/blob/b397555d7023fd05f8e56326905fe7b185109de/src/treelearner/serial_tree_learner.cpp#L282.
- [7] Scikit-hep/uproot. URL <https://github.com/scikit-hep/uproot>.
- [8] Datashader: Revealing the Structure of Genuinely Big Data. URL <https://github.com/holoviz/datashader>.
- [9] O . Wwww.OIS.dk - Din genvej til ejendomsdata. URL <https://www.ois.dk/>.
- [10] M. Abadi et al. TensorFlow: Large-scale machine learning on heterogeneous systems. URL <http://tensorflow.org/>.
- [11] Y. S. Abu-Mostafa, M. Magdon-Ismail, and H.-T. Lin. *Learning From Data*. AMLBook. ISBN 978-1-60049-006-4.
- [12] D. Albanese, S. Riccadonna, C. Donati, and P. Franceschi. A practical tool for maximal information coefficient analysis. 7. ISSN 2047-217X. doi: 10.1093/gigascience/giy032. URL <https://doi.org/10.1093/gigascience/giy032>.
- [13] H. Albrecht, H. Ehrlichmann, T. Hamacher, R. P. Hofmann, T. Kirchhoff, A. Nau, S. Nowak, H. Schröder, H. D. Schulz, M. Walter, R. Wurth, C. Hast, H. Kolanoski, A. Kosche,

- A. Lange, A. Lindner, R. Mankel, M. Schieber, T. Siegmund, B. Spaan, H. Thurn, D. Töpfer, D. Wegener, M. Bittrner, P. Eckstein, M. Paulini, K. Reim, H. Wegener, R. Eckmann, R. Mundt, T. Oest, R. Reiner, W. Schmidt-Parzefall, W. Funk, J. Stiewe, S. Werner, K. Ehret, W. Hofmann, A. Hüpper, S. Khan, K. T. Knöpfle, M. Seeger, J. Spengler, D. I. Britton, C. E. K. Charlesworth, K. W. Edwards, E. R. F. Hyatt, H. Kapitza, P. Krieger, D. B. MacFarlane, P. M. Patel, J. D. Prentice, P. R. B. Saull, K. Tzamariudaki, R. G. Van de Water, T. S. Yoon, D. Reßing, M. Schmidtler, M. Schneider, K. R. Schubert, K. Strahl, R. Waldi, S. Weseler, G. Kernel, P. Križnič, T. Podobnik, T. Živko, V. Balagura, I. Belyaev, S. Chechelnitsky, M. Danilov, A. Droutskoy, Y. Gershtein, A. Golutvin, G. Kostina, D. Litvintsev, V. Lubimov, P. Pakhlov, F. Ratnikov, S. Semenov, A. Snizhko, V. Soloshenko, I. Tichomirov, and Y. Zaitsev. A model-independent determination of the inclusive semileptonic decay fraction of B mesons. 318(2): 397–404. ISSN 0370-2693. doi: 10.1016/0370-2693(93)90146-9. URL <http://www.sciencedirect.com/science/article/pii/0370269393901469>.
- [14] E. Anderson. The Species Problem in Iris. 23(3):457–509. ISSN 00266493. doi: 10.2307/2394164. URL www.jstor.org/stable/2394164.
 - [15] B. Andersson, G. Gustafson, G. Ingelman, and T. Sjöstrand. Parton fragmentation and string dynamics. 97(2): 31–145. ISSN 0370-1573. doi: 10.1016/0370-1573(83)90080-7. URL <http://www.sciencedirect.com/science/article/pii/0370157383900807>.
 - [16] S. R. Armstrong. A Search for the standard model Higgs boson in four jet final states at center-of-mass energies near 183-GeV with the ALEPH detector at LEP. URL <http://wwwlib.umi.com/dissertations/fullcit?p9910371>.
 - [17] M. Awad and R. Khanna. Support Vector Regression. In M. Awad and R. Khanna, editors, *Efficient Learning Machines: Theories, Concepts, and Applications for Engineers and System Designers*, pages 67–80. Apress. ISBN 978-1-4302-5990-9. doi: 10.1007/978-1-4302-5990-9_4. URL https://doi.org/10.1007/978-1-4302-5990-9_4.
 - [18] R. J. Barlow. *Statistics: A Guide to the Use of Statistical Methods in the Physical Sciences (Manchester Physics Series)*. WileyBlackwell, reprint edition. ISBN 0-471-92295-1. URL <http://www.amazon.co.uk/Statistics-Statistical-Physical-Sciences-Manchester/dp/0471922951%3FSubscriptionId%3D13CT5CVB80YFWJEPWS02%26tag%3Dws%26linkCode%3Dxm%26camp%3D2025%26creative%3D165953%26creativeASIN%3D0471922951>.

- [19] J. T. Barron. A General and Adaptive Robust Loss Function. URL <http://arxiv.org/abs/1701.03077>.
- [20] W. Bartel et al. Experimental study of jets in electron-positron annihilation. 101(1):129–134. ISSN 0370-2693. doi: 10.1016/0370-2693(81)90505-0. URL <http://www.sciencedirect.com/science/article/pii/0370269381905050>.
- [21] E. Becht, C.-A. Dutertre, I. W. H. Kwok, L. G. Ng, F. Ginhoux, and E. W. Newell. Evaluation of UMAP as an alternative to t-SNE for single-cell data. page 298430, . doi: 10.1101/298430. URL <https://www.biorxiv.org/content/10.1101/298430v1>.
- [22] E. Becht, L. McInnes, J. Healy, C.-A. Dutertre, I. W. H. Kwok, L. G. Ng, F. Ginhoux, and E. W. Newell. Dimensionality reduction for visualizing single-cell data using UMAP. 37(1):38–44, . ISSN 1546-1696. doi: 10.1038/nbt.4314. URL <https://www.nature.com/articles/nbt.4314>.
- [23] J. Bergstra and Y. Bengio. Random Search for Hyper-parameter Optimization. 13:281–305. ISSN 1532-4435. URL <http://dl.acm.org/citation.cfm?id=2188385.2188395>.
- [24] C. Bierlich. Rope hadronization, geometry and particle production in pp and pA Collisions. URL <https://lup.lub.lu.se/search/ws/files/18474576/thesis.pdf>.
- [25] Bolighed. Bolighed - usikkerhed i data-vurderingen. URL <https://bolighed.dk/om-bolighed/spoergsmaal-og-svar/#boligvaerdi>.
- [26] L. Breiman. Random Forests. 45(1):5–32. ISSN 1573-0565. doi: 10.1023/A:1010933404324. URL <https://doi.org/10.1023/A:1010933404324>.
- [27] R. P. Brent. *Algorithms for Minimization without Derivatives*. Prentice-Hall Series in Automatic Computation. Prentice-Hall. URL <https://cds.cern.ch/record/113464>.
- [28] E. Brochu, V. M. Cora, and N. de Freitas. A Tutorial on Bayesian Optimization of Expensive Cost Functions, with Application to Active User Modeling and Hierarchical Reinforcement Learning. URL <http://arxiv.org/abs/1012.2599>.
- [29] R. Brun and F. Rademakers. ROOT — An object oriented data analysis framework. 389(1):81–86. ISSN 0168-9002. doi: 10.1016/S0168-9002(97)00048-X. URL <http://www.sciencedirect.com/science/article/pii/S016890029700048X>.
- [30] A. Buckley, J. Butterworth, S. Gieseke, D. Grellscheid, S. Hoche, H. Hoeth, F. Krauss, L. Lonnblad, E. Nurse, P. Richardson, S. Schumann, M. H. Seymour, T. Sjostrand,

- P. Skands, and B. Webber. General-purpose event generators for LHC physics. 504(5):145–233. ISSN 03701573. doi: 10.1016/j.physrep.2011.03.005. URL <http://arxiv.org/abs/1101.2599>.
- [31] C. Burgard. Standard model of physics | TikZ example. URL <http://www.texample.net/tikz/examples/model-physics/>.
 - [32] D. Buskulic et al. An investigation of B_{d0} and B_{s0} oscillation. 322(4):441–458. ISSN 0370-2693. doi: 10.1016/0370-2693(94)91177-0. URL <http://www.sciencedirect.com/science/article/pii/0370269394911770>.
 - [33] M. Cacciari, G. P. Salam, and G. Soyez. FastJet user manual. 72(3):1896. ISSN 1434-6044, 1434-6052. doi: 10.1140/epjc/s10052-012-1896-2. URL <http://arxiv.org/abs/1111.6097>.
 - [34] S. Catani, Y. L. Dokshitzer, M. H. Seymour, and B. R. Webber. Longitudinally-invariant kt-clustering algorithms for hadron-hadron collisions. 406(1):187–224, . ISSN 0550-3213. doi: 10.1016/0550-3213(93)90166-M. URL <http://www.sciencedirect.com/science/article/pii/055032139390166M>.
 - [35] S. Catani, G. Turnock, and B. R. Webber. Jet broadening measures in e+ e- annihilation. B295:269–276, . doi: 10.1016/0370-2693(92)91565-Q.
 - [36] T. Chen and C. Guestrin. XGBoost: A Scalable Tree Boosting System. pages 785–794. doi: 10.1145/2939672.2939785. URL <http://arxiv.org/abs/1603.02754>.
 - [37] A. Collaboration. Electron efficiency measurements with the ATLAS detector using 2012 LHC proton-proton collision data. 77(3):195, . ISSN 1434-6044, 1434-6052. doi: 10.1140/epjc/s10052-017-4756-2. URL <http://arxiv.org/abs/1612.01456>.
 - [38] C. Collaboration. Search for a Higgs boson in the decay channel H to ZZ(*) to q qbar l-l+ in pp collisions at sqrt(s) = 7 TeV. 2012(4):36, . ISSN 1029-8479. doi: 10.1007/JHEP04(2012)036. URL <http://arxiv.org/abs/1202.1416>.
 - [39] T. A. Collaboration. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. 716(1):1–29, . ISSN 03702693. doi: 10.1016/j.physletb.2012.08.020. URL <http://arxiv.org/abs/1207.7214>.
 - [40] T. C. Collaboration. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. 716(1):30–61, . ISSN 03702693. doi: 10.1016/j.physletb.2012.08.021. URL <http://arxiv.org/abs/1207.7235>.
 - [41] M. Dam. An upper limit for Br(Z->ggg) from symmetric 3-jet zo hadronic decays. 389(2):405–415. ISSN 0370-2693. doi: 10.1016/S0370-2693(96)01450-5.

- [42] A. Diaz-Papkovich, L. Anderson-Trocmé, C. Ben-Eghan, and S. Gravel. UMAP reveals cryptic population structure and phenotype heterogeneity in large genomic cohorts. 15(11). ISSN 1553-7390. doi: 10.1371/journal.pgen.1008432. URL <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6853336/>.
- [43] Y. L. Dokshitzer, G. D. Leder, S. Moretti, and B. R. Webber. Better Jet Clustering Algorithms. 1997(08):001–001. ISSN 1029-8479. doi: 10.1088/1126-6708/1997/08/001. URL <http://arxiv.org/abs/hep-ph/9707323>.
- [44] S. D. DST. Price Index (EJ14) - Statistics Denmark. URL <https://www.dst.dk/en/Statistik/emner/priser-og-forbrug/ejendomme>.
- [45] S. D. Ellis and D. E. Soper. Successive combination jet algorithm for hadron collisions. 48(7):3160–3166. doi: 10.1103/PhysRevD.48.3160. URL <https://link.aps.org/doi/10.1103/PhysRevD.48.3160>.
- [46] D. et al. Buskulic. A precise measurement of hadrons. 313(3):535–548. ISSN 0370-2693. doi: 10.1016/0370-2693(93)90028-G. URL <http://www.sciencedirect.com/science/article/pii/037026939390028G>.
- [47] F. Faye. Frederik Faye / deepcalo. URL <https://gitlab.com/ffaye/deepcalo>.
- [48] R. A. Fisher. The Use of Multiple Measurements in Taxonomic Problems. 7(2):179–188. ISSN 2050-1439. doi: 10.1111/j.1469-1809.1936.tb02137.x. URL <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1469-1809.1936.tb02137.x>.
- [49] Y. Freund and R. E. Schapire. A desicion-theoretic generalization of on-line learning and an application to boosting. In P. Vitányi, editor, *Computational Learning Theory*, pages 23–37. Springer Berlin Heidelberg. ISBN 978-3-540-49195-8. Adaboost.
- [50] S. L. Glashow. Partial-symmetries of weak interactions. 22(4):579–588. ISSN 0029-5582. doi: 10.1016/0029-5582(61)90469-2. URL <http://www.sciencedirect.com/science/article/pii/0029558261904692>.
- [51] P. Gras, S. Höche, D. Kar, A. Larkoski, L. Lönnblad, S. Plätzer, A. Siódmok, P. Skands, G. Soyez, and J. Thaler. Systematics of quark/gluon tagging. 2017(7):91. ISSN 1029-8479. doi: 10.1007/JHEP07(2017)091. URL <http://arxiv.org/abs/1704.03878>.
- [52] N. Guttenberg, N. Virgo, O. Witkowski, H. Aoki, and R. Kanai. Permutation-equivariant neural networks applied to dynamics prediction. URL <http://arxiv.org/abs/1612.04530>.

- [53] A. E. Harvey and S. Peters. Estimation Procedures for Structural Time Series Models. doi: [10.1002/for.3980090203](https://doi.org/10.1002/for.3980090203).
- [54] T. Hastie and R. Tibshirani. Generalized Additive Models: Some Applications. 82(398):371–386. ISSN 01621459. doi: [10.2307/2289439](https://doi.org/10.2307/2289439). URL www.jstor.org/stable/2289439.
- [55] T. Hastie, R. Tibshirani, and J. Friedman. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Second Edition*. Springer Series in Statistics. Springer-Verlag, 2 edition. ISBN 978-0-387-84857-0. URL [//www.springer.com/la/book/9780387848570](http://www.springer.com/la/book/9780387848570).
- [56] K. Hornik. Approximation capabilities of multilayer feedforward networks. 4(2):251–257. ISSN 0893-6080. doi: [10.1016/0893-6080\(91\)90009-T](https://doi.org/10.1016/0893-6080(91)90009-T). URL <http://www.sciencedirect.com/science/article/pii/089360809190009T>.
- [57] P. Huber and E. Ronchetti. *Robust Statistics*. Wiley Series in Probability and Statistics. Wiley. ISBN 978-1-118-21033-8. URL https://books.google.dk/books?id=j10hquR_j88C.
- [58] S. Hviid, Juul. Working Paper: A regional model of the Danish housing market. URL <http://www.nationalbanken.dk/en/publications/Pages/2017/11/Working-Paper-A-regional-model-of-the-Danish-housing-market.aspx>.
- [59] F. James and M. Roos. Minuit – a system for function minimization and analysis of the parameter errors and correlations. 10:343–367. doi: [10.1016/0010-4655\(75\)90039-9](https://doi.org/10.1016/0010-4655(75)90039-9).
- [60] R. E. Kalman. A new approach to linear filtering and prediction problems. 82:35–45.
- [61] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, and T.-Y. Liu. LightGBM: A Highly Efficient Gradient Boosting Decision Tree. In I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, editors, *Advances in Neural Information Processing Systems 30*, pages 3146–3154. Curran Associates, Inc. URL <http://papers.nips.cc/paper/6907-lightgbm-a-highly-efficient-gradient-boosting-decision-tree.pdf>.
- [62] D. P. Kingma and J. Ba. Adam: A Method for Stochastic Optimization. URL <http://arxiv.org/abs/1412.6980>.
- [63] G. Klambauer, T. Unterthiner, A. Mayr, and S. Hochreiter. Self-Normalizing Neural Networks. URL <http://arxiv.org/abs/1706.02515>.
- [64] A. J. Larkoski, J. Thaler, and W. J. Waalewijn. Gaining (Mutual) Information about Quark/Gluon Discrimination. 2014 (11). ISSN 1029-8479. doi: [10.1007/JHEP11\(2014\)129](https://doi.org/10.1007/JHEP11(2014)129). URL <http://arxiv.org/abs/1408.3122>.

- [65] C. Leys, C. Ley, O. Klein, P. Bernard, and L. Licata. Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. 49(4):764–766. ISSN 0022-1031. doi: 10.1016/j.jesp.2013.03.013. URL <http://www.sciencedirect.com/science/article/pii/S0022103113000668>.
- [66] S. M. Lundberg and S.-I. Lee. A unified approach to interpreting model predictions. In *Proceedings of the 31st International Conference on Neural Information Processing Systems*, NIPS’17, pages 4768–4777. Curran Associates Inc. ISBN 978-1-5108-6096-4. URL <http://dl.acm.org/citation.cfm?id=3295222.3295230>.
- [67] S. M. Lundberg, G. G. Erion, and S.-I. Lee. Consistent Individualized Feature Attribution for Tree Ensembles - SHAP. . URL <http://arxiv.org/abs/1802.03888>.
- [68] S. M. Lundberg, G. G. Erion, and S.-I. Lee. Consistent Individualized Feature Attribution for Tree Ensembles. . URL <http://arxiv.org/abs/1802.03888>.
- [69] A. L. Maas. Rectifier nonlinearities improve neural network acoustic models.
- [70] L. McInnes and J. Healy. UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. URL <http://arxiv.org/abs/1802.03426>.
- [71] T. C. Mills. *Time Series Techniques for Economists / Terence c. Mills*. Cambridge University Press Cambridge [England] ; New York. ISBN 0-521-34339-9 0-521-40574-2. URL <http://www.loc.gov/catdir/toc/cam031/89007187.html>.
- [72] N. Mohd Razali and B. Yap. Power Comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling Tests. 2.
- [73] I. Mulalic, H. Rasmussen, J. Rouwendal, and H. H. Woltmann. The Financial Crisis and Diverging House Prices: Evidence from the Copenhagen Metropolitan Area. ISSN 1556-5068. doi: 10.2139/ssrn.3041272. URL <https://www.ssrn.com/abstract=3041272>.
- [74] Particle Data Group et al. Review of Particle Physics. 98(3):030001. doi: 10.1103/PhysRevD.98.030001. URL <https://link.aps.org/doi/10.1103/PhysRevD.98.030001>.
- [75] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. 12:2825–2830.

- [76] E. Polley and M. van der Laan. Super Learner In Prediction. URL <https://biostats.bepress.com/ucbbiostat/paper266>.
- [77] J. Proriol, J. Jousset, C. Guicheney, A. Falvard, P. Henrard, D. Pallin, P. Perret, and B. Brandl. TAGGING B QUARK EVENTS IN ALEPH WITH NEURAL NETWORKS (comparison of different methods : Neural Networks and Discriminant Analysis). page 27.
- [78] A. Purcell. Go on a particle quest at the first CERN webfest. URL <https://cds.cern.ch/record/1473657>.
- [79] S. Ravanbakhsh, J. Schneider, and B. Poczos. Deep Learning with Sets and Point Clouds. URL <http://arxiv.org/abs/1611.04500>.
- [80] D. N. Reshef, Y. A. Reshef, H. K. Finucane, S. R. Grossman, G. McVean, P. J. Turnbaugh, E. S. Lander, M. Mitzenmacher, and P. C. Sabeti. Detecting Novel Associations in Large Data Sets. 334(6062):1518–1524. ISSN 0036-8075, 1095-9203. doi: 10.1126/science.1205438. URL <http://science.sciencemag.org/content/334/6062/1518>.
- [81] J. W. Rohlfs. *Modern Physics from A to Z*. John Wiley and Sons. ISBN 978-0-471-57270-1.
- [82] P. J. Rousseeuw and C. Croux. Alternatives to the Median Absolute Deviation. 88(424):1273–1283. ISSN 0162-1459. doi: 10.1080/01621459.1993.10476408. URL <https://www.tandfonline.com/doi/abs/10.1080/01621459.1993.10476408>.
- [83] A. Salam. Weak and electromagnetic interactions. In *Selected Papers of Abdus Salam*, volume Volume 5 of *World Scientific Series in 20th Century Physics*, pages 244–254. WORLD SCIENTIFIC. ISBN 978-981-02-1662-7. doi: 10.1142/9789812795915_0034. URL https://www.worldscientific.com/doi/abs/10.1142/9789812795915_0034.
- [84] S. Schmitt. Data Unfolding Methods in High Energy Physics. 137:11008. ISSN 2100-014X. doi: 10.1051/epjconf/201713711008. URL <http://arxiv.org/abs/1611.01927>.
- [85] L. Scodellaro. B tagging in ATLAS and CMS. URL <http://arxiv.org/abs/1709.01290>.
- [86] L. Shapley. A value for n-person games. In *The Shapley Value*, volume 28 of *Annals of Math Studies*, pages 307–317. doi: 10.1017/CBO9780511528446.003.
- [87] T. Sjöstrand, S. Ask, J. R. Christiansen, R. Corke, N. De-sai, P. Ilten, S. Mrenna, S. Prestel, C. O. Rasmussen, and P. Z. Skands. An Introduction to PYTHIA 8.2. 191:159–177. ISSN 00104655. doi: 10.1016/j.cpc.2015.01.024. URL <http://arxiv.org/abs/1410.3012>.

- [88] P. Skands. Peter Skands.
- [89] S. J. Taylor and B. Letham. Forecasting at scale. doi: 10.7287/peerj.preprints.3190v2. URL <https://peerj.com/preprints/3190>.
- [90] i. team. Iminuit – A python interface to minuit. URL <https://github.com/scikit-hep/iminuit>.
- [91] J. Thaler. Report of the Les Houches Quark/Gluon Subgroup. (1):28.
- [92] R. Tibshirani. Regression Shrinkage and Selection via the Lasso. 58(1):267–288. ISSN 0035-9246. URL www.jstor.org/stable/2346178.
- [93] A. Tikhonov. *On the Stability of Inverse Problems*, volume vol. 39 of *Doklady Akademii Nauk SSSR*.
- [94] S. Toghi Eshghi, A. Au-Yeung, C. Takahashi, C. R. Bolen, M. N. Nyachienga, S. P. Lear, C. Green, W. R. Mathews, and W. E. O’Gorman. Quantitative Comparison of Conventional and t-SNE-guided Gating Analyses. 10. ISSN 1664-3224. doi: 10.3389/fimmu.2019.01194. URL <https://www.frontiersin.org/articles/10.3389/fimmu.2019.01194/full>.
- [95] d. L. M. J. van, E. C. Polley, and A. E. Hubbard. Super Learner. 6(1). ISSN 1544-6115. doi: 10.2202/1544-6115.1309. URL <https://www.degruyter.com/view/j/sagmb.2007.6.issue-1/sagmb.2007.6.1.1309/sagmb.2007.6.1.1309.xml>.
- [96] J. J. van der Bij and E. W. N. Glover. Z boson production and decay via gluons. 313(2):237–257. ISSN 0550-3213. doi: 10.1016/0550-3213(89)90317-9. URL <http://www.sciencedirect.com/science/article/pii/0550321389903179>.
- [97] L. van der Maaten and G. Hinton. Visualizing Data using t-SNE. 9:2579–2605. ISSN ISSN 1533-7928. URL <http://www.jmlr.org/papers/v9/vandermaaten08a.html>.
- [98] F. van Veen. The Neural Network Zoo. URL <http://www.asimovinstitute.org/neural-network-zoo/>.
- [99] V. Vapnik. Principles of Risk Minimization for Learning Theory. In *Proceedings of the 4th International Conference on Neural Information Processing Systems*, NIPS’91, pages 831–838. Morgan Kaufmann Publishers Inc. ISBN 978-1-55860-222-9. URL <http://dl.acm.org/citation.cfm?id=2986916.2987018>.
- [100] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. Jarrod Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey, I. Polat, Y. Feng, E. W. Moore,

J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, P. van Mulbregt, and S. . . Contributors. SciPy 1.0—Fundamental Algorithms for Scientific Computing in Python. page arXiv:1907.10121. URL <https://ui.adsabs.harvard.edu/abs/2019arXiv190710121V/abstract>.

- [101] I. Wallach and R. Lilien. The protein–small-molecule database, a non-redundant structural resource for the analysis of protein-ligand binding. 25(5):615–620. ISSN 1367-4803. doi: 10.1093/bioinformatics/btp035. URL <https://academic.oup.com/bioinformatics/article/25/5/615/183421>.
- [102] S. Weinberg. A Model of Leptons. 19(21):1264–1266. doi: 10.1103/PhysRevLett.19.1264. URL <https://link.aps.org/doi/10.1103/PhysRevLett.19.1264>.
- [103] H. Wickham. Tidy data. 59(10):1–23. ISSN 1548-7660. doi: 10.18637/jss.v059.i10. URL <https://www.jstatsoft.org/v059/i10>.
- [104] M. Wobisch and T. Wengler. Hadronization Corrections to Jet Cross Sections in Deep-Inelastic Scattering. URL <http://arxiv.org/abs/hep-ph/9907280>.
- [105] M. Zaheer, S. Kottur, S. Ravanbakhsh, B. Poczos, R. Salakhutdinov, and A. Smola. Deep Sets. URL <http://arxiv.org/abs/1703.06114>.