

Centro Brasileiro de Pesquisas Físicas

COHEP - Coordenação de Física de Altas Energias

Tese de Doutorado

**Busca por gluinos de vida longa em cenários de
supersimetria comprimida usando redes neurais
artificiais**

Gilson Correia Silva

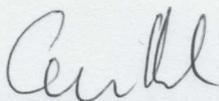
Rio de Janeiro

2019

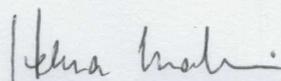
“BUSCA POR GLUINOS DE VIDA LONGA EM CENÁRIOS DE
SUPERSIMETRIA COMPRIMIDA USANDO REDES NEURAIAS
ARTIFICIAIS”

GILSON CORREIA SILVA

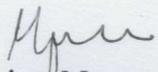
Tese de Doutorado em Física apresentada no
Centro Brasileiro de Pesquisas Físicas do
Ministério da Ciência Tecnologia e Inovação.
Fazendo parte da banca examinadora os
seguintes professores:



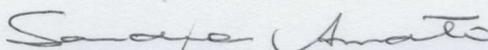
Carsten Hensel – Presidente/Orientador/CBPF



Helena Brandão Malbouisson – Coorientador/UERJ



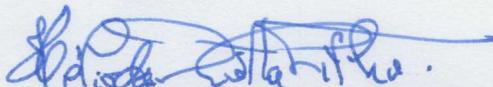
Irina Nasteva - UFRJ



Sandra Filippa Amato - UFRJ



Martin Makler - CBPF



Hélio da Motta Filho - CBPF

Rio de Janeiro, 29 de novembro de 2019.

Agradecimentos

Agradeço aos meus familiares pelo apoio dado a mim durante todos esses anos, entre eles Adailson Correia, Larissa Quirino, Tatiane Silva, Thaís Patrício, Gustavo Francisco, Josenite Alves e principalmente aos meus pais Glaciete Correia e José da Costa, que sempre se esforçaram para garantir que eu continuasse meus estudos, respeitando e apoiando minhas decisões.

Agradeço a todos os professores que tive durante minha carreira acadêmica, entre eles: Luís Fernando, Maria de Fátima, Marcelo Chiapparini, Daniel Barci, José Roberto Mahon, Rudnei Ramos, James Skea, José Abdalla Helayël, Sebastião Alves e Antônio José Accioly. Um agradecimento especial ao professor Carsten Hensel e professora Helena Brandão Malbouisson que me orientaram durante este trabalho.

Agradeço a Matthias Hamer pelos esclarecimentos em alguns pontos da análise. Agradeço a Fábio Lucio pelas inúmeras discussões, auxílios e dicas durante os anos de doutorado, por ter sido um excelente companheiro de trabalho no banco de dados do HCAL no CMS e por sua valorosa amizade.

Agradeço aos cientistas do CMS que me auxiliaram em diversos momentos, entre eles: Kenichi Hatakeyama, Federico De Guio, Salavat Abdoulline, Kevin Pedro, Christopher West, Mariarosaria D'Alfonso, Malik Shahzad Muzaffar e Gabriele Benelli.

Agradeço ao Epica, The Gathering, Nightwish, Raça Negra, Autumn, Lacuna Coil, Anneke Van Giersbergen, Sting, Guns N'Roses, Pearl Jam, Linkin Park, Anathema entre outras bandas que proporcionaram a trilha sonora de diversas jornadas de trabalho.

Agradeço a todos meus amigos que proporcionam alegria e diversão a minha vida, em especial a Ivania Maturana, Leonardo Brasil, Anderson Batista, Ana Luiza, Allan Lima, Michel Pereira, Fábio Costa, André Luiz, Karolyne Roberta, João Gustavo, Alessandra Tiburcio, Jéssica Lyrio, Laís Soares, Rodrigo Genzo, Flávio Santiago, Vinicius Marques, Marvin Ascencio-sosa, Gian Caceres e Adrian Lozano.

Por fim, agradeço ao CNPq pelo financiamento recebido durante os anos de doutorado.

*“The effort to understand the universe is one of the very few things
which lifts human life a little above the level of farce and gives it
some of the grace of tragedy.”*

— STEVEN WEINBERG

Resumo

Esta análise é definida no contexto dos Cenários Supersimétricos Comprimidos, considerando o caso em que quase todas as partículas supersimétricas têm grandes massas, acima da escala TeV e fora do alcance do LHC, mas com gluinos e neutralinos capazes de serem criados nas colisões próton-próton do experimento. Usa-se dados do *Run II* do CMS, com uma energia de centro de massa de $\sqrt{s} = 13$ TeV, para realizar uma pesquisa no canal caracterizado pela produção de dois gluinos de vida longa, cada um deles decaindo em quark, anti-quark e no neutralino mais leve. Embora esse canal já tenha sido pesquisado antes, a característica crucial desse cenário é a, relativamente pequena, diferença de massa entre os gluinos e os neutralinos ($\sim 40 - 100$ GeV) que implica na produção de quarks e anti-quarks de baixa energia, originando jatos no limite da sensibilidade de reconstrução do CMS. O sinal possui uma assinatura difícil de detectar, caracterizada pela presença de energia transversa perdida, vértices secundários e jatos de baixa energia de tal maneira, que esse tipo de evento seria facilmente perdido por buscas anteriores. Na maioria dos eventos, não é possível reconstruir todos os jatos associados ao decaimento do gluino, de modo que se tornou necessário criar uma estratégia para extrair o máximo de informação possível desses eventos, a fim de aumentar sua sensibilidade. A estratégia aplicada usa análise multivariada baseada em métodos de aprendizado de máquina para processar a informação disponível nos objetos reconstruídos a fim de maximizar a sensibilidade da busca.

Palavras-chave: Supersimetria; gluinos de vida longa; LHC; CMS; vértices secundários; jatos; energia transversa perdida; análise multivariada; aprendizado de máquina.

Abstract

This analysis is set in the context of the Supersymmetric Compressed Scenarios considering the case where almost all the supersymmetric particles have big masses above the TeV scale and out of reach of the LHC, but with gluinos and neutralinos able to be created in the proton-proton collisions of the experiment. We use CMS Run II data with a center-of-mass energy of $\sqrt{s} = 13$ TeV to perform a search on the channel characterized by the production of two long-lived gluinos, each one decaying into quark, anti-quark and the lightest neutralino. Although this channel has already been searched before, the crucial feature of this scenarios is the relatively small difference of mass between the gluinos and neutralinos ($\sim 40 - 100$ GeV) which implies in the production of quark and anti-quark with very low energy originating jets in the limit of the reconstruction sensitivity of the CMS. The signal leads to a hard to detect signature, characterized by the presence of missing transverse energy, secondary vertices, and low energy jets in such way that this kind of events would be easily missed by previous searches. In most events, it's not possible to reconstruct all the jets associated from the gluino decay so that became necessary creating a strategy to extract as much information as possible from these events in order to increase their sensitivity. The strategy applied uses multivariate analysis based on machine learning methods to process the information available from reconstructed objects to maximize the sensitivity of the search.

Keywords: supersymmetry; long-lived gluinos; LHC; CMS; secondary vertices; jets; missing transverse energy; multivariate analysis; machine learning.

Lista de Figuras

FIGURA 2.1 – Visão geral de todas as partículas fundamentais descritas pelo Modelo Padrão da física de partículas [9].	9
FIGURA 2.2 – Potencial antes da quebra de simetria, $\nu^2 = 0$ (esquerda) e potencial após a quebra espontânea de simetria, $\nu^2 = -\mu^2/2\lambda$ (direita).	15
FIGURA 2.3 – Diagrama de Feynman da correção quântica dominante para o massa de Higgs proveniente de <i>loops</i> do quark top.	20
FIGURA 2.4 – Exemplo de um espectro de massa “típico” da mSUGRA [44]. Na primeira coluna estão os campos oriundo do setor de Higgs, na segunda estão os sléptons, na terceira estão os charginos e neutralinos e na quarta o gluino e os squarks.	28
FIGURA 2.5 – Evolução das constantes de acoplamento das interações fundamentais no Modelo Padrão (esquerda), onde α_1 é a constante da interação eletromagnética, α_2 é da interação fraca e α_3 é da interação forte, e no MSSM (direita) [47].	29
FIGURA 2.6 – Esquema da aparência típica do espectro de massa das partículas supersimétrica em cenários padrões de SUSY (esquerda) e em cenários de SUSY comprimida (direita).	30
FIGURA 2.7 – Pontos da amostra que resultam em PGU (cinza). Também são mostradas restrições na massa do gluino para os modos de decaimento $\tilde{g} \rightarrow q\bar{q} + \tilde{\chi}_1^0$ (verde) e $\tilde{g} \rightarrow b\bar{b} + \tilde{\chi}_1^0$ (azul). A área amarela denota a região onde o gluino seria a LSP, na qual contém uma restrição (linha vermelha) proveniente de buscas por R-hádrons estáveis no ATLAS [52].	32
FIGURA 2.8 – Distribuição da largura de decaimento na amostra de referência. As linhas tracejadas verticais indicam onde o comprimento de decaimento do gluino, $c\tau_{\tilde{g}}$, atinge $10 \mu\text{m}$ e 1mm [52].	33

- FIGURA 2.9 – Modelo simplificado da produção de dois gluinos na colisão próton-próton decaindo ambos em quark, antiquark e no neutralino LSP. O círculo cinza representa os potenciais canais de produção dos gluinos ($pp \rightarrow \tilde{g}\tilde{g}$). 35
- FIGURA 2.10 – Gráfico da dependência do comprimento de decaimento do gluino com respeito à massa do gluino e à diferença de massa entre o gluino e o neutralino. Gerado com o pacote SDECAY 1.5a [60]. 35
- FIGURA 2.11 – Produção de um par de gluinos na colisão próton-próton, os neutralinos são produzidos em direções opostas quase colineares resultando em uma baixa energia transversa perdida (esquerda). O par de gluinos é produzido junto com um jato do estado de radiação inicial que gera um desalinhamento maior entre os neutralinos produzidos, aumentando a quantidade de energia transversa perdida (direita). 37
- FIGURA 2.12 – Seções de choque de produção de pares de partículas supersimétricas no LHC em energias de centro de massa de $\sqrt{s} = 8 \text{ TeV}$ e $\sqrt{s} = 13 - 14 \text{ TeV}$ [74]. 38
- FIGURA 2.13 – Visão geral dos resultados de buscas por SUSY no CMS envolvendo a produção de pares de gluinos. Os limites observados da massa dos gluinos com 95% C.L. são representados pela barra laranja [75]. 39
- FIGURA 2.14 – Visão geral dos resultados de buscas por SUSY no ATLAS. Os limites observados da massa das partículas produzidas na colisão próton-próton com 95% C.L. são representados pela barra azul [77]. 40
- FIGURA 3.1 – Visão abrangente dos experimentos do LHC localizados em um túnel subterrâneo na fronteira da França com a Suíça. 43
- FIGURA 3.2 – Representação esquemática do complexo do LHC [82]. Os feixes de prótons são acelerados, passo a passo, por uma cadeia de aceleradores. 44
- FIGURA 3.3 – Luminosidade integrada entregue pelo LHC ao experimento CMS em colisões pp durante os anos referentes aos *Runs* 1 e 2 [83]. 46
- FIGURA 3.4 – Visão seccionada do experimento CMS identificando o posicionamento de seus principais componentes [87]. 48
- FIGURA 3.5 – Exemplo de um vetor \vec{p} representando o momento de uma partícula e como suas componentes estão relacionadas ao sistema de coordenadas do CMS. 49

- FIGURA 3.6 – Representação esquemática do detector de píxeis da “Fase 0”, composto por três camadas centrais e quatro tampas, duas de cada lado [91]. 51
- FIGURA 3.7 – Representação esquemática do sistema de trajetórias do CMS no plano $r - z$. A parte roxa representa o detector de píxeis e a parte rosa representa o detector de tiras [93]. 52
- FIGURA 3.8 – Representação esquemática do Calorímetro Eletromagnético no plano $y - z$. As linhas tracejadas denotam a cobertura em pseudorapidez para a região de barril (EB) e região das tampas (EE). [94]. 52
- FIGURA 3.9 – Representação esquemática do Calorímetro Hadrônico no plano $r - z$ e seus subdetectores: HB, HE, HO e HF. Os segmentos roxos são partes do sistema de múons [96]. 54
- FIGURA 3.10 – Representação esquemática do Sistema de Múons no plano $r - z$ e seus subdetectores: DT, RPC e CSC [99]. 56
- FIGURA 3.11 – Os gráficos acima estão normalizados e mostram as distribuições de interações verdadeiras dos dados coletados em 2016 (esquerda) e da amostra de $t\bar{t}$ +jatos produzida para 2016 (direita). O gráfico abaixo à esquerda mostra os pesos de *pileup*, que serão aplicados a cada evento simulado, em função do número de interações verdadeiras do evento. Ele é obtido pela divisão das distribuições acima. No gráfico abaixo à direita, estão reunidas as distribuições anteriores dos dados e da amostra de simulação, e também a nova distribuição da amostra de $t\bar{t}$ +jatos corrigida (vermelho) após serem aplicados os pesos de *pileup* nos eventos. 61
- FIGURA 3.12 – Figura esquemática de uma fatia transversal do detector CMS, da região de interação do feixe ao subdetector de múons, mostrando a interação de partículas específicas com os subdetectores do CMS. O múon e o píon são carregados positivamente e o elétron é carregado negativamente. (crédito: David Barney/CERN) 62
- FIGURA 3.13 – Esquema de uma colisão entre prótons resultando na produção de um jato composto por partículas estáveis que deixam sinais nos subdetectores do CMS [110]. 63
- FIGURA 3.14 – Exemplo do desempenho dos três algoritmos de agrupamento sequencial, com $R = 1$, na resolução de jatos [112]. 65
- FIGURA 3.15 – Resposta da energia dos jatos no detector CMS em função de $|\eta|$ para jatos com diferentes momentos transversos [113]. 66

FIGURA 3.16 – Representação esquemática da definição de \vec{E}_T^{miss} e E_T^{miss} [112].	68
FIGURA 3.17 – Exemplo de um evento com a reconstrução de um vértice secundário. L_{xy} é a distância entre o vértice secundário e o vértice primário no plano transversal e d_0 é o parâmetro de impacto transversal do traço, definido como a distância de maior proximidade entre um traço e o vértice primário no plano transversal [122].	70
FIGURA 4.1 – Jato composto por três constituintes, cujos traços estão representados pelas setas, localizados próximos aos vértices A e B. Os valores relativos ao p_T dos constituintes estão especificados acima das setas.	75
FIGURA 4.2 – Exemplos de eventos de QCD e SUSY com energia transversal perdida [132]. O vetor cinza representa o p_T correto do jato e o vetor preto o p_T reconstruído. O vetor azul representa \vec{H}_T^{miss} (ou \vec{E}_T^{miss}) e o vetor tracejado preto representa o neutralino LSP nos eventos de sinal.	78
FIGURA 4.3 – Comportamento das variáveis angulares com respeito a $\Delta\varphi_i$ para diversos valores de f_i [132]. A linha pontilhada azul representa um típico corte no limiar inferior γ_0 aplicado para suprimir eventos de QCD.	79
FIGURA 4.4 – Distribuições dos eventos de QCD e do sinal(1000,100) nas variáveis $\Delta\varphi_{\text{min}}^*$, $\hat{\omega}_{\text{min}}$ e χ_{min} , e a curva ROC medindo a performance das variáveis na discriminação entre as duas classes de eventos.	80
FIGURA 4.5 – Assinatura do sinal - um jato energético originário do vértice primário (ISR jato), dois vértices secundários com jatos pouco energéticos associados a eles e energia transversal perdida.	82
FIGURA 4.6 – Distribuições do momento transversal dos quarks produzidos no decaimento dos gluinos. À esquerda estão as distribuições correspondentes aos quarks com menor p_T e à direita estão as distribuições correspondentes aos quarks de maior p_T	83
FIGURA 4.7 – Distribuições de E_T^{miss} correspondentes a eventos com 0, 1 ou 2 pártons na radiação de estado inicial.	84
FIGURA 4.8 – Distribuições normalizadas da multiplicidade de SV1Jet's e SV2Jets's reconstruídos por evento.	86
FIGURA 4.9 – Distribuição do ΔR de todos os pares quark-jato formados na etapa dois do algoritmo de combinação.	87

- FIGURA 4.10 – Gráficos de eficiência na reconstrução de SV1Jet’s, SV2Jets’s e SVs, considerando-se a definição forte para GD jatos e PU jatos. 88
- FIGURA 4.11 – Gráficos de eficiência na reconstrução de SV1Jet’s, SV2Jets’s e SVs, considerando-se a definição fraca para GD jatos e PU jatos. 89
- FIGURA 4.12 – Gráficos de eficiência do algoritmo de associação a SV1Jet’s, SV2Jets’s e SVs. Nos gráficos de cima, os GD jatos são definidos pela situação 1, e nos gráficos de baixo, eles são definidos pela situação 2. 91
- FIGURA 4.13 – Gráficos bidimensionais da eficiência dos *triggers* de busca com respeito as variáveis *offline* H_T^{miss} e E_T^{miss} , sem a aplicação do corte em Δ_T^{miss} (esquerda) e com a aplicação do corte em Δ_T^{miss} (direita). 95
- FIGURA 4.14 – Curva de eficiência com respeito à variável E_T^{miss} após ter sido aplicado o corte $H_T^{\text{miss}} \geq 220$ GeV (esquerda). Curva de eficiência com respeito à variável H_T^{miss} após ter sido aplicado o corte $E_T^{\text{miss}} \geq 240$ GeV (direita). As linhas tracejadas verticais denotam os limiares inferiores da região em que se alcança a eficiência máxima. 95
- FIGURA 4.15 – Comportamento anômalo do *trigger* de buscas para $E_T^{\text{miss}} > 500$ GeV, com respeito aos dados reais, em que a eficiência decresce de forma sistemática. O mesmo comportamento não é observado na amostra de dados de MC em marrom. 97
- FIGURA 4.16 – Fatores de escala aplicados nos eventos de MC para que fiquem em acordo com a eficiência dos dados reais. 97
- FIGURA 4.17 – Distribuições de eficiência dos *triggers* de busca nas variáveis presentes na tabela 4.3. O valor nominal obtido para a eficiência e seu erro sistemático estão representados em vermelho nos gráficos. 98
- FIGURA 4.18 – O gráfico à esquerda mostra a distribuição dos eventos de *background*, que passaram pela seleção inicial, com respeito ao número de elétrons e múons no evento. A linha vertical tracejada em 1 mostra o ponto de corte, sendo selecionados apenas os eventos com $N_{e,\mu} = 0$. O gráfico à direita mostra a distribuição dos eventos de *background*, que passaram pelo corte em $N_{e,\mu}$, com respeito a variável $\hat{\omega}_{\text{min}}$. A linha vertical tracejada em 0,4 aponta a localização do corte, sendo selecionados apenas os eventos com $\hat{\omega}_{\text{min}} \geq 0,4$ 100
- FIGURA 4.19 – Representação esquemática das regiões de busca. Os cones representam os jatos e os círculos representam os vértices secundários do SV2Jets ou SV1Jet’s presentes em cada sub-região. 103

- FIGURA 4.20 – Fração do número de eventos de sinal em cada sub-região de busca. No gráfico à esquerda são expostos três pontos do *grid* de sinal com $\Delta m = 60 \text{ GeV}$ e diferentes valores de $m_{\tilde{g}}$, já no gráfico à direita são expostos três pontos do *grid* de sinal com $m_{\tilde{g}} = 1000 \text{ GeV}$ e diferentes valores de Δm 104
- FIGURA 4.21 – Estimativa, baseada na simulação, dos percentuais das contribuições de cada *background* nas quatro sub-regiões de busca. 105
- FIGURA 4.22 – Esquema das regiões definidas na análise, todas elas consideram que os eventos tenham passado pela seleção inicial. A região amarela é a região de busca. As regiões de controle dos eventos de $t\bar{t}$ +Jatos, $W(l\nu_l)$ +Jatos, QCD e $Z(\nu\nu)$ +Jatos estão, respectivamente, em azul, roxo, vermelho e verde. 106
- FIGURA 4.23 – Correções de forma dos eventos de $W(l\nu_l)$ +Jatos e suas incertezas, aplicadas em sub-regiões definidas pelas variáveis $N_{\text{jets}}^{\text{PV}}$ e H_{T}^{PV} 108
- FIGURA 4.24 – Distribuições das variáveis globais na região de controle de $W(l\nu_l)$ +Jatos, após as correções da simulação. 109
- FIGURA 4.25 – Correções de forma dos eventos de $t\bar{t}$ +Jatos e suas incertezas, aplicadas em sub-regiões definidas pelas variáveis $H_{\text{T}}^{\text{miss}}$ e H_{T}^{PV} 111
- FIGURA 4.26 – Distribuições das variáveis globais na região de controle de $t\bar{t}$ +Jatos, após as correções da simulação. 112
- FIGURA 4.27 – Gráfico de controle da pureza e eficiência dos eventos de QCD com relação aos pontos de corte na variável $\hat{\omega}_{\text{min}}$ (esquerda). Gráfico de controle da contaminação por eventos da amostra de sinal, com $m_{\tilde{g}} = 1000 \text{ GeV}$ e $\Delta m = 100 \text{ GeV}$, com relação aos pontos de corte na variável $\hat{\omega}_{\text{min}}$ (direita). A linha tracejada vertical denota o valor de $\hat{\omega}_{\text{min}}$ em que foi aplicado o corte que define a região de controle de QCD. 113
- FIGURA 4.28 – Correções de forma dos eventos de QCD e suas incertezas, aplicadas em sub-regiões definidas pelas variáveis N_{jets} e f_{max} 114
- FIGURA 4.29 – Distribuições das variáveis globais na região de controle de QCD após as correções da simulação. 115
- FIGURA 4.30 – Distribuição dos eventos de sinal e *background* no intervalo $0,78 \leq \hat{\omega}_{\text{min}} < 0,8$, descrito na figura 4.22, com respeito à variável discriminante obtida por uma análise multivariada usando redes neurais, a qual será descrita em detalhes no capítulo 5. 117

- FIGURA 4.31 –Gráfico de controle da pureza e eficiência dos eventos de $Z(\nu\nu)+\text{Jatos}$ com relação aos pontos de corte na variável discriminante (esquerda). Gráfico de controle da contaminação por eventos da amostra de sinal, com $m_{\tilde{g}} = 1000 \text{ GeV}$ e $\Delta m = 100 \text{ GeV}$, e com $m_{\tilde{g}} = 1000 \text{ GeV}$ e $\Delta m = 30 \text{ GeV}$, com relação aos pontos de corte na variável discriminante (direita). A linha tracejada vertical denota o valor da variável discriminante em que foi aplicado o corte que define a região de controle de $Z(\nu\nu)+\text{Jatos}$ 117
- FIGURA 4.32 –Correções de forma dos eventos de $Z(\nu\nu)+\text{Jatos}$ e suas incertezas, aplicadas em sub-regiões definidas pelas variáveis $E_{\text{T}}^{\text{miss}}$ e $\text{Jet}_{\text{S}}^{\text{PV}} p_{\text{T}}$. . . 118
- FIGURA 4.33 –Distribuições das variáveis globais na região de controle de $Z(\nu\nu)+\text{Jatos}$ após as correções da simulação. 119
- FIGURA 4.34 –Distribuições dos *backgrounds* e sinais, com relação às variáveis regionais, na sub-região de busca R1. 121
- FIGURA 4.35 –Distribuições dos *backgrounds* e sinais, com relação às variáveis globais, na sub-região de busca R1. 122
- FIGURA 4.36 –Distribuições dos *backgrounds* e sinais, com relação às variáveis regionais, na sub-região de busca R2. 124
- FIGURA 4.37 –Distribuições dos *backgrounds* e sinais, com relação às variáveis globais, na sub-região de busca R2. 125
- FIGURA 4.38 –Distribuições dos *backgrounds* e sinais, com relação às variáveis regionais, na sub-região de busca R3. 126
- FIGURA 4.39 –Distribuições dos *backgrounds* e sinais, com relação às variáveis globais, na sub-região de busca R3. 127
- FIGURA 4.40 –Distribuições dos *backgrounds* e sinais, com relação às variáveis globais, na sub-região de busca R4. 128
- FIGURA 4.41 –Região Leptônica: possui pelo menos um múon ou elétron reconstruído com $p_{\text{T}} \geq 10 \text{ GeV}$ e o valor de $\hat{\omega}_{\text{min}}$ é maior que 0,4. 129
- FIGURA 4.42 –Região de $Z(\nu\nu)+\text{Jatos}$: não possui múons ou elétrons reconstruídos, o valor de $\hat{\omega}_{\text{min}}$ se encontra no intervalo $0,78 \leq \hat{\omega}_{\text{min}} < 0,8$ e o valor do discriminante da análise multivariada é menor que 0,6. 129
- FIGURA 4.43 –Região de QCD: não possui múons ou elétrons reconstruídos e o valor de $\hat{\omega}_{\text{min}}$ é menor que 0,15. 129

- FIGURA 5.1 – Exemplo de um sistema que possui duas classes de eventos (A e B), em que as informações conhecidas sobre cada classe estão armazenadas nas variáveis x_1 e x_2 . A curva de separação (em verde) é obtida da análise multivariada realizada nos dados simulados e posteriormente é aplicada nos dados reais para realizar a classificação dos eventos. 132
- FIGURA 5.2 – Processo completo do treinamento e seleção de modelos em análises multivariadas usando técnicas de aprendizagem de máquina. 133
- FIGURA 5.3 – Exemplo de uma situação de *overtraining* em que os pontos azuis e vermelhos representam duas classes diferentes de eventos na amostra de treinamento. A curva preta representa a curva de separação ideal, que considera somente as propriedades gerais do sistema, já a curva verde é resultado de um treinamento excessivo em que o modelo considera na classificação as flutuações estatísticas da amostra de treinamento, que não serão observadas nos dados reais. 134
- FIGURA 5.4 – Exemplo da distribuição do discriminante de um modelo de classificação entre duas classes de eventos, com respeito aos eventos da amostra de treinamento. As diferentes classes estão representadas pelas cores vermelha e azul. A reta vertical em um valor ligeiramente menor que zero, representa o ponto de corte escolhido para separar as duas classes de eventos do problema. 134
- FIGURA 5.5 – Exemplo de uma árvore de decisão aplicada sobre um sistema formado por duas classes de eventos, com 500 eventos na classe de sinal (azul) e 800 eventos na classe de *background* (vermelho). A árvore possui 9 nódulos, formados com cortes em três variáveis diferentes (x_1 , x_2 e x_3), e gera 3 regiões de sinal e duas regiões de *background* no espaço das variáveis. 135
- FIGURA 5.6 – Representação de um neurônio biológico [153]. 139
- FIGURA 5.7 – Representação esquemática de um Neurônio Artificial. 140
- FIGURA 5.8 – Exemplo do funcionamento de uma rede Perceptron que recebe entradas compostas por duas variáveis $\{x_1, x_2\}$ 141
- FIGURA 5.9 – Três casos possíveis do produto escalar $\mathbf{w} \cdot \mathbf{x}^a$ 142
- FIGURA 5.10 – Os dois casos possíveis em que a saída da rede Perceptron não é igual à saída esperada. 143
- FIGURA 5.11 – Evolução da função erro quadrático com relação ao vetor de pesos (esquerda) e com relação ao número de épocas (direita). 145

- FIGURA 5.12 – Exemplo de um sistema com duas classes de eventos, A e B, cuja a informação está contida em um conjunto de duas variáveis, x_1 e x_2 . A reta diagonal representa o hiperplano de separação ótimo encontrado pela rede Adaline. 145
- FIGURA 5.13 – Exemplo da forma geral de uma rede Perceptron Multicamadas com 3 camadas. O quadrado verde no final da rede representa o pós-processamento dos sinais de saída para a identificação da classe do evento, quando há mais de um neurônio na camada de saída. 146
- FIGURA 5.14 – Rede Perceptron Multicamadas com três camadas e apenas um neurônio na última camada. Abaixo da rede é mostrado o fluxo completo de processamento da informação na rede, desde as variáveis de entrada, x_i , até o sinal de saída, Y^3 149
- FIGURA 5.15 – Funções de ativação usadas nas redes MLP. 150
- FIGURA 5.16 – Função utilizada na ponderação dos eventos da classe de sinal, descrita pela equação (5.17). 153
- FIGURA 5.17 – Figura esquemática representando a divisão dos dados simulados entre as amostras de treinamento, validação e teste. 154
- FIGURA 5.18 – Distribuições das classes de sinal e *background* com respeito as 12 variáveis de entrada na amostra de treinamento. 155
- FIGURA 5.19 – Matriz de correlação linear das classes de sinal (acima) e *background* (abaixo) com respeito as 12 variáveis de entrada na amostra de treinamento. 156
- FIGURA 5.20 – Valor da função de perda no decorrer das épocas de treinamento da rede com 2 camadas, 48 neurônios por camada e normalização dos eventos da classe de *background* igual a 2. O treinamento foi encerrado na época 54 e o resultado é a rede neural com configuração de pesos da época 24, denotada no gráfico pela linha vertical tracejada. 157
- FIGURA 5.21 – Curva ROC das duas redes MLP e das duas BDTs com melhor performance. 158
- FIGURA 5.22 – Validação da compatibilidade entre as estimativas da amostra de treinamento e as estimativas das outras cinco amostras de dados simulados. 159

- FIGURA 5.23 – Gráficos de R1: Histograma de controle da contaminação de sinal (acima à esquerda). Curva de significância dos modelos com $m_{\tilde{g}} = 1000$ GeV (acima à direita). Distribuição da variável discriminante correspondente à RNA com melhor performance (abaixo). A região de controle e os 13 *bins* de busca são delimitados pelas linhas verticais em 0,4, 0,45, 0,5, 0,55, 0,6, 0,65, 0,7, 0,75, 0,8, 0,85, 0,9, 0,94 e 0,97. 162
- FIGURA 5.24 – Gráficos de R2: Histograma de controle da contaminação de sinal (acima à esquerda). Curva de significância dos modelos com $m_{\tilde{g}} = 1000$ GeV (acima à direita). Distribuição da variável discriminante correspondente à RNA com melhor performance (abaixo). A região de controle e os 13 *bins* de busca são delimitados pelas linhas verticais em 0,32, 0,4, 0,45, 0,5, 0,55, 0,6, 0,65, 0,7, 0,75, 0,8, 0,85, 0,9 e 0,95. 163
- FIGURA 5.25 – Gráficos de R3: Histograma de controle da contaminação de sinal (acima à esquerda). Curva de significância dos modelos com $m_{\tilde{g}} = 1000$ GeV (acima à direita). Distribuição da variável discriminante correspondente à RNA com melhor performance (abaixo). A região de controle e os 9 *bins* de busca são delimitados pelas linhas verticais em 0,47, 0,55, 0,6, 0,65, 0,7, 0,75, 0,8, 0,85 e 0,9. 164
- FIGURA 5.26 – Gráficos de R4: Histograma de controle da contaminação de sinal (acima à esquerda). Curva de significância dos modelos com $m_{\tilde{g}} = 1000$ GeV (acima à direita). Distribuição da variável discriminante correspondente à RNA com melhor performance (abaixo). A região de controle e os 10 *bins* de busca são delimitados pelas linhas verticais em 0,50, 0,55, 0,6, 0,65, 0,7, 0,75, 0,8, 0,85, 0,9 e 0,95. 165
- FIGURA 6.1 – Estimativa das contribuições de processos de *background* e sinal em cada *bin* de busca. 166
- FIGURA 6.2 – Estimativa nominal do número de eventos de $t\bar{t}$ +Jatos nos *bins* de busca junto com algumas estimativas de suas incertezas sistemáticas. 169
- FIGURA 6.3 – Estimativa nominal do número de eventos de $W(l\nu)$ +Jatos nos *bins* de busca junto com algumas estimativas de suas incertezas sistemáticas. 170
- FIGURA 6.4 – Estimativa nominal do número de eventos de $Z(\nu\nu)$ +Jatos nos *bins* de busca junto com algumas estimativas de suas incertezas sistemáticas. 171

FIGURA 6.5 – Limites de exclusão esperados com 95%C.L. e suas faixas de incerteza para os modelos com massa do gluino igual a 750, 1000 e 1250 GeV.	173
FIGURA A.1 – Figura esquemática mostrando o fluxo de ações do processo automatizado.	194
FIGURA C.1 – Validação das variáveis regionais na sub-região CR1 da região leptônica.	199
FIGURA C.2 – Validação das variáveis globais na sub-região CR1 da região leptônica. 200	
FIGURA C.3 – Validação das variáveis regionais na sub-região CR2 da região leptônica.	201
FIGURA C.4 – Validação das variáveis globais na sub-região CR2 da região leptônica. 202	
FIGURA C.5 – Validação das variáveis regionais na sub-região CR3 da região leptônica.	203
FIGURA C.6 – Validação das variáveis globais na sub-região CR3 da região leptônica. 204	
FIGURA C.7 – Validação das variáveis globais na sub-região CR4 da região leptônica. 205	
FIGURA C.8 – Validação das variáveis regionais na sub-região CR1 da região de $Z(\nu\nu)+Jatos$	206
FIGURA C.9 – Validação das variáveis globais na sub-região CR1 da região de $Z(\nu\nu)+Jatos$	207
FIGURA C.10 – Validação das variáveis regionais na sub-região CR2 da região de $Z(\nu\nu)+Jatos$	208
FIGURA C.11 – Validação das variáveis globais na sub-região CR2 da região de $Z(\nu\nu)+Jatos$	209
FIGURA C.12 – Validação das variáveis regionais na sub-região CR3 da região de $Z(\nu\nu)+Jatos$	210
FIGURA C.13 – Validação das variáveis globais na sub-região CR3 da região de $Z(\nu\nu)+Jatos$	211
FIGURA C.14 – Validação das variáveis globais na sub-região CR4 da região de $Z(\nu\nu)+Jatos$	212
FIGURA C.15 – Validação das variáveis regionais na sub-região CR1 da região de QCD.	213
FIGURA C.16 – Validação das variáveis globais na sub-região CR1 da região de QCD. 214	

FIGURA C.17 –Validação das variáveis regionais na sub-região CR2 da região de QCD.	215
FIGURA C.18 –Validação das variáveis globais na sub-região CR2 da região de QCD.	216
FIGURA C.19 –Validação das variáveis regionais na sub-região CR3 da região de QCD.	217
FIGURA C.20 –Validação das variáveis globais na sub-região CR3 da região de QCD.	218
FIGURA C.21 –Validação das variáveis globais na sub-região CR4 da região de QCD.	219
FIGURA D.1 –Matrizes de correlação linear das variáveis de entrada em R1.	221
FIGURA D.2 –Distribuições das variáveis de entrada em R1.	222
FIGURA D.3 –Gráfico de evolução da função de perda (acima), checagem de <i>over-training</i> (centro) e avaliação da performance (abaixo) da RNA usada em R1.	223
FIGURA D.4 –Matrizes de correlação linear das variáveis de entrada em R2.	224
FIGURA D.5 –Distribuições das variáveis de entrada em R2.	225
FIGURA D.6 –Gráfico de evolução da função de perda (acima), checagem de <i>over-training</i> (centro) e avaliação da performance (abaixo) da RNA usada em R2.	226
FIGURA D.7 –Matrizes de correlação linear das variáveis de entrada em R3.	227
FIGURA D.8 –Distribuições das variáveis de entrada em R3.	228
FIGURA D.9 –Gráfico de evolução da função de perda (acima), checagem de <i>over-training</i> (centro) e avaliação da performance (abaixo) da RNA usada em R3.	229
FIGURA D.10 –Matrizes de correlação linear das variáveis de entrada em R4.	230
FIGURA D.11 –Distribuições das variáveis de entrada em R4.	231
FIGURA D.12 –Gráfico de evolução da função de perda (acima), checagem de <i>over-training</i> (centro) e avaliação da performance (abaixo) da RNA usada em R4.	232

Lista de Tabelas

TABELA 2.1 – Composição dos supercampos no MSSM.	25
TABELA 4.1 – Fração de gluinos decaindo em intervalos de comprimento distintos para diferentes pontos do <i>grid</i> de sinal.	85
TABELA 4.2 – Valores obtidos para os parâmetros da função f , após ter sido realizado o ajuste com a curva de eficiência dos <i>triggers</i> de busca do gráfico à esquerda na figura 4.14.	96
TABELA 4.3 – Estimativa do erro sistemático da eficiência dos <i>triggers</i> de busca. São apresentados os valores obtidos nos ajustes (coluna 2), referentes a cada variável descrita na coluna 1. Nas colunas 3 e 4 são expostos, respectivamente, os resultados dos ajustes mais 1σ de suas incertezas e menos 1σ de suas incertezas. A coluna 5, mostra os valores absolutos da maior diferença entre os valores das colunas 3 e 4 com relação ao valor nominal, 97,72%.	99
TABELA 4.4 – Fluxo de cortes sobre alguns pontos do <i>grid</i> de sinal e sobre os <i>backgrounds</i> , começando com os eventos que passaram pela seleção inicial.	101
TABELA 4.5 – Fatores de escala global para eventos de $t\bar{t}$ +Jatos e $W(l\nu_l)$ +Jatos.	107
TABELA 4.6 – Fatores de escala regionais para eventos de $W(l\nu_l)$ +Jatos.	108
TABELA 4.7 – Fatores de correção da radiação de estado inicial dos eventos de $t\bar{t}$ +Jatos.	110
TABELA 4.8 – Fatores de escala regionais para eventos de $t\bar{t}$ +Jatos.	111
TABELA 4.9 – Fator de escala global para eventos de QCD.	114
TABELA 4.10 – Fatores de escala regionais para eventos de QCD.	114
TABELA 4.11 – Fator de escala global para eventos de $Z(\nu\nu)$ +Jatos.	118
TABELA 4.12 – Fatores de escala regionais para eventos de $Z(\nu\nu)$ +Jatos.	118

TABELA 5.1 – Exposição das características de diferentes tipos de redes neurais.	151
TABELA B.1 – Amostras de Monte Carlo dos Backgrounds.	195
TABELA B.2 – Amostras de Monte Carlo dos Backgrounds. (Continuação)	196
TABELA B.3 – Amostras de Monte Carlo do Sinal.	197
TABELA B.4 – Amostras de Dados.	198

Sumário

1	INTRODUÇÃO	5
2	MOTIVAÇÃO FÍSICA	7
2.1	Modelo Padrão da Física de Partículas	7
2.1.1	O Espectro de Partículas	7
2.1.2	Simetrias e Interações	9
2.1.2.1	Simetria de Poincaré	10
2.1.2.2	Teorias de gauge	11
2.1.2.3	A Teoria Eletrofraca	12
2.1.2.4	O Mecanismo de Higgs	14
2.1.2.5	Acoplamentos de Yukawa	16
2.1.2.6	A Cromodinâmica Quântica	17
2.1.3	Questões Abertas no Modelo Padrão	19
2.2	Supersimetria	20
2.2.1	Modelo Padrão Supersimétrico Mínimo	23
2.2.1.1	Quebra de SUSY e o espectro de partículas	24
2.2.1.2	Soluções para questões abertas no Modelo Padrão	28
2.2.2	Supersimetria Comprimida	29
2.2.2.1	Modelo Simplificado	33
2.2.2.2	Caracterização e Cinemática dos Estados Finais	35
2.3	Buscas por Supersimetria no LHC	37
3	EXPERIMENTO	42
3.1	O LHC	42

3.1.1	Colisões próton-próton	46
3.2	O Detector CMS	47
3.2.1	Coordenadas do Detector	48
3.2.2	Sistema de Trajetórias	49
3.2.3	Calorímetro Eletromagnético (ECAL)	51
3.2.4	Calorímetro Hadrônico (HCAL)	53
3.2.5	Solenóide Supercondutor	55
3.2.6	Sistema de Múons	55
3.2.7	<i>Triggers</i>	58
3.3	Reconstrução de Eventos	59
3.3.1	Pileup	59
3.3.2	Fluxo de Partículas	60
3.3.3	Jatos	63
	3.3.3.1 Algoritmos de reconstrução de jatos	63
	3.3.3.2 Correções à energia dos jatos	64
3.3.4	Energia Transversa Perdida	67
3.3.5	Filtros de MET	68
3.3.6	Vértices Primários e Secundários	69
3.3.7	Marcadores de b-jatos	71
4	SELEÇÃO E TRATAMENTO DE DADOS	73
4.1	Variáveis Globais	73
4.1.1	Variáveis dos Vértices	74
4.1.2	Variáveis Angulares	76
4.2	Amostras de dados	80
4.3	Estudo do Sinal	82
4.4	Eficiência do Trigger	92
4.5	Região de busca	99
4.6	Modelagem do Background	103
4.6.1	$W(l\nu_l)$ +Jatos	105

4.6.2	$t\bar{t}$ +Jatos	110
4.6.3	QCD	113
4.6.4	$Z(\nu\nu)$ +Jatos	116
4.7	Variáveis nas sub-regiões de busca	120
4.7.1	Variáveis de R1	120
4.7.2	Variáveis de R2	123
	4.7.2.1 Variáveis <i>razor</i>	123
4.7.3	Variáveis de R3	126
4.7.4	Variáveis de R4	128
4.7.5	Validação das Variáveis	129
5	ANÁLISE MULTIVARIADA	130
5.1	Aprendizado de Máquina	131
5.2	Árvores de Decisão	133
5.3	Redes Neurais Artificiais	137
5.3.1	Neurônio Biológico e Artificial	139
5.3.2	Rede Perceptron	140
5.3.3	Rede Adaline	143
5.3.4	Aspectos gerais das RNAs	145
	5.3.4.1 Treinamento	146
	5.3.4.2 Pré-processamento das entradas	147
5.3.5	Rede Perceptron Multicamadas	148
5.4	Região de controle de $Z(\nu\nu)$+Jatos	151
5.4.1	Amostras de entrada	152
5.4.2	Treinamento	157
5.4.3	Checagem de <i>overtraining</i>	158
5.5	Regiões de sinal	159
5.5.1	Análise multivariada em R1	162
5.5.2	Análise multivariada em R2	163
5.5.3	Análise multivariada em R3	164

5.5.4	Análise multivariada em R4	165
6	RESULTADOS	166
6.1	Incertezas Sistemáticas	166
6.2	Limites de Exclusão	172
7	CONCLUSÕES E PERSPECTIVAS FUTURAS	174
	REFERÊNCIAS	177
	APÊNDICE A – SERVIÇO NO BANCO DE DADOS DO HCAL NO CMS	189
A.1	Trabalho no banco de dados	189
A.2	Trabalho no projeto de automatização	191
	APÊNDICE B – AMOSTRAS DE MONTE CARLO E DADOS	195
B.1	Amostras de Monte Carlo dos Backgrounds	195
B.2	Amostras de Monte Carlo do Sinal	197
B.3	Amostras de Dados Reais	198
	APÊNDICE C – VALIDAÇÃO DAS VARIÁVEIS DE BUSCA	199
C.1	Região Leptônica	199
C.1.1	Variáveis de CR1	199
C.1.2	Variáveis de CR2	201
C.1.3	Variáveis de CR3	203
C.1.4	Variáveis de CR4	205
C.2	Região de $Z(\nu\nu)+\text{Jatos}$	206
C.2.1	Variáveis de CR1	206
C.2.2	Variáveis de CR2	208
C.2.3	Variáveis de CR3	210
C.2.4	Variáveis de CR4	212
C.3	Região de QCD	213
C.3.1	Variáveis de CR1	213
C.3.2	Variáveis de CR2	215

C.3.3	Variáveis de CR3	217
C.3.4	Variáveis de CR4	219
APÊNDICE D	– MVA GRÁFICOS	220
D.1	Gráficos de R1	221
D.2	Gráficos de R2	224
D.3	Gráficos de R3	227
D.4	Gráficos de R4	230

1 Introdução

A física de partículas e altas energias é a área da física que busca descrever a composição e interação da matéria no seu nível mais fundamental. Os constituintes básicos da matéria são chamados de partículas elementares. Exemplos de partículas atualmente consideradas elementares são os elétrons, múons, quarks e neutrinos. No caso particular dos quarks, eles não são vistos isolados na natureza, apenas em partículas compostas e neutras de cor formadas pelo agrupamento de quarks, chamadas de hádrons, como os prótons, nêutrons e píons. Os experimentos criados com o objetivo de detectar essas partículas fizeram com que a física de partículas evoluísse muito nas últimas décadas se tornando uma área complexa e intrigante. Os conhecimentos adquiridos ao longo dos anos resultaram na formulação de um modelo matemático que usando o conceito de campos e simetrias descreve muito bem as características das partículas observadas até agora e como elas interagem entre si; tal modelo foi denominado Modelo Padrão (MP) da física de partículas.

O MP está longe de ser uma teoria completa e apresenta deficiências pontuais e questões em aberto. Devido a isso, surgiram no decorrer dos anos várias propostas de extensões para ele, sendo a Supersimetria (*Supersymmetry* - SUSY) uma das propostas que se mantém com posição de destaque devido à sua consistência teórica e por ser capaz de dar uma solução natural a alguns problemas do MP como divergências que aparecem no cálculo da massa do bóson de Higgs (problema da Hierarquia), que só são contornadas no MP fazendo um ajuste fino, não natural, nos parâmetros da teoria; também temos a falta de candidatos, dentro do espectro do MP, a serem os constituintes da matéria escura; e a não unificação das forças fundamentais eletrofraca e forte em extrapolações para escalas de energia elevadas, onde espera-se que elas se unifiquem.

A busca por Supersimetria é um dos objetivos principais do maior acelerador de partículas existente atualmente, o LHC (*Large Hadron Collider*), porém até o momento não foi descoberta nenhuma evidência de sua existência. Visto isso, uma das conjecturas que podemos fazer é que talvez as análises padrões que vêm sendo feitas não sejam sensíveis a SUSY. Esta pesquisa é estabelecida considerando essa possibilidade, sendo assim, utiliza-se dados provenientes do CMS (*Compact Muon Solenoid*) [1], um dos quatro grandes experimentos que fazem parte do LHC, e analisamos um canal de busca promissor no

contexto dos Cenários Supersimétricos Comprimidos (*Compressed Supersymmetric Scenarios* - CSS), onde quase todas as partículas supersimétricas possuem massas grandes, acima da escala dos TeV e fora do alcance de produção no LHC, mas com o gluino e neutralino mais leve possuindo massas acessíveis ao LHC e com uma diferença de massa da ordem de 40 – 100 GeV, considerada relativamente pequena em espectros supersimétricos e que dá origem à denominação de cenários comprimidos.

O canal de busca desta análise consiste na produção de dois gluinos, que naturalmente possuem vida longa em cenários comprimidos, com cada um deles decaindo em quark, antiquark e no neutralino mais leve. Os jatos produzidos nesses decaimentos não são sempre possíveis de serem reconstruídos por terem uma energia muito baixa, uma vez que a energia transferida para eles é limitada pela diferença de massa entre o gluino e o neutralino mais leve. Análises anteriores utilizando o mesmo canal de busca são pouco sensíveis a eventos de cenários comprimidos por tipicamente requererem a reconstrução de 2 ou mais jatos com $p_T > 30$ GeV [2,3]. Tendo em vista a dificuldade em se detectar eventos de cenários comprimidos, são usadas técnicas de análise multivariada na busca, de modo a aumentar sua sensibilidade. Mais especificamente são utilizados métodos de aprendizado de máquina que estão cada vez mais presentes nas análises físicas e reconstrução de objetos na área de física de partículas e altas energias.

No capítulo 2, é apresentada a motivação física para buscas por simetria em cenários comprimidos. No capítulo 3, é realizada a descrição do aparato experimental e aspectos de reconstrução do CMS. No capítulo 4, descreve-se a estratégia da análise, as características dos eventos de sinal e a modelagem das simulações de *background* dominantes. No capítulo 5, é realizada uma introdução à análise multivariada e redes neurais artificiais, posteriormente são apresentados os resultados das análises multivariadas na definição de regiões estratégicas da busca. No capítulo 6, são apresentados os resultados da análise em seu estágio atual. No capítulo 7, são discutidas as conclusões e são relatados os próximos passos da análise.

2 Motivação Física

2.1 Modelo Padrão da Física de Partículas

O Modelo Padrão da física de partículas é uma teoria quântica de campos que descreve as propriedades das partículas fundamentais e suas interações. Os campos são estruturas matemáticas definidas em todos os pontos do espaço, as quais carregam propriedades de simetria ao serem submetidos a determinados grupos de transformações, por exemplo rotações. As partículas são entendidas como estados excitados desses campos, logo, compartilham das mesmas propriedades atribuídas a eles. Ao detectarmos tais partículas nos experimentos, podemos obter informações dessas propriedades.

O MP descreve bem como três das quatro forças fundamentais da natureza (forte, fraca e eletromagnética) interagem. Ele é baseado em duas teorias quânticas de campos: Cromodinâmica Quântica (*Quantum chromodynamics* - QCD) e Dinâmica Quântica de Sabor (*Quantum flavordynamics* - QFD). A QCD descreve as interações das partículas carregadas de cor através da força forte e a QFD descreve a teoria das interações eletrofracas, que após o fenômeno conhecido com Quebra Espontânea de Simetria (QES) se desdobra nas forças fraca e eletromagnética. O Modelo Padrão, apesar de bem comprovado experimentalmente, não é uma teoria completa das interações fundamentais, principalmente porque não inclui a força da gravidade.

O Modelo Padrão foi construído ao longo de décadas de trabalho intenso e sua versão atual foi formulada em meados da década de 1970. Muitos renomados físicos de todo o mundo contribuíram para sua elaboração, e muitas dessas contribuições foram agraciadas com o prêmio Nobel. As sucessivas descobertas do quark top [4] em 1995, do neutrino do tau [5] em 2000 e do bóson de Higgs [6,7] em 2012 são alguns exemplos de seu grande poder preditivo.

2.1.1 O Espectro de Partículas

Algumas propriedades das partículas elementares são denotadas por números quânticos. Os números quânticos são propriedades quantizadas que podem ser usadas para

classificar partículas. Um número quântico importante é o “spin”, que é frequentemente descrito como o momento angular interno de uma partícula. As partículas representam flutuações positivas do campo, já as flutuações negativas são representadas por antipartículas que possuem os mesmos números quânticos, mas com sinal oposto. A massa das partículas e antipartículas deve ser a mesma porque corresponde a flutuações simetricamente opostas no campo.

Dentro do Modelo Padrão, o spin distingue duas classes de partículas, os férmions e os bósons. Férmions são partículas de spin semi-inteiro obedecendo a estatística de Fermi-Dirac e formam o que se denomina normalmente como matéria. Um sistema de partículas idênticas obedecendo a estatística de Fermi-Dirac é totalmente antissimétrico sob o intercâmbio de qualquer par de partículas. Todos os doze férmions contidos no MP têm spin $1/2$. Os férmions podem ser divididos em dois grupos com base no seu número quântico de carga de cor, os quarks e os léptons. A QCD governa a interação forte entre os quarks, já os léptons são singletos de cor e, portanto, não interagem com a força forte.

Bósons obedecem a estatística de Bose-Einstein em que um sistema de bósons idênticos é totalmente simétrico sob o intercâmbio de qualquer par. Bósons de *gauge* têm spin inteiro e são frequentemente apelidados de portadores de força porque mediam as diferentes interações descritas pelo MP. O fóton (γ) é o mediador da força eletromagnética, os bósons Z^0 , W^+ e W^- mediam a força fraca, enquanto os glúons (g) são responsáveis pela força forte. Por fim, temos o bóson de Higgs (H) que não é um bóson de *gauge* mas sim um subproduto do mecanismo responsável pela forma como as partículas adquirem massa (mecanismo de Higgs). O bóson de Higgs possui spin 0 enquanto os bósons de *gauge* possuem spin 1. Uma descrição mais detalhada das teorias de *gauge* responsáveis pelas interações no MP é apresentada na seção 2.1.2.

Os férmions são divididos em 3 gerações (I, II e III). Partículas de diferentes gerações compartilham propriedades idênticas, mas a massa das partículas aumenta em cada geração. Isso significa que partículas de gerações mais altas não são estáveis, podendo decair em partículas de gerações inferiores. Assim, toda matéria comum é composta de partículas de primeira geração. Uma geração de léptons consiste sempre de um lépton com carga elétrica negativa e seu correspondente neutrino eletricamente neutro. A primeira geração é constituída pelo elétron (e) e pelo neutrino do elétron (ν_e), a segunda geração pelo múon (μ) e pelo neutrino do múon (ν_μ) e a terceira geração contém o tau (τ) e o neutrino do tau (ν_τ). Uma geração de quarks consiste de um quark tipo-*up* e um tipo-*down* com carga elétrica fracionada de $+2/3$ e $-1/3$, respectivamente. A primeira geração consiste no quark *up* (u) e *down* (d), a segunda geração do quark *charm* (c) e do *strange* (s) e da terceira geração dos quarks *top* (t) e *bottom* (b).

No Modelo Padrão, supõe-se que os neutrinos não têm massa, embora isso tenha sido experimentalmente provado estar errado, é uma suposição conveniente, uma vez que, as

massas são muito pequenas e sua natureza exata é desconhecida [8]. É considerado que a verdadeira teoria dos neutrinos pode ser adicionada ao MP com poucas mudanças nas previsões experimentais dos colisores de partículas.

A figura 2.1 é um esquema do espectro de partículas do Modelo Padrão, mostrando a classificação das partículas e propriedades como massa, spin, carga elétrica e carga de cor. A carga elétrica está expressa em unidades absolutas da carga do elétron, a carga de cor pode assumir três valores para as partículas (vermelho, verde e azul) e três para as antipartículas (anti-vermelho, anti-verde e anti-azul). A massa de uma partícula corresponde à sua energia quando está em repouso. A unidade de massa no Sistema Internacional (SI) é o quilograma, kg. No entanto, o sistema natural de unidades é usado nesta tese, definindo a constante de Planck e velocidade da luz no vácuo como $\hbar = c = 1$. Ao fazer isso, massas, momentos e energias podem ser expressos em GeV.

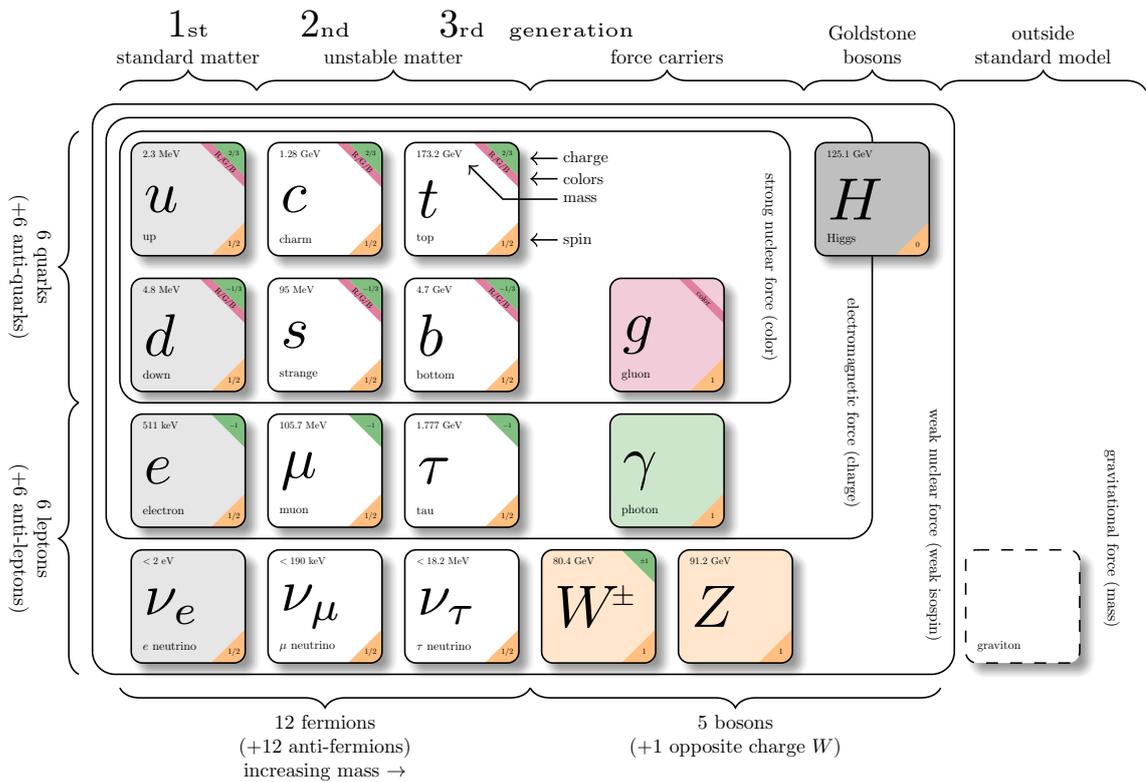


FIGURA 2.1 – Visão geral de todas as partículas fundamentais descritas pelo Modelo Padrão da física de partículas [9].

2.1.2 Simetrias e Interações

Simetrias estão diretamente ligadas à existência de quantidades conservadas na natureza, por exemplo a conservação de momento linear e de momento angular são consequên-

cia da invariância dos sistemas lagrangeanos sob as operações de translação e rotação, respectivamente. Além disso, as interações do MP podem ser derivadas matematicamente do princípio da invariância de *gauge* local dos bósons portadores de força. Simetrias locais são aquelas que dependem da posição enquanto as simetrias globais não dependem.

2.1.2.1 Simetria de Poincaré

Todos os campos físicos existentes na natureza, aos quais estão associadas as partículas elementares, possuem uma simetria espaço-temporal conhecida como simetria de Poincaré. A denominação espaço-temporal significa que as transformações de simetria que eles possuem são caracterizadas por transformações nas coordenadas espaciais e temporal. O grupo de Poincaré é composto pelas transformações de Lorentz, $\Lambda^\mu{}_\nu$, que englobam rotações e *boosts* na velocidade, mais as translações espaço-temporais, a^μ ,

$$x^\mu \rightarrow x'^\mu = \Lambda^\mu{}_\nu x^\nu + a^\mu, \quad (2.1)$$

sendo x^μ um ponto na variedade espaço-tempo.¹

A cada transformação do grupo de Poincaré é associado um operador unitário no espaço de Hilbert, $U(\Lambda, a)$, que atuará sobre os campos,

$$\psi(x) \rightarrow \psi'(x') = U(\Lambda, a)\psi(x). \quad (2.2)$$

A forma de $U(\Lambda, a)$ depende de como os campos se transformam perante o grupo de Poincaré, apresentando propriedades bem distintas nos casos em que os campos são de natureza fermiônica (spin semi-inteiro) ou bosônica (spin inteiro).

Podemos escrever uma transformação do grupo de Poincaré na forma infinitesimal, de modo a explicitar os geradores do grupo de Lorentz e das translações, respectivamente, $M^{\mu\nu}$ e P^μ ,

$$g(\Lambda, a) = I - \frac{i}{2}\omega_{\mu\nu}M^{\mu\nu} + ia_\mu P^\mu, \quad (2.3)$$

onde a_μ e o tensor antissimétrico $\omega_{\mu\nu}$ são parâmetros infinitesimais. O grupo de Poincaré é completamente definido pela álgebra que seus geradores obedecem, dada por²

$$\begin{aligned} [M_{\mu\nu}, M_{\rho\sigma}] &= -i(\eta_{\mu\rho}M_{\nu\sigma} - \eta_{\mu\sigma}M_{\nu\rho} - \eta_{\nu\rho}M_{\mu\sigma} + \eta_{\nu\sigma}M_{\mu\rho}), \\ [M_{\mu\nu}, P_\rho] &= -i(\eta_{\mu\rho}P_\nu - \eta_{\nu\rho}P_\mu), \\ [P_\mu, P_\nu] &= 0, \end{aligned} \quad (2.4)$$

¹ $x^\mu = (x^0, x^1, x^2, x^3)$, em que $x^0 = t, x^1 = x, x^2 = y, x^3 = z$ são as coordenadas espaço-temporais do espaço de Minkowski. Na equação 2.1 e no decorrer desta tese é adotada a notação de somatório de Einstein.

²O comutador de dois operadores A e B é definido como $[A, B] \equiv AB - BA$.

sendo $\eta_{\mu\nu}$ a métrica do espaço-tempo de Minkowski.

2.1.2.2 Teorias de gauge

Além das simetrias espaço-temporais, os campos físicos também possuem as simetrias de *gauge* (simetrias de calibre), que estão relacionadas às simetrias na estrutura interna dos campos. Esses grupos de simetria definem a forma como as partículas interagem entre si, e estão associados a uma classe de campos conhecida como campos de *gauge*, responsáveis por mediar as interações entre partículas. O MP descreve as interações dos campos físicos em termos da junção de três grupos de simetria, $SU(3)_C \times SU(2)_L \times U(1)_Y$. A força forte é descrita pelo grupo $SU(3)_C$ que age sobre a carga de cor C . Os grupos $SU(2)_L \times U(1)_Y$ denotam a força eletrofraca, o grupo $U(1)_Y$ atua na hipercarga fraca Y e $SU(2)_L$ atua somente em férmions com quiralidade³ *left-handed* (L).

Consideremos um campo representado por um multipletto, ψ , com N componentes, que se transforme por uma transformação de *gauge* dada por,

$$\psi' = e^{i\omega_a G_a} \psi, \quad (2.5)$$

sendo ω_a parâmetros constantes e G_a os geradores da simetria (matrizes $N \times N$). Caso o grupo de simetria possua M geradores não comutantes entre si, dizemos que essa é uma transformação de *gauge* não-abeliana global. A teoria de Yang-Mills [10], desenvolvida de maneira independente por Shaw [11], descreve como construir uma teoria invariante por transformações de *gauge* no caso em que os parâmetros da simetria sejam funções das coordenadas espaço-temporais, denominadas de transformações de *gauge* não-abelianas locais,

$$\psi' = e^{i\omega_a(x) G_a} \psi \equiv S\psi, \quad (2.6)$$

com $\omega_a(x)$ sendo parâmetros dependentes das coordenadas. Dois elementos fundamentais na construção de uma teoria não-abeliana invariante de *gauge* local são a derivada covariante e o tensor intensidade de campo, descritos a seguir.

- **Derivada covariante**

Para construir uma teoria invariante por transformações de *gauge* locais modifica-se o termo cinético da lagrangeana dos campos de matéria. Os termos cinéticos de teorias fermiônicas e bosônicas têm, respectivamente, as formas $\bar{\psi}\gamma^\mu\partial_\mu\psi$ e $\partial_\mu\phi^\dagger\partial^\mu\phi$.

³A quiralidade é uma propriedade da partícula que determina sob qual de duas representações distintas do grupo de Poincaré, *left-handed* ou *right-handed*, ela se transforma. No caso de partículas sem massa, a quiralidade é o mesmo que helicidade, de maneira que uma partícula *right-handed* tem sempre seu spin alinhado no mesmo sentido do movimento e uma partícula *left-handed* tem seu spin alinhado no sentido oposto.

Substitui-se a derivada normal, ∂_μ , pela derivada covariante, D_μ , definida como,

$$D_\mu \equiv \partial_\mu - ig\mathbf{B}_\mu, \quad (2.7)$$

com $\mathbf{B}_\mu \equiv B_{a\mu}G_a$ e g sendo uma constante de acoplamento, de modo que são introduzidos M campos de gauge, $B_{a\mu}$, na teoria. Os termos cinéticos ficam invariantes por transformações de *gauge* quando a derivada covariante do campo se transforma igual ao campo em si, sob transformações de gauge, de modo que

$$D'_\mu \psi' = S D_\mu \psi. \quad (2.8)$$

Conseqüentemente a transformação de *gauge* que o campo \mathbf{B}_μ deve obedecer para que a relação (2.8) seja satisfeita é dada por

$$\mathbf{B}'_\mu = S \mathbf{B}_\mu S^{-1} - \frac{i}{g} (\partial_\mu S) S^{-1}. \quad (2.9)$$

- **Tensor intensidade de campo**

O termo cinético dos campos de *gauge* é construído em termos de um tensor intensidade de campo, $\mathbf{W}_{\mu\nu} \equiv W_{a\mu\nu}G_a$ que deve se transformar como $\mathbf{W}'_{\mu\nu} = S \mathbf{W}_{\mu\nu} S^{-1}$ para garantir a invariância de *gauge* do termo cinético. A forma mais simples que satisfaz tal requerimento é

$$\mathbf{W}_{\mu\nu} \equiv \partial_\mu \mathbf{B}_\nu - \partial_\nu \mathbf{B}_\mu - ig[\mathbf{B}_\mu, \mathbf{B}_\nu]. \quad (2.10)$$

O termo cinético do campo de *gauge* será então proporcional ao traço $\text{Tr}(\mathbf{W}^{\mu\nu}\mathbf{W}_{\mu\nu})$ para que seja invariante de gauge. Tem-se $\text{Tr}(\mathbf{W}^{\mu\nu}\mathbf{W}_{\mu\nu}) = cW_{a\mu\nu}W_a^{\mu\nu}$, de maneira que c é uma constante que depende do grupo de simetria e $W_{a\mu\nu} = \partial_\mu B_{a\nu} - \partial_\nu B_{a\mu} + g f_{abc}B_{b\mu}B_{c\nu}$, com f_{abc} sendo a constante de estrutura do grupo, que no caso do SU(2) é igual ao tensor Levi-Civita, ϵ_{abc} , e no U(1) é igual a zero.

2.1.2.3 A Teoria Eletrofraca

O modelo de Glashow-Weinberg-Salam (GWS) é construído como uma teoria de *gauge* associada ao grupo de simetria $SU(2)_L \times U(1)_Y$, descrevendo a interação eletrofraca no MP [12,13,14]. A teoria eletrofraca atua de maneira distinta nos férmions com quiralidade *left-handed* e *right-handed* devido à não-conservação de paridade nas interações fracas [15,16]. Pode-se decompor um campo espinorial, ψ , nas suas componentes *left-handed*,

ψ_L , e *right-handed*, ψ_R , através dos operadores de quiralidade P_R e P_L ,

$$\begin{aligned}\psi_R &= P_R\psi = \frac{1}{2}(1 + \gamma^5)\psi, \\ \psi_L &= P_L\psi = \frac{1}{2}(1 - \gamma^5)\psi,\end{aligned}\tag{2.11}$$

onde γ^5 é o produto das quatro matrizes de Dirac. No MP, os férmions aparecem em famílias de dubletos *left-handed* de quarks Q_L^i e léptons L_L^i e singletos *right-handed* de quarks u_R^i (d_R^i) e léptons e_R^i (ν_R^i):

$$\begin{aligned}Q_L^i &= \begin{pmatrix} u \\ d \end{pmatrix}_L, \begin{pmatrix} c \\ t \end{pmatrix}_L, \begin{pmatrix} s \\ b \end{pmatrix}_L; & L_L^i &= \begin{pmatrix} \nu_e \\ e \end{pmatrix}_L, \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}_L, \begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}_L; \\ u_R^i &= u_R, c_R, t_R; & d_R^i &= d_R, s_R, b_R; \\ e_R^i &= e_R, \mu_R, \tau_R; & \nu_R^i &= \nu_R^e, \nu_R^\mu, \nu_R^\tau.\end{aligned}\tag{2.12}$$

O índice $i = 1, 2, 3$ denota a geração das partículas. Os campos *left-handed* são agrupados em dubletos de isospin ($T = 1/2$) enquanto os campos *right-handed* formam singletos de isospin ($T = 0$) e são invariantes sob transformações fracas de isospin. Portanto sob $SU(2)_L$, os campos *left-handed* transformam-se como dubletos, enquanto os *right-handed* não se transformam.

Seguindo a prescrição das teorias de Yang-Mills, a teoria eletrofraca, sendo uma teoria invariante de *gauge* localmente, é construída introduzindo derivadas covariantes com os campos de *gauge*, as quais serão distintas nos termos *left-handed* e *right-handed*, visto que o grupo de simetria $SU(2)_L$ atua somente no setor *left-handed*,

$$\partial_\mu\psi_L \rightarrow D_\mu\psi_L = (\partial_\mu - ig\mathbf{W}_\mu - ig'YB_\mu)\psi_L\tag{2.13}$$

$$\partial_\mu\psi_R \rightarrow D_\mu\psi_R = (\partial_\mu - ig'YB_\mu)\psi_R\tag{2.14}$$

sendo \mathbf{W}_μ e B_μ os campos de *gauge* introduzidos na teoria e g e g' as respectivas constantes de acoplamento. Os campos de *gauge* introduzidos não têm massa e possuem quatro modos de propagação, ou seja, inserem quatro partículas transportadoras sem massa, duas carregadas eletricamente e duas neutras, para mediar a interação eletrofraca. No entanto, o curto alcance da força fraca indica que ela é mediada por partículas massivas. Isso implica que a simetria da teoria é quebrada por algum mecanismo que dá massa às partículas mediadoras das interações fracas, os três bósons vetoriais W^+ , W^- e Z^0 , mas não para os fótons trocados nas interações eletromagnéticas. As provas experimentais [17,18,19,20] da existência de transportadores da força eletrofraca apoiaram a teoria unificada de interações fracas e eletromagnéticas, visto que as massas dos bósons estavam de acordo com seus valores previstos. Após a quebra da simetria eletrofraca, a simetria $U(1)_{EM}$ da eletrodinâmica aparece como uma simetria remanescente, de maneira que a

carga elétrica, Q , pode ser expressa como a soma de um dos geradores do grupo $SU(2)$ (o terceiro componente do isospin fraco T_3) e a hipercarga fraca como: $Q = (T_3 + Y/2)$. O mecanismo através do qual é quebrada a simetria eletrofraca é descrito na subseção seguinte.

2.1.2.4 O Mecanismo de Higgs

Na natureza, os bósons de *gauge* da interação fraca são observados como massivos e, portanto, um mecanismo explícito para quebrar a simetria eletrofraca é introduzido através do mecanismo de Higgs proposto em 1964 quase que simultaneamente por Peter Higgs e outros dois grupos independentes [21,22,23,24]. O mecanismo propõe que antes da quebra de simetria eletrofraca, a lagrangeana do modelo é escrita como

$$\mathcal{L} = (D_\mu \phi)^\dagger D^\mu \phi - V(\phi) - \frac{1}{4} W_{a\mu\nu} W_a^{\mu\nu} - \frac{1}{4} F_{\mu\nu} F^{\mu\nu}, \quad (2.15)$$

onde $W_{a\mu\nu} \equiv \partial_\mu W_{a\nu} - \partial_\nu W_{a\mu} + g\epsilon_{abc} W_{b\mu} W_{c\nu}$ é o tensor intensidade de campo relativo à simetria $SU(2)_L$ e $F_{\mu\nu} \equiv \partial_\mu B_\nu - \partial_\nu B_\mu$ é o relativo à simetria $U(1)_Y$. O campo, ϕ , é um dubleto de isospin fraco composto por dois campos complexos, logo, possui quatro modos de propagação,

$$\phi = \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix}, \quad (2.16)$$

onde ϕ_1, ϕ_2, ϕ_3 e ϕ_4 são campos reais.

A energia potencial (ou apenas “o potencial”) do modelo é dada por

$$V(\phi) = \mu^2 \phi^\dagger \phi + \lambda (\phi^\dagger \phi)^2, \quad (2.17)$$

tal que μ e λ são parâmetros que podem variar seus valores ao longo do tempo. O primeiro e o segundo termo são, respectivamente, quadrático e quártico em ϕ_i . O parâmetro λ é considerado sempre positivo e a quebra espontânea de simetria (*spontaneous symmetry breaking* - SSB) acontece quando μ^2 , considerado inicialmente positivo, adquire um valor negativo em determinado momento, deslocando o mínimo do potencial para uma das infinitas configurações de vácuo do potencial, figura 2.2, denotadas por ϕ_0 , que satisfazem a relação,

$$\phi_0^\dagger \phi_0 \equiv \nu^2 = \frac{-\mu^2}{2\lambda}. \quad (2.18)$$

Por definição ν^2 é positivo, logo, ν é um número real.

Devido à liberdade existente na parametrização do campo, ϕ , e sua invariância por transformações de *gauge* locais, independentemente da configuração ϕ_0 que o vácuo adquira após a quebra de simetria, pode-se sempre escolher uma parametrização apropriada,

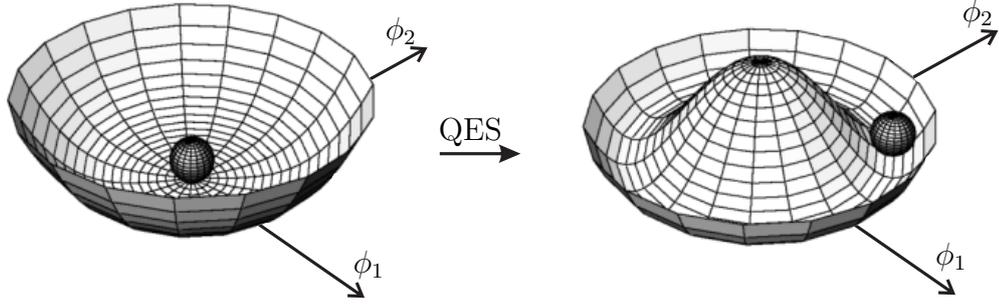


FIGURA 2.2 – Potencial antes da quebra de simetria, $\nu^2 = 0$ (esquerda) e potencial após a quebra espontânea de simetria, $\nu^2 = -\mu^2/2\lambda$ (direita).

$$\Phi = e^{i\xi_a G_a/\nu} \begin{pmatrix} 0 \\ H + \nu \end{pmatrix}, \quad (2.19)$$

em que o campo Φ fique descrito em termos dos campos reais ξ_1 , ξ_2 , ξ_3 e H . Assim, ao se aplicar uma transformação de *gauge* local, $\Omega = e^{i\theta_a G_a}$, em Φ e \mathbf{W}_μ , tal que $\theta_1(x) = -\xi_1(x)/\nu$, $\theta_2(x) = -\xi_2(x)/\nu$ e $\theta_3(x) = -\xi_3(x)/\nu$, os campos após a transformação de *gauge* ficam,

$$\Phi' = \Omega\Phi = \begin{pmatrix} 0 \\ H + \nu \end{pmatrix} \quad (2.20a)$$

$$\mathbf{W}'_\mu = \Omega \mathbf{W}_\mu \Omega^{-1} - \frac{1}{g}(\partial_\mu \Omega)\Omega^{-1}. \quad (2.20b)$$

Com isso, os três modos de propagação, relativos aos campos ξ_a , do dubleto escalar são transferidos para os bósons de *gauge*, \mathbf{W}'_μ . O modo de propagação remanescente no dubleto escalar, H , é conhecido como campo de Higgs. Usando as definições em (2.20) obtém-se,

$$V(\Phi') = -2\mu^2 H^2 + 4\lambda\nu H^3 + \lambda H^4 + \mu^2 \nu^2 + \lambda \nu^4, \quad (2.21)$$

e

$$(D'_\mu \Phi')^\dagger D'^\mu \Phi' = \partial_\mu H \partial^\mu H + g^2 (H^2 + 2\nu H + \nu^2) \left(\frac{g^2}{4} W_\mu^+ W^{-\mu} + \frac{g^2 + g'^2}{4} Z_\mu Z^\mu \right), \quad (2.22)$$

onde os campos de *gauge* Z^μ , $W^{+\mu}$ e $W^{-\mu}$, que são os bósons vetoriais da teoria eletrofraca, são definidos como

$$Z^\mu \equiv \cos \theta_W W_3^\mu - \sin \theta_W B^\mu \quad (2.23a)$$

$$W^{\pm\mu} \equiv \frac{1}{\sqrt{2}}(W_1^\mu \pm iW_2^\mu), \quad (2.23b)$$

com $\cos \theta_W \equiv g/\sqrt{g^2 + g'^2}$ e $\tan \theta_W \equiv g'/g$, sendo θ_W conhecido como ângulo de Wein-

berg. A lagrangeana do modelo escrita em termos do campo, H , não é mais explicitamente invariante sob as transformações do grupo de simetria $SU(2)_L \times U(1)_Y$, logo, é dito que o sistema sofreu uma quebra espontânea de simetria.

A lagrangeana, sem os termos constantes, que descreverá a interação entre o campo escalar H , o famoso bóson de Higgs, e os bósons vetoriais assume a forma

$$\begin{aligned} \mathcal{L} = & \partial_\mu H \partial^\mu H + 2\mu^2 H^2 - 4\lambda\nu H^3 - \lambda H^4 - \frac{1}{4} A_{\mu\nu} A^{\mu\nu} - \frac{1}{4} Z_{\mu\nu} Z^{\mu\nu} + \frac{g^2 + g'^2}{4} \nu^2 Z_\mu Z^\mu + \\ & - \frac{1}{2} W_{\mu\nu}^+ W^{-\mu\nu} + \frac{g^2}{2} \nu^2 W_\mu^+ W^{-\mu} + \mathcal{L}_I^{\text{BH}} + \mathcal{L}_I^{\text{BB}}, \end{aligned} \quad (2.24)$$

sendo

$$\mathcal{L}_I^{\text{BH}} = (2\nu H + H^2) \left(\frac{g^2}{2} W_\mu^+ W^{-\mu} + \frac{g^2 + g'^2}{4} Z_\mu Z^\mu \right), \quad (2.25)$$

e

$$\begin{aligned} \mathcal{L}_I^{\text{BB}} = & -\frac{ig}{2} (\cos\theta_W Z_\nu + \sin\theta_W A_\nu) [(\partial^\mu W^{+\nu} - \partial^\nu W^{+\mu}) W_\mu^- + (\partial^\mu W^{-\nu} - \partial^\nu W^{-\mu}) W_\mu^+] + \\ & -\frac{ig}{2} (W_\mu^+ W_\nu^- - W_\mu^- W_\nu^+) (\cos\theta_W \partial^\mu Z^\nu + \sin\theta_W \partial^\mu A^\nu) + \\ & -\frac{g^2}{2} W_\mu^- W_\nu^+ (W^{-\mu} W^{+\nu} - W^{+\mu} W^{-\nu}) + \\ & -2g^2 \sin^2\theta_W (W_\mu^- W^{+\nu} A^\mu A_\nu - W_\mu^- W^{+\mu} A^\nu A_\nu) + \\ & -2g^2 \cos^2\theta_W (W_\mu^- W^{+\nu} Z^\mu Z_\nu - W_\mu^- W^{+\mu} Z^\nu Z_\nu) + \\ & -2g^2 \sin\theta_W \cos\theta_W [W_\mu^- W^{+\nu} (A^\mu Z_\nu + Z^\mu A_\nu) - W_\mu^- W^{+\mu} Z_\nu A^\nu]. \end{aligned} \quad (2.26)$$

O novo campo de *gauge* sem massa, A^μ , que aparece na lagrangeana é definido como,

$$A^\mu \equiv \sin\theta_W W_\mu^{\prime 3} + \cos\theta_W B^\mu, \quad (2.27)$$

e nada mais é que o campo eletromagnético. Os tensores intensidade de campo são definidos como $A_{\mu\nu} \equiv \partial_\mu A_\nu - \partial_\nu A_\mu$, $Z_{\mu\nu} \equiv \partial_\mu Z_\nu - \partial_\nu Z_\mu$ e $W_{\mu\nu}^\pm \equiv \partial_\mu W_\nu^\pm - \partial_\nu W_\mu^\pm$.

Após a quebra de simetria observamos a existência de um campo escalar massivo, H , três campos vetoriais massivos, Z^μ , $W^{+\mu}$ e $W^{-\mu}$, e um campo vetorial sem massa, A^μ . Os campos ξ_1 , ξ_2 e ξ_3 , colocados dentro da definição de \mathbf{W}'_μ , resultam nos três modos de propagação adquiridos pelos bósons vetoriais massivos.

2.1.2.5 Acoplamentos de Yukawa

O fato do grupo $SU(2)_L$ atuar apenas nas componentes *left-handed* dos campos fermiônicos cria um problema fundamental com relação às massas dessas partículas. O termo

de massa para um férmion de Dirac $m(\bar{\psi}_L\psi_R + \bar{\psi}_R\psi_L)$ acopla os componentes *left-handed* e *right-handed*, portanto, não é invariante sob $SU(2)_L$. No entanto, em uma teoria com QES, há uma maneira de dar massas aos férmions através do chamado acoplamento Yukawa (*Yukawa coupling*) a um campo escalar. No MP, a lagrangeana renormalizável das interações de Yukawa que descreve o acoplamento entre o campo de Higgs e os campos de quarks e léptons, possui a forma

$$\mathcal{L}_{\text{Yukawa}} = -g_i \bar{L}_L^i \Phi e_R^i - \Gamma_{ij}^d \bar{Q}_L^i \Phi d_R^j - \Gamma_{ij}^u \bar{Q}_L^i \tilde{\Phi} u_R^j + \text{h.c.}, \quad (2.28)$$

sendo g_i , Γ_{ij}^d e Γ_{ij}^u as constantes de acoplamento, e $\tilde{\Phi} = i\sigma_2\Phi^*$ o conjugado de carga do campo de Higgs, com σ_2 denotando a segunda matriz de Pauli. O expressão “h.c.” representa os conjugados hermitianos dos termos expostos.

Considerando, como exemplo, o acoplamento entre o dubleto de elétrons L_L^e , o dubleto de Higgs Φ e a componente *right-handed* do campo do elétron e_R , a lagrangeana efetiva de Yukawa após a quebra espontânea de simetria torna-se:

$$\mathcal{L}_{\text{Yukawa}}^e = -\frac{g_e \nu}{\sqrt{2}}(\bar{e}_L e_R + \bar{e}_R e_L) - \frac{g_e}{\sqrt{2}}(\bar{e}_L e_R + \bar{e}_R e_L)H. \quad (2.29)$$

O primeiro termo é identificado como o termo de massa dos elétrons ($m_e \equiv g_e \nu / \sqrt{2}$), enquanto o segundo termo descreve a interação entre o campo de Higgs e o par elétron-pósitron. O mesmo mecanismo é considerado para gerar a massa dos outros léptons carregados. Por tanto, os férmions adquirem massa dentro do MP através das interações de Yukawa entre o campo de Higgs e os férmions. Uma consequência direta do mecanismo QES é que a intensidade do acoplamento g_f do campo de Higgs com um férmion massivo é proporcional à sua massa $g_f = \sqrt{2}m_f/\nu$.

2.1.2.6 A Cromodinâmica Quântica

A Cromodinâmica Quântica (*Quantum Chromodynamics* - QCD) é a teoria renormalizável que descreve as interações fortes entre quarks e glúons, ou seja, entre partículas que possuem cargas de cor. Ela é uma teoria de *gauge* conhecida por $SU(3)_C$ descrita pela lagrangeana,

$$\mathcal{L}_{\text{QCD}} = -\frac{1}{4}G_{a\mu\nu}G_a^{\mu\nu} + \sum_f \bar{q}_f(i\gamma^\mu D_\mu - m_f)q_f, \quad (2.30)$$

tal que $D_\mu = \partial_\mu - ig_s \frac{\lambda_a}{2} G_{\mu a}$ é a derivada covariante que por simplicidade está considerando somente a interação dos quarks, q_f , com os campos de *gauge* de $SU(3)_C$, mas como sabemos da subseção anterior, eles também interagem segundo o grupo $SU(2)_L \times SU(1)_Y$. Temos que g_s é a constante de acoplamento forte e as massas dos quarks na soma dos

sabores, f , são adicionadas a mão e são oriundas dos acoplamentos de Yukawa. O símbolo $G_{\mu\nu}^a$ representa o tensor intensidade de campo gluônico dado por,

$$G_{a\mu\nu} = \partial_\mu G_{a\nu} - \partial_\nu G_{a\mu} + g_s f_{abc} G_{b\mu} G_{c\nu}, \quad (2.31)$$

com $G_{a\mu}$ correspondendo os 8 campos bosônicos sem massa dos glúons, que são associados aos 8 geradores do grupo representados pelas matrizes de Gell-Mann, $\lambda_a (a \in [1, 8])$. Por fim, f_{abc} são as constantes de estrutura do grupo $SU(3)$.

Sabendo-se que cada um dos seis sabores de quarks se manifesta em três diferentes cores $i = 1, 2, 3$, pode-se representar o estado de um quark no espaço de cores como

$$q_f = \begin{pmatrix} q_f^1 \\ q_f^2 \\ q_f^3 \end{pmatrix} \quad (2.32)$$

em que cada componente q_f^i é um espinor de Dirac. Devido à natureza não-abeliana de $SU(3)_C$, os glúons também são coloridos e, portanto, têm auto-interação. Isso pode ser visto na forma explícita dos termos com 3 e 4 glúons oriundos do primeiro termo da lagrangeana (2.30),

$$\mathcal{L}_{\text{QCD}}^{\text{auto-interação}} = -\frac{g_s}{2} f_{abc} (\partial^\mu G_a^\nu - \partial^\nu G_a^\mu) G_\mu^b G_\nu^c - \frac{g_s^2}{4} f_{abc} f_{ade} G_b^\mu G_c^\nu G_{\mu d} G_{\nu e} \quad (2.33)$$

Na natureza, os quarks estão sempre agrupados formando os chamados hádrons. O hádron é uma partícula não elementar que possui carga de cor zero. Ele é formado na maioria das vezes pela composição de um quark e um antiquark (cargas de cor opostas), ou pela composição de três quarks com cargas de cor distintas, dando origem, respectivamente, aos mésons e aos bárions. Contudo, combinações com mais quarks também são permitidas, por exemplo quatro quarks e um antiquark podem formar um estado ligado chamado “pentaquark”, os dois primeiros pentaquarks [25] foram descobertos em 2015 pela colaboração LHCb (*Large Hadron Collider beauty*) [26]. O fato dos quarks só serem observados formando hádrons é chamado de confinamento de cor, o qual ainda não possui uma explicação analítica. À medida que duas cargas de cor são separadas, em algum ponto se torna energeticamente favorável à produção de um par quark-antiquark, nesse processo de formação de pares os quarks se reagrupam formando partículas neutras de cor. Como consequência disso, quando quarks são produzidos nos aceleradores de partículas, invés de se observar quarks individuais nos detectores, são observados jatos de partículas neutras de cor, oriundas desse processo conhecido como hadronização.

Apesar de confinados no interior dos hádrons por um acoplamento forte muito intenso, os quarks, assintoticamente, com o aumento da energia, se comportam como partículas

livres. Desse modo, em altas energias é possível desenvolver a QCD de forma perturbativa, análoga à Eletrodinâmica Quântica (*Quantum Electrodynamics* - QED).

2.1.3 Questões Abertas no Modelo Padrão

Embora o MP se mostre uma teoria extremamente bem sucedida e útil, sabe-se que é uma teoria incompleta como sugerido por uma variedade de observações. Alguns dos fatos experimentais mais importantes são:

- **Gravidade:** o MP não descreve a gravidade. Até agora, nenhuma descrição consistente da quantização da interação gravitacional foi formulada.
- **Matéria Escura e Energia Escura:** uma variedade de observações astronômicas provam que apenas cerca de 5% do conteúdo energético do universo pode ser explicado pela matéria ordinária descrita pelo MP [27, 28]. Cerca de 27% consiste em um tipo desconhecido de matéria, referida como Matéria Escura, que só interage gravitacionalmente e talvez fracamente. Evidências de matéria escura podem ser deduzidas diretamente da curva de rotação de galáxias [29], efeitos de lentes gravitacionais [30, 31] ou anisotropias na radiação cósmica de fundo [32], bem como indiretamente a partir do desenvolvimento de processos como a formação de estruturas no universo primordial [33]. O conteúdo de energia restante no universo é um tipo desconhecido de energia, a chamada Energia Escura. Essa energia é responsável por efeitos como a atual expansão acelerada do universo.

Do ponto de vista teórico, o MP também tem algumas deficiências. Entre elas podemos citar:

- **Problema de Hierarquia:** As massas das partículas descritas pelo MP recebem correções radiativas que surgem de diagramas de Feynman de ordem superior com férmions ou bósons virtuais. Em relação ao bóson de Higgs, que é um escalar fundamental que não se origina de uma simetria de *gauge*, essas chamadas contribuições de *loop* resultam em grandes correções. Visto que os acoplamentos do bóson de Higgs às partículas no *loop* são proporcionais à sua massa, as correções principais surgem dos *loops* com quarks top. O diagrama de Feynman correspondente a um *loop* é mostrado na figura 2.3.

Somando todas as contribuições existentes a um *loop*, a massa observável do bóson de Higgs, m_H^{obs} , [34] é dada por:

$$(m_H^{\text{obs}})^2 = (m_H^{\text{bare}})^2 + c \cdot \Lambda^2 + \mathcal{O}(\log(\Lambda^2)), \quad \text{with } c \in \mathbb{R}. \quad (2.34)$$

A massa nua do bóson de Higgs, m_H^{bare} , é um parâmetro livre da teoria enquanto o parâmetro de corte Λ é necessário para avaliar as integrais divergentes (“regularização”). O parâmetro de corte é frequentemente interpretado como a escala em que a teoria não é mais válida. Este é o caso da escala de Planck $\Lambda_P \approx 10^{19}$ GeV, em que a força da gravidade torna-se comparável às outras interações fundamentais. Supondo que não ocorra a manifestação de nenhum efeito físico desconhecido nessa escala, a massa nua do bóson de Higgs teria que estar perto da escala de Planck. Na verdade, a massa nua teria que cancelar as correções enormes do *loop* quase que perfeitamente, para que a massa do bóson de Higgs fosse observada com 125 GeV em baixas escalas de energia. Esse procedimento é referido como ajuste fino e é considerado não natural.

- **Unificação das forças:** a unificação das forças eletromagnéticas e fracas sugere que pode haver uma teoria única, uma Teoria da Grande Unificação (GUT), que pode descrever as três forças do MP. Para conseguir isso, espera-se que os valores das constantes de acoplamento das interações convirjam em alguma escala alta de energia, e isso não é observado no MP.

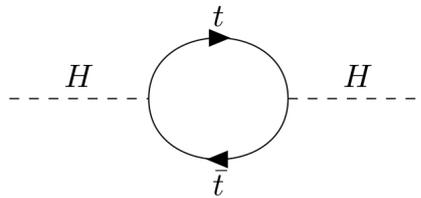


FIGURA 2.3 – Diagrama de Feynman da correção quântica dominante para o massa de Higgs proveniente de *loops* do quark top.

2.2 Supersimetria

O teorema de Coleman e Mandula [35], proposto em 1967, mostrou que ao exigir certas propriedades razoáveis de amplitudes de espalhamento, restringia-se misturas não triviais entre o espaço-tempo e simetrias internas. Contudo, uma extensão desse teorema proposta por Haag et al. [36], mostrou que é possível inserir uma simetria fermiônica, anticomutante, que se mistura com simetrias espaço-temporais e que ainda satisfaz todos os requerimentos razoáveis para amplitudes de espalhamento. Esse tipo de simetria foi chamada de “Supersimetria” (*supersymmetry* - SUSY). A estrutura algébrica da Supersimetria é uma extensão das álgebras de Lie, chamada álgebra de Lie graduada, na qual há elementos que dizemos ser de natureza par, os quais obedecem relações de comutação (álgebra de Lie), enquanto os de natureza ímpar obedecem relações de anti-comutação

entre eles e relações de comutação com os de natureza par. No contexto de SUSY, os elementos de natureza par são os geradores do grupo de Poincaré e os de natureza ímpar são os geradores de SUSY. Uma transformação supersimétrica transforma um estado bosônico em um estado fermiônico e vice-versa, com o gerador dado pelo operador espinorial anticomutante, Q , de modo que,

$$\text{férmion} \leftarrow Q \rightarrow \text{bóson}. \quad (2.35)$$

A princípio poderíamos construir uma álgebra com N geradores de SUSY, porém apenas teorias com uma carga espinorial Q_α , conhecidas como N=1 SUSY, permitem férmions quirais [37], ou seja, férmions em que suas partes *left-handed* e *right-handed* transformam-se diferentemente sob transformações de simetria. A Supersimetria do tipo N=1 é a extensão supersimétrica mínima da álgebra (2.4), de maneira que

$$\begin{aligned} [P_\mu, Q_\alpha] &= 0 \\ [Q_\alpha, M^{\mu\nu}] &= \frac{1}{2}(\sigma^{\mu\nu})_\alpha^\beta Q_\beta, \\ \{Q_\alpha, \bar{Q}_\beta\} &= 2(\gamma^\mu)_{\alpha\beta} P_\mu. \end{aligned} \quad (2.36)$$

O operador P^2 , sendo P o gerador das translações, comuta com todos os geradores, de modo que todas as partículas presentes em um supermultiplete (supercampo), representação irredutível da álgebra de SUSY, terão os mesmos autovalores de P^2 , portanto, a mesma massa. Os geradores de Supersimetria também comutam com os geradores das transformações de *gauge*, o que significa que os membros de um supermultiplete pertencem à mesma representação do grupo de *gauge*, por isso, possuem a mesma carga elétrica, isospin fraco e graus de liberdade de cor. O supercampo seria composto por uma componente bosônica e uma fermiônica. Logo, a Supersimetria propõe que para cada partícula bosônica do MP, existe uma partícula supersimétrica fermiônica com as mesmas características e vice-versa.

A extensão supersimétrica do espaço-tempo é chamada de superespaço e é parametrizada pelas coordenadas do espaço tempo e duas coordenadas espinoriais θ_α e $\bar{\theta}_{\dot{\alpha}}$. Os índices, α e $\dot{\alpha}$, são referentes à representação de $SU(2)$ do espinor. Uma transformação geral de SUSY, com parâmetros $(a^\mu, \xi, \bar{\xi})$, atuando nas coordenadas $(x^\mu, \theta, \bar{\theta})$ correspondem às transformações,

$$\begin{aligned} x'^\mu &= x^\mu + a^\mu - i\xi\sigma^\mu\bar{\theta} + i\theta\sigma^\mu\bar{\xi}; \\ \theta' &= \theta + \xi, \quad \bar{\theta}' = \bar{\theta} + \bar{\xi}. \end{aligned} \quad (2.37)$$

Por sua vez, a transformação de um supercampo, $S(x^\mu, \theta, \bar{\theta})$, ocorre segundo a expressão

$$S'(x^\mu, \theta, \bar{\theta}) = \exp\{i(\xi Q + \bar{\xi}\bar{Q} - a^\mu P_\mu)\}S(x^\mu, \theta, \bar{\theta}). \quad (2.38)$$

Nos supermultipletos, os estados bosônicos e fermiônicos têm o mesmo número de graus de liberdade. As regras na construção de teorias de *gauge* N=1 SUSY consistem em atribuir um supercampo quirial para cada campo de matéria e um supercampo vetorial para cada campo de gauge. Os supercampos quirais *left-handed* e *right-handed* são definidos por

$$\bar{D}_L \Phi_L = 0, \quad D_R \Phi_R = 0, \quad (2.39)$$

onde

$$(D_L)_\alpha = \frac{\partial}{\partial \theta^\alpha} + 2i\sigma_{\alpha\dot{\beta}}^\mu \bar{\theta}^{\dot{\beta}} \partial_\mu, \quad (\bar{D}_L)_{\dot{\alpha}} = -\frac{\partial}{\partial \bar{\theta}^{\dot{\alpha}}}, \quad (2.40)$$

$$(\bar{D}_R)_{\dot{\alpha}} = \frac{\partial}{\partial \bar{\theta}^{\dot{\alpha}}} - 2i\sigma_{\dot{\alpha}\beta}^\mu \bar{\theta}^\beta \partial_\mu, \quad (D_R)_\alpha = \frac{\partial}{\partial \theta^\alpha}. \quad (2.41)$$

Nesta representação o supercampo quirial *left-handed* é independente de $\bar{\theta}$ enquanto o campo quirial *right-handed* é independente de θ . O supercampo pode ser escrito em termos de suas componentes como

$$\Phi_L = \phi(x) + \sqrt{2}\theta^\alpha \psi_\alpha(x) + \theta^\alpha \theta^\beta \epsilon_{\alpha\beta} F(x). \quad (2.42)$$

A componente $\phi(x)$ é um campo escalar complexo de spin 0, $\psi(x)$ é um campo fermiônico de Weyl, spin 1/2, e $F(x)$ é um campo escalar auxiliar. Assim, o conteúdo físico do supercampo corresponde aos dois modos físicos do campo escalar complexo, $\phi(x)$, e aos dois modos físicos do férmion de Weyl, $\psi(x)$. Sendo $\psi(x)$ um férmion do Modelo Padrão, o campo $\phi(x)$ seria seu superparceiro escalar, um sférmion. Os supercampos quirais também são usados para descrever a extensão supersimétrica do campo de Higgs, nesse caso a parte escalar é o campo de Higgs e $\psi(x)$ seria o seu parceiro supersimétrico, o higgsino.

O supercampo vetorial, satisfaz $V(x, \theta, \bar{\theta}) = V^\dagger(x, \theta, \bar{\theta})$, e no *gauge* de Wess-Zumino, a expansão do campo vetorial em componentes é dada por

$$V = i\theta\theta\bar{\theta}\bar{\lambda}(x) + i\bar{\theta}\bar{\theta}\theta\lambda(x) - \theta\sigma^\mu\bar{\theta}A_\mu(x) + \frac{1}{2}\theta\theta\bar{\theta}\bar{\theta}D(x), \quad (2.43)$$

onde $A(x)$ é o campo de *gauge* de spin 1 e λ seu superparceiro fermiônico de spin 1/2, o gaugino, já $D(x)$ é um campo auxiliar. Sendo $A(x)$ um bóson sem massa, ele possui dois modos físicos, e $\lambda(x)$ sendo um férmion de Weyl, também possui dois modos físicos.

A lagrangeana que compõe a generalização supersimétrica do MP é construída escre-

vendo todos os termos invariantes de *gauge* com dimensão de massa igual a quatro,

$$\begin{aligned} \mathcal{L}_{\text{SUSY}} = & \sum_i \int d^2\theta d^2\bar{\theta} \Phi_i^\dagger e^{g_3 \hat{V}_3 + g_2 \hat{V}_2 + g_1 \hat{V}_1} \Phi_i + \left(\int d^2\theta (\mathcal{W}(\Phi) + h.c.) \right) + \\ & + \sum_{U(1), SU(2), SU(3)} \frac{1}{4} \left(\int d^2\theta [W^\alpha W_\alpha] + h.c. \right). \end{aligned} \quad (2.44)$$

O primeiro termo corresponde aos termos de energia cinética dos férmions e sférmions, e a interação deles com os campos de *gauge* e os gauginos. O segundo termo contém todos os termos possíveis com respeito às interações de férmions quirais, onde \mathcal{W} é chamado de superpotencial. E o terceiro termo corresponde à parte cinética dos campos de *gauge* em termo do supercampo W^α definido como,

$$W_\alpha = -\frac{1}{4}(\bar{D}\bar{D})e^{-V}D_\alpha e^V. \quad (2.45)$$

2.2.1 Modelo Padrão Supersimétrico Mínimo

O Modelo Padrão Supersimétrico Mínimo (*Minimal Supersymmetric Standard Model* - MSSM) é uma extensão da teoria de *gauge* $SU(3)_C \times SU(2)_L \times U(1)_Y$ adicionando a $N=1$ SUSY e inserindo o conteúdo mínimo de supercampos necessários para gerar o espectro de partículas do MP [38,39,40]. Como nenhum dos férmions de matéria do MP se transforma como o adjunto do grupo de gauge, eles não podem ser identificados como parceiros supersimétricos dos bósons de gauge. Além do mais, os superparceiros dos campos do MP devem carregar os mesmos números quânticos globais e de gauge. Então, conclui-se que deve-se introduzir um supercampo para cada campo do MP. Nos supermultipletos da álgebra de SUSY utilizados na construção do MSSM, os parceiros escalares (spin-0) dos férmions são chamados sléptons e squarks, já os parceiros fermiônicos (spin 1/2) do bóson de Higgs são chamados de higgisinos, os dos bósons vetoriais são o bino- \tilde{B} e os winos- $\tilde{W}s$, e os dos glúons são os gluinos- \tilde{g} . O bóson de Higgs e os higgisinos estão contidos dentro de dois dubletos de campos escalares, os quais são necessários para manter a renormalizabilidade da teoria e gerar apropriadamente os termos de Yukawa que dão massa aos campos. Os superparceiros das partículas do MP são denotados com o uso do til acima do símbolo da partícula, por exemplo o superparceiro do elétron, e , é denotado por \tilde{e} .

Entre os campos que constituem o MSSM, temos primeiramente os supercampos *left-handed* contendo os quarks e squarks, Q_i (i denota as gerações), e aqueles contendo os léptons e sléptons, L_i , sendo ambos dubletos de $SU(2)$. Os supercampos *right-handed* correspondentes, U_i e D_i , contêm, respectivamente, os quarks tipo-up e tipo-down e E_i contém os léptons carregados. Similarmente ao MP, não existe supercampo associado

aos neutrinos *right-handed*. Os bósons de *gauge* e os gauginos estão contidos dentro dos campos vetoriais, G para os glúons e gluinos, W para os bósons W e Winos e B para o bóson B e o Bino. Por fim, temos os dois dubletos, H_1 e H_2 que contêm os campos de Higgs e higgsinos.

Um superpotencial, que deverá conter os acoplamentos de Yukawa dentro dele, é construído com os multipletos de matéria, exigindo apenas que seja formado por todos os operadores invariantes de *gauge* de dimensão não maior que 3, ficando da seguinte forma

$$\begin{aligned} \mathcal{W} = & (\lambda_u^{ij} Q^i H_1 \bar{U}^j + \lambda_d^{ij} Q^i H_2 \bar{D}^j + \lambda_e^{ij} L^i H_2 \bar{E}^j + \mu H_2 H_1) + \\ & + (\alpha_1^{ijk} Q^i L^j \bar{D}^k + \alpha_2^{ijk} L^i L^j \bar{E}^k + \alpha_3^i L^i H_1 + \alpha_4^{ijk} \bar{D}^i \bar{D}^j \bar{U}^k). \end{aligned} \quad (2.46)$$

Os termos do segundo par de parêntesis quebram a conservação do número de bárions, B , e léptons, L , sendo assim, é requerida uma simetria adicional para livrar-se dessas interações que violam B e L , as quais são fenomenologicamente inaceitáveis. A maneira mais simples de se obter isso, é requerer que \mathcal{W} seja invariante por uma transformação de simetria chamada paridade-R, definida como $P_R \equiv (-1)^{3(B-L)+2S}$, de maneira que sob paridade-R os supercampos se transformam como

$$\begin{aligned} H_1, H_2 & \rightarrow H_1, H_2, \\ Q, \bar{U}, \bar{D}, L, \bar{E} & \rightarrow -(Q, \bar{U}, \bar{D}, L, \bar{E}), \end{aligned} \quad (2.47)$$

que implica em campos do Modelo Padrão sendo inalterados sob a transformação e seus parceiros supersimétricos mudando de sinal. Uma consequência da conservação da paridade-R é que partículas supersimétricas interagem apenas em pares, implicando em que as partículas supersimétricas mais leves (*lightest supersymmetric particle* - LSP) sejam estáveis se a paridade-R for exata. A tabela B.4 mostra como estão organizadas as partículas do MP e suas superparceiras no contexto do MSSM.

2.2.1.1 Quebra de SUSY e o espectro de partículas

Um ponto importante é que se a Supersimetria fosse uma simetria exata, as partículas supersimétricas teriam a mesma massa de suas parceiras do MP. Isso não acontece ou tais partículas já teriam sido detectadas, logo, a Supersimetria deve ser quebrada em algum momento. A quebra de SUSY deve ser suave (*soft*), ou seja, não deve reintroduzir divergências quadráticas na teoria. Isso é alcançado através de uma quebra de simetria espontânea. Como nenhum dos campos do MSSM pode desenvolver valores esperados não nulos no vácuo (vevs) para quebrar SUSY sem estragar a invariância de *gauge*, é esperado que a quebra espontânea de Supersimetria aconteça através de algum outro campo. O

Supermultipletos		Bósons	Férmions	$SU(3)_C$	$SU(2)_L$	$U(1)_Y$
Gauge/gaugino	G	g	\tilde{g}	8	1	0
	W	W^\pm, W^0	$\tilde{W}^\pm, \tilde{W}^0$	1	3	0
	B	B^0	\tilde{B}^0	1	1	0
Slépton/lépton	L	$(\tilde{\nu}_e, \tilde{e})_L$	$(\nu_e, e)_L$	1	2	-1
	\bar{E}	\tilde{e}_R^*	e_R^\dagger	1	1	-2
Squarks/quarks	Q	$(\tilde{u}, \tilde{d})_L$	$(u, d)_L$	3	2	1/3
	\bar{U}	\tilde{u}_R^*	u_R^\dagger	3	1	4/3
	\bar{D}	\tilde{d}_R^*	d_R^\dagger	3	1	-2/3
Higgs/higgsinos	H_1	(H_1^+, H_1^0)	$(\tilde{H}_1^+, \tilde{H}_1^0)$	1	2	1
	H_2	(H_2^0, H_2^-)	$(\tilde{H}_2^0, \tilde{H}_2^-)$	1	2	-1

TABELA 2.1 – Composição dos supercampos no MSSM.

cenário mais comum para propor a quebra de Supersimetria estabelece a existência de dois setores, o setor usual (“visível”) onde residem os campos do MSSM e um setor “escondido” que contém os campos que levam à quebra da Supersimetria. Os dois setores interagem via campos conhecidos como “mensageiros” que transmitem a quebra de SUSY para os campos do MSSM e geram as massas das partículas supersimétricas.

Os mecanismos de quebra de SUSY são bem diferentes mas produzem resultados comuns. Os termos de quebra suave de SUSY para o MSSM apresentam a forma

$$\begin{aligned}
-\mathcal{L}_{\text{quebra}} = & \frac{1}{2}M_1\tilde{B}\tilde{B} + \frac{1}{2}M_2\tilde{W}\tilde{W} + \frac{1}{2}M_3\tilde{g}\tilde{g} + \\
& + M_{H_1}^2|H_1|^2 + M_{H_2}^2|H_2|^2 + \sum_i M_{Q_i}^2|\tilde{q}_{L_i}|^2 + \sum_i M_{U_i}^2|\tilde{u}_{R_i}^c|^2 + \sum_i M_{D_i}^2|\tilde{d}_{R_i}^c|^2 + \\
& + \sum_i M_{L_i}^2|\tilde{l}_{L_i}|^2 + \sum_i M_{E_i}^2|\tilde{e}_{R_i}^c|^2 + \\
& + (\lambda_E A_E)_{ij}H_1\tilde{l}_{L_i}\tilde{e}_{R_j}^c + (\lambda_U A_U)_{ij}H_1\tilde{q}_{L_i}\tilde{u}_{R_j}^c + (\lambda_D A_D)_{ij}H_1\tilde{q}_{L_i}\tilde{d}_{R_j}^c + h.c. + \\
& + B\mu H_1 H_2 + h.c.. \tag{2.48}
\end{aligned}$$

Na primeira linha temos os termos de massa dos gauginos, na segunda e terceira linhas estão os termos de massa dos superparceiros escalares, na linha quatro estão os termos das interações trilineares e na última linha estão os termos das interações bilineares. Existem modelos nada triviais que tentam explicar como ocorre essa quebra de simetria. Três mecanismos bem conhecidos que tentam descrever como os superparceiros do MSSM adquirem massas são:

- Quebra de SUSY por mediação da gravidade (*Gravity Mediated SUSY Breaking* - SUGRA): nessa abordagem interações de Supergravidade⁴ são responsáveis por mediar a quebra de simetria. Espera-se que as interações mais gerais consistentes

⁴Supergravidade é uma teoria de campos que combina princípios de Supersimetria e Relatividade Geral.

com as simetrias dos setores visível e escondido sejam geradas em uma teoria efetiva com os termos de acoplamento sendo suprimidos pela escala de Planck, M_{Pl} .

- Quebra de SUSY por mediação de *gauge* (*Gauge Mediated SUSY Breaking* - GMSB): considera-se que a SUSY é quebrada em baixas escalas de energia de modo que contribuições de SUGRA para os termos suaves sejam desprezíveis. A comunicação da quebra de SUSY entre os setores visível e escondido ocorre apenas através de termos de interação de *gauge* do MP. Nesse cenário, o problema das correntes neutras com mudança de sabor (*flavor-changing neutral current* - FCNC) não aparece.
- Quebra de SUSY por anomalia (*Anomaly Mediation SUSY Breaking* - AMSB): nesse cenário, a supressão de correntes neutras com mudança de sabor ocorre sem a necessidade de supor a quebra de SUSY em baixas escalas de energia. As FCNC podem ser suprimidas se os potenciais da lagrangeana de interação⁵ dos supercampos quirais possuem a forma

$$\begin{aligned} K &= -3M_{\text{Pl}}^2 \ln \left(1 - \frac{f_{\text{vis}}}{3M_{\text{Pl}}^2} - \frac{f_{\text{hid}}}{3M_{\text{Pl}}^2} \right) \\ \mathcal{W} &= \mathcal{W}_{\text{vis}} + \mathcal{W}_{\text{hid}} + \mathcal{W}_0, \end{aligned} \quad (2.50)$$

onde “vis” e “hid” denotam, respectivamente, as contribuições do setor visível e escondido. A lagrangeana nessa forma, resulta no desaparecimento dos termos suaves a nível de árvore. Contudo, as massas dos gauginos e os acoplamentos trilineares são gerados a um *loop* e as massas escalares a dois *loops*.

A quebra de simetria eletrofraca ocorre de maneira similar ao MP, quando os dois campos escalares de Higgs adquirem um valor esperado no vácuo, a razão entre os vevs dos dois campos de Higgs é denotada por $\tan \beta$. Entre os oito graus de liberdade dos campos escalares, três se tornam modos longitudinais do bóson de *gauge* e geram os bósons massivos W^\pm e Z^0 . Os cinco graus de liberdade restantes dão origem a um par de bósons de Higgs carregados, H^\pm , um bóson de Higgs pseudo-escalar neutro CP⁶-ímpar, A^0 e dois bósons de Higgs neutros CP-par, h^0 e H^0 , onde considera-se o mais leve entre eles, h^0 , o bóson de Higgs do MP.

O conjunto de superpartículas fermiônicas, auto-estados de massa, do MSSM é formado por gluinos, neutralinos ($\chi_{1,2,3,4}^0$) e charginos ($\chi_{1,2}^\pm$), onde os neutralinos são formados

⁵Forma geral da lagrangeana em uma teoria de interação de supercampos quirais:

$$\mathcal{L} = \int d^4\theta K(\Phi_i) + \int d^2\theta \mathcal{W}(\Phi_i) + \int d^2\bar{\theta} \bar{\mathcal{W}}(\Phi_i^\dagger), \quad (2.49)$$

onde K é o potencial de Kähler e \mathcal{W} é o superpotencial.

⁶Operação de conjugação de carga e paridade (CP): troca de partícula por antipartícula mais a inversão de suas coordenadas espaciais.

por misturas dos gauginos e higgsinos neutros ($\tilde{B}, \tilde{W}^3, \tilde{H}_1^0, \tilde{H}_2^0$) e os charginos são misturas dos gauginos e higgsinos carregados ($\tilde{W}^+, \tilde{H}_1^+, \tilde{W}^-, \tilde{H}_2^-$). Dentro da fenomenologia das superpartículas temos também a existência de partículas compostas, chamadas de R-hádrons, que são objetos formados por squarks ou gluinos e pártons do Modelo Padrão (quarks e glúons), analogamente aos hádrons normais. R-hádrons são formados quando o tempo de vida médio dos squark ou gluino é maior que a escala de tempo típica da hadronização.

O MSSM estende os 19 parâmetros do MP adicionando mais 105 parâmetros. Contudo, restrições teóricas provenientes dos mecanismos propostos para a quebra de SUSY, junto com requerimentos de supressão de correntes neutras de mudança de sabor e obtenção da escala correta de quebra de simetria eletrofraca sugerem uma origem comum para alguns dos parâmetros em altas escalas de energia. Isso resulta em modelos altamente restringidos como o modelo mínimo de Supergravidade (*minimal supergravity model* - mSUGRA) [41], o MSSM restringido (*constrained MSSM* - cMSSM) [42], a SUSY mediada por *gauge* (*gauge mediated SUSY*) [43] entre outros. Outra abordagem consiste em restringir os parâmetros requerendo que eles satisfaçam os limites aceitáveis das observações experimentais, sem nenhuma restrição teórica prévia. Essa segunda abordagem é conhecida como MSSM fenomenológica (*phenomenological MSSM* - pMSSM).

Os modelos cMSSM e mSUGRA são modelos muito populares de SUSY por sua simplicidade. A maioria das buscas experimentais têm sido projetadas com esses modelos, uma vez que, o pequeno número de parâmetros torna fácil a restrição do espaço de parâmetros baseando-se nos dados observados. Visto que, com a Supersimetria, os acoplamentos de *gauge* se unificam na escala de GUT⁷, o mSUGRA, por exemplo, supõe que todos os termos suaves para as massas dos gauginos são os mesmos em altas escalas de energia (universalidade), sendo a massa unificada dos gauginos em altas energias denotada por $M_{1/2}$. Suas massas na escala eletrofraca são então calculadas através de equações do grupo de renormalização. O modelo mSUGRA também supõe que as massas das superpartículas escalares se unificam em altas energias, sendo denotada por M_0 . No setor de Higgs, temos que a quebra de simetria eletrofraca determina a magnitude do parâmetro de Higgs, μ , mas não seu sinal e também não determina a razão entre os vevs, $\tan \beta$. Todos os termos trilineares também são considerados unificados em escalas de altas energias com o parâmetro comum denotado por A_0 . Logo, o mSUGRA consiste em um modelo com apenas cinco parâmetros livres: $M_{1/2}$, M_0 , $\tan \beta$, A_0 e o sinal de μ . Visto que as equações do grupo de renormalização são fixadas ao se definir o grupo de *gauge* e o espectro de partículas, o espectro de massas é unicamente determinado pela definição dos valores dos cinco parâmetros livres do modelo. Um exemplo de um espectro de massa do mSUGRA,

⁷ Escala de GUT ou escala das Grandes Teorias Unificadas (*Grand Unified Theories* - GUT) é o nível de energia de aproximadamente 10^{16} GeV, no qual se acredita que as forças eletromagnética, fraca e forte se tornam iguais em intensidade e se unificam em uma única força descrita por um grupo de Lie.

ponto de referência SPS⁸ 1a [44], pode ser visto na figura 2.4.

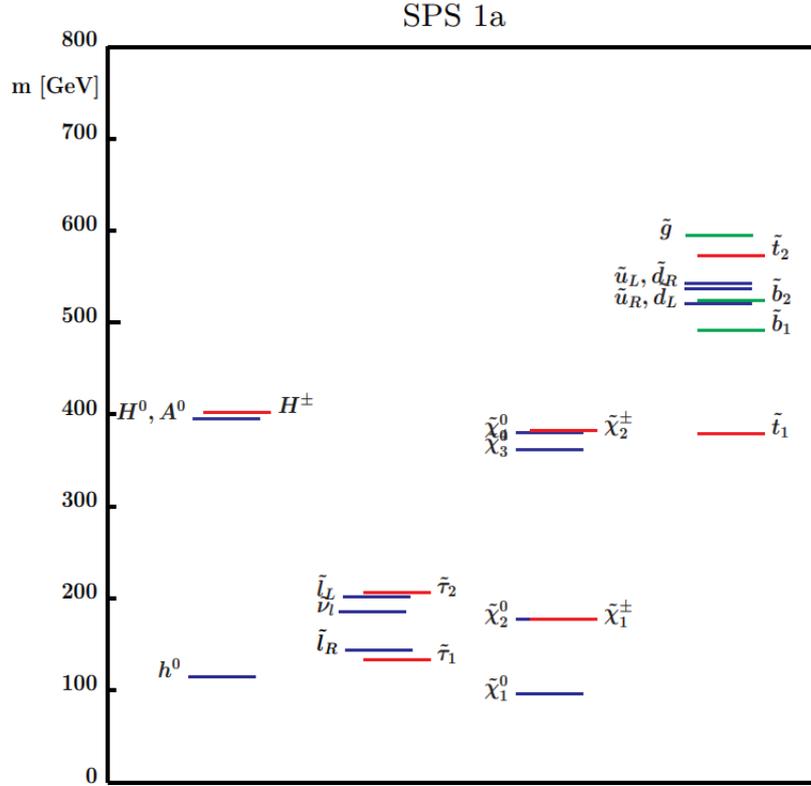


FIGURA 2.4 – Exemplo de um espectro de massa “típico” da mSUGRA [44]. Na primeira coluna estão os campos oriundo do setor de Higgs, na segunda estão os sléptons, na terceira estão os charginos e neutralinos e na quarta o gluino e os squarks.

2.2.1.2 Soluções para questões abertas no Modelo Padrão

A Supersimetria possui o mérito de apresentar soluções naturais para alguns dos problemas em aberto do MP. Algumas dessas soluções são descritas abaixo.

- O requerimento adicional da conservação da paridade-R resulta que a partícula supersimétrica mais leve é estável, e em alguns modelos ela é uma candidata promissora para a matéria escura, visto que espera-se que seja massiva, eletricamente neutra e que interaja apenas gravitacionalmente e fracamente.
- Em Supersimetria, cada partícula do MP tem uma superparceira que contribuirá com correções radiativas de mesma magnitude e sinal oposto no cálculo da massa do bóson de Higgs, eliminando assim as divergências quadráticas e dando uma solução natural para o problema da hierarquia [45].

⁸SPS é a sigla de “*Snowmass Points and Slope*” que forma um conjunto de cenários de referência que surgiu em 2001 no evento “*Snowmass Workshop on the Future of Particle Physics*” como um consenso baseado em propostas de diferentes grupos.

- O espectro de partículas de SUSY é capaz de modificar a evolução das constantes de acoplamento de *gauge* segundo o grupo de normalização, de modo que elas se encontrem na escala de GUT [46], figura 2.5.

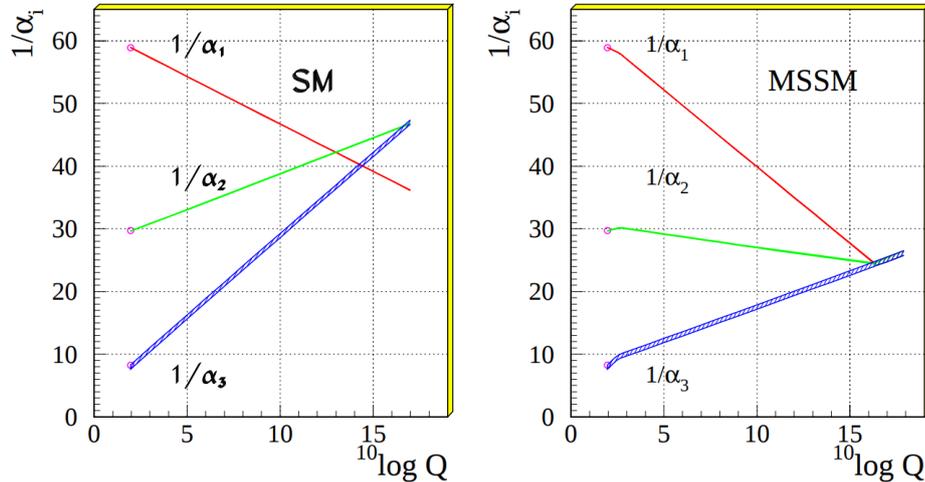


FIGURA 2.5 – Evolução das constantes de acoplamento das interações fundamentais no Modelo Padrão (esquerda), onde α_1 é a constante da interação eletromagnética, α_2 é da interação fraca e α_3 é da interação forte, e no MSSM (direita) [47].

2.2.2 Supersimetria Comprimida

Muitas buscas por SUSY já foram realizadas, sendo um dos tópicos físicos de maior interesse no LHC. Entretanto, nenhum sinal de SUSY foi descoberto até o momento nas buscas “padrões”, baseadas em modelos como mSUGRA e cMSSM. Uma maneira de manter a esperança na existência de Supersimetria é dedicar buscas a versões mais “exóticas” de SUSY. O LHC tem encorajado essas novas buscas e esta tese trata de uma delas, na qual é estudado um cenário de SUSY conhecido como SUSY comprimida, que já possui diversos trabalhos realizados com justificativas teóricas e análises fenomenológicas [48,49,50,51,52]. A SUSY comprimida trata de padrões altamente comprimidos das massas do gauginos e higgsinos na escala dos TeV ou abaixo, enquanto os parceiros escalares ficam fora do alcance de detecção pelo LHC. Em eventos envolvendo a produção de pares de gluinos, a sensibilidade do LHC é drasticamente reduzida quando há uma pequena diferença de massa entre o gluino e a partícula supersimétrica mais leve, oferecendo uma explicação de porque a Supersimetria não foi descoberta ainda.

Na figura 2.6 são mostrados dois esquemas da aparência do espectro de massa de SUSY em modelos supersimétricos típicos e comprimidos. Em buscas por gluinos em modelos típicos, como o modelo de mSUGRA SPS 1a, o gluino é uma das partículas de maior massa, originando um longa cadeia de decaimentos até chegar a LSP. Eventos desse

tipo resultam na produção de vários jatos e léptons. Já em modelos comprimidos, a massa do gluino é só um pouco maior que a massa da LSP, conseqüentemente o decaimento do gluino resulta na produção de jatos pouco energéticos.

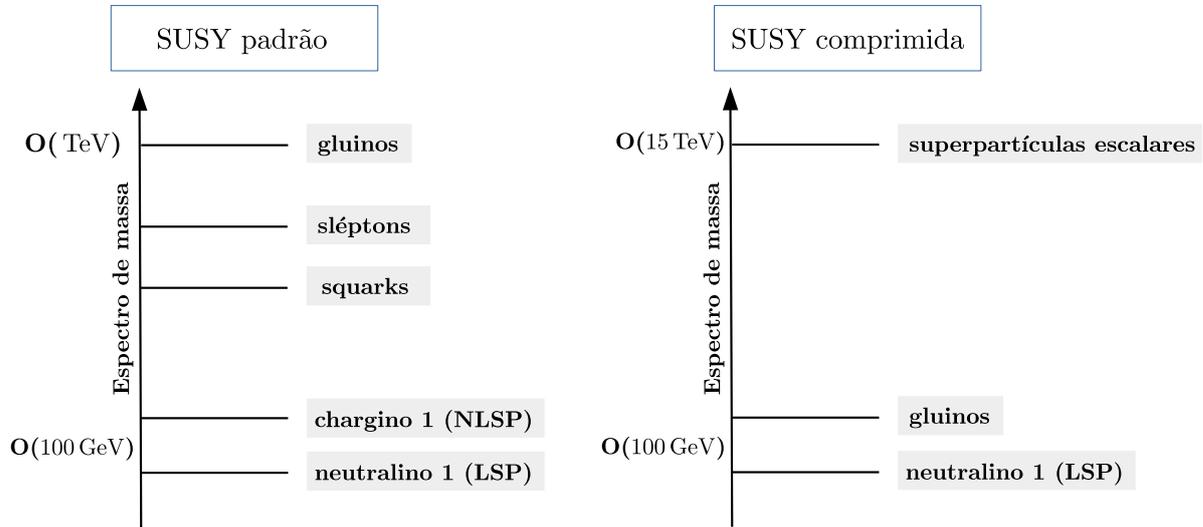


FIGURA 2.6 – Esquema da aparência típica do espectro de massa das partículas supersimétrica em cenários padrões de SUSY (esquerda) e em cenários de SUSY comprimida (direita).

Já foi visto que a SUSY possibilita a unificação dos acoplamentos de *gauge*, alcançada geralmente com a aplicação de correções de limiares de altas escalas de energias. Mesmo em cenários populares como o cMSSM, a unificação dos acoplamentos de *gauge* não é precisa, o acoplamento forte ($\alpha_3 = g_3^2/(4\pi)$) é 3% menor que α_1 e α_2 na escala de grande unificação, GUT, o que pode gerar limiares em altas escalas [50,51]. Um dos atrativos teóricos para os cenários de SUSY comprimida é a possibilidade de conseguirmos a unificação sem a necessidade de aplicar essas correções, situação conhecida como unificação precisa dos acoplamentos de *gauge* (*Precise Gauge Unification* - PGU), oferecendo uma solução natural para os problemas da hierarquia e unificação dos acoplamentos de *gauge*. Existem na literatura vários exemplos de espectros atingindo a PGU. Grande parte deles são excluídos pela não observação de superparceiros ou pela previsão de uma massa relativamente grande para o bóson de Higgs. Em [52] é discutido quais seriam as conseqüências de uma unificação precisa dos acoplamentos de *gauge* na ausência de correções de limiares de altas escalas de energia (PGU) para o espectro de massas das partículas supersimétricas e se a PGU pode ser atingida ou não em modelos realísticos de SUSY.

Sabendo que o espectro de massa depende do padrão de quebra de SUSY, ao se considerar a possibilidade de realizar PGU por massas não universais dos gauginos, uma possibilidade consiste em usar a quebra de simetria mediada por anomalias enquanto espectros mais comprimidos das massas dos gauginos podem ser obtidos em esquemas com uma mistura de mediação por anomalias e pela gravidade, conhecido como esquema

de miragem [53,54,55]. No contexto da mediação por miragem, as massas escalares e os acoplamentos trilineares são dependentes do modelo. As massas escalares estão geralmente na ordem da massa do gravitino e os acoplamentos trilineares são suprimidos podendo ser desprezados. O espectro dos gauginos é tipicamente mais comprimido e suas massas em baixas energias podem ser aproximadas por,

$$\begin{aligned} m_{\tilde{B}} &\simeq 0,45(\varrho + 3,3) \frac{m_{3/2}}{16\pi^2}, \\ m_{\tilde{W}} &\simeq 0,9(\varrho + 0,5) \frac{m_{3/2}}{16\pi^2}, \\ m_{\tilde{g}} &\simeq 2,4(\varrho - 1,5) \frac{m_{3/2}}{16\pi^2}, \end{aligned} \quad (2.51)$$

onde $m_{3/2}$ denota a massa do gravitino e ϱ parametriza a contribuição da mediação da gravidade para a massa dos gauginos. No limite em que ϱ tende a zero, a mediação por anomalia é recuperada. Esse modelo possui então cinco parâmetros livres, $m_{3/2}$, ϱ , m_0 , μ e $\tan\beta$. Contudo, m_0 é geralmente eliminado requerendo que a massa do Higgs mais leve seja $m_h = 126$ GeV, que resulta em $m_0 \sim 15$ TeV.

Embora não seja possível prever a hierarquia entre os gauginos e higgsinos dentro do esquema de mediação por miragem, há uma tendência de que a PGU prefira espectros de massa nos quais o gluino é a partícula mais próxima da partícula supersimétrica mais leve (*next-to-lightest supersymmetric particle* - NLSP) e que a diferença entre as massas do gluino e do neutralino seja pequena ($\lesssim 100$ GeV). Se o gluino é a segunda partícula mais leve, os modos de decaimento relevantes são

$$\tilde{g} \rightarrow q\bar{q} + \tilde{\chi}_1^0, \quad (2.52)$$

$$\tilde{g} \rightarrow g + \tilde{\chi}_1^0. \quad (2.53)$$

O primeiro processo acontece via um squark *off-shell* e o segundo via um *loop* de quark-antiquark. A diferença de massa entre os squarks *left-handed* e *right-handed* é tipicamente pequena, ou seja, a paridade é aproximadamente preservada no setor de squarks e o decaimento (2.53), o qual requer violação de paridade [56], tem tipicamente uma contribuição muito pequena, podendo ser desprezado. Modos de decaimento adicionais apareceriam se neutralinos ou charginos adicionais fossem mais leves que o gluino.

A figura 2.7 mostra o padrão de massas do gluino e neutralino, gerada com escolhas aleatórias dos parâmetros livres nos intervalos: $\varrho = 0,5 - 30$, $m_{3/2} = (40 - 200)\varrho$ TeV, $\mu = 0,1 - 2$ TeV e $\tan\beta = 10 - 50$, que tenderiam para uma bem sucedida realização da PGU e foi obtido gerando uma amostra com escolhas aleatórias dos parâmetros livres. Dessa forma, é gerado o gráfico de espalhamento onde os pontos cinzas correspondem aos valores de $m_{\tilde{g}}$ e $m_{\tilde{\chi}_1^0}$ que resultaram em PGU. De forma ilustrativa, são incluídos limites nas massas dos gluinos de uma busca do experimento ATLAS (*A Toroidal LHC*

ApparatuS) [57] por jatos mais energia transversa perdida (*Missing Transverse Energy* - MET) [58], e uma busca do CMS por b-jatos mais energia transversa perdida [59]. Ambas as análises são baseadas em modelos simplificados com o decaimento (2.52) possuindo razão de ramificação de 100%. O limite do ATLAS fornece apenas uma estimativa razoável e o limite do CMS não é aplicável, uma vez que, não há preferência por b-jatos nos estados finais em comparação a jatos oriundos de quarks leves e glúons. Uma fração pequena dos pontos caracterizam os gluinos como a partícula mais leve, uma restrição é incluída no gráfico oriunda de buscas por R-hádrons. Em mais de 90% dos pontos de referência satisfazem $m_{\tilde{\chi}_1^0} > 0,5m_{\tilde{g}}$, mostrando uma tendência por espectros comprimidos.

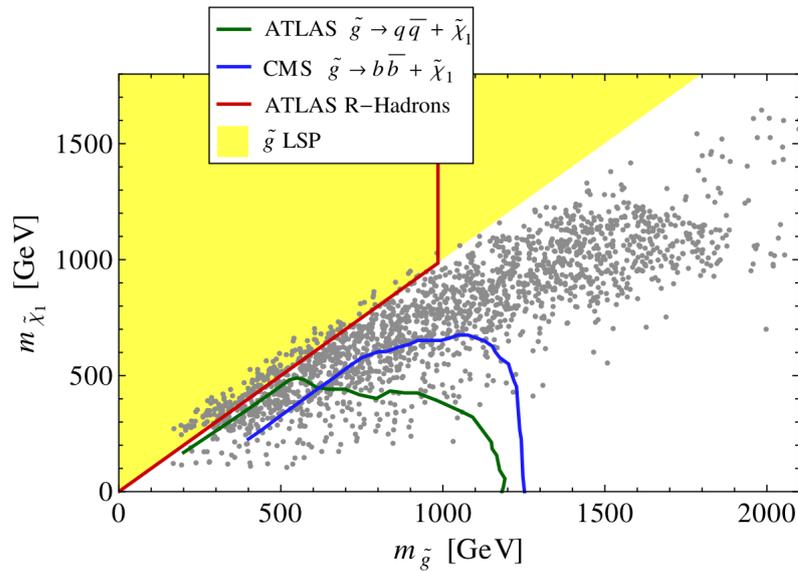


FIGURA 2.7 – Pontos da amostra que resultam em PGU (cinza). Também são mostradas restrições na massa do gluino para os modos de decaimento $\tilde{g} \rightarrow q\bar{q} + \tilde{\chi}_1^0$ (verde) e $\tilde{g} \rightarrow b\bar{b} + \tilde{\chi}_1^0$ (azul). A área amarela denota a região onde o gluino seria a LSP, na qual contém uma restrição (linha vermelha) proveniente de buscas por R-hádrons estáveis no ATLAS [52].

Nesse cenário o gluino pode ter vida longa, parcialmente porque o espaço de fase é suprimido devido à pequena diferença de massa entre o gluino e o neutralino e parcialmente devido à grande massa do squark que media o decaimento do gluino (~ 15 TeV). Na figura 2.8 é exposta a distribuição da largura de decaimento total do gluino, $\Gamma_{\tilde{g}}$, gerada com a ferramenta SDECAY 1.3 [60], para a mesma amostra usada na figura 2.7. A reta vertical em $10 \mu\text{m}$ denota a resolução aproximada do LHC. É observado que um pouco mais de 10% dos pontos de referência possuem um comprimento de decaimento (*decay length*), $c\tau_{\tilde{g}}$, maior que $10 \mu\text{m}$, o que sugere que em uma pequena fração do espaço de parâmetros, pode ser possível observar vértices deslocados, também chamados de vértices secundários. Gluinos de vida longa são particularmente obtidos se a partícula supersimétrica mais leve for composta quase que totalmente por gauginos (*gauginolike*) [61]. A distribuição da figura 2.8 é um cálculo aproximado apenas, uma estimativa da ordem de grandeza devido

à incerteza na massa do squak.

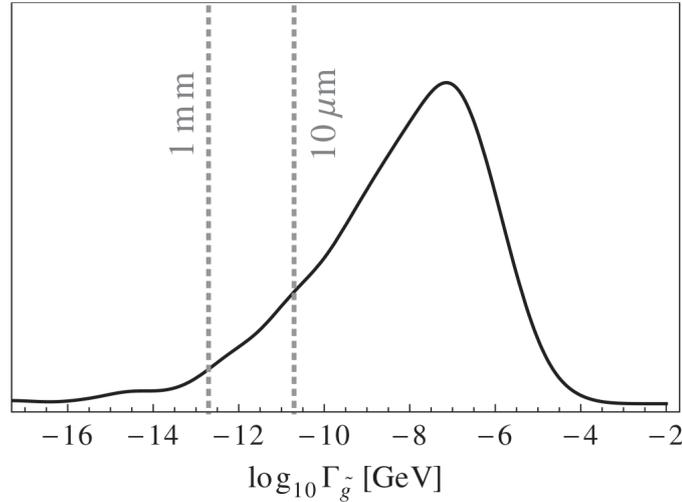


FIGURA 2.8 – Distribuição da largura de decaimento na amostra de referência. As linhas tracejadas verticais indicam onde o comprimento de decaimento do gluino, $c\tau_{\tilde{g}}$, atinge $10 \mu\text{m}$ e 1mm [52].

2.2.2.1 Modelo Simplificado

Historicamente, a interpretação de buscas por SUSY tem sido realizada no contexto dependente de modelo [62]. Isso maximiza a sensibilidade, visto que todos os potenciais canais de produção e estados finais de um modelo são levados em conta. Os resultados, no entanto, são muito difíceis de reinterpretar em termos de outro modelo. Uma solução para esse problema são os chamados modelos simplificados [63], que fornecem uma ferramenta simples para procurar por assinaturas de SUSY.

Modelos simplificados focam em um único modo de produção e uma cadeia de decaimento única, com uma ou duas etapas, até um estado final relevante para um determinado modelo ou análise. Apenas algumas partículas Além do Modelo Padrão (*Beyond the Standard Model* - BSM) são consideradas, enquanto o resto é considerado dissociado do espectro. Sendo assim, os modelos simplificados não podem capturar todos os detalhes de uma teoria, mas eles são construídos para refletir as propriedades mais importantes de um determinado modelo. Um modelo simplificado pode então ser descrito por um pequeno número de parâmetros que estão diretamente relacionados aos observáveis físicos, tais como as massas, seções de choque de produção ou fração de ramificação das partículas. Isso é especialmente útil no caso de cenários de SUSY, que geralmente têm um alto número de parâmetros livres de modo que é impossível interpretar os resultados de uma busca no espaço de parâmetros completo da teoria.

Embora modelos simplificados não sejam independentes de modelo, eles permitem reinterpretções em termos de teorias completamente diferentes, desde que possuam es-

tados finais similares [63]. Um exemplo são as buscas por eventos com jatos e energia transversa perdida, que são esperados em uma variedade de modelos supersimétricos e outras teorias BSM [63,64]. Geralmente, os modelos simplificados têm as seguintes aplicações pretendidas [63]:

- **Identificar os limites da sensibilidade da busca** - qualquer pesquisa precisa incluir uma identificação clara dos limites da sensibilidade. Um exemplo importante é a dependência da eficiência de seleção de eventos de sinal com respeito à diferença de massa entre uma partícula e seus produtos de decaimento, visto que isso muitas vezes tem uma grande influência nas propriedades cinemáticas de um estado final. Essa informação ajuda a identificar modelos bem motivados de cenários nos quais as estratégias de busca existentes não são eficientes.
- **Caracterizar novos sinais físicos** - caso seja observada evidência para um física BSM, é importante caracterizar o intervalo de massas, topologias de decaimentos e possíveis números quânticos das partículas que podem resultar no desvio observado do esperado no Modelo Padrão. Com base nisso, a consistência de um dado modelo de sinal ou uma teoria completa podem ser avaliados.

A análise desenvolvida nesta tese é baseada em um modelo simplificado caracterizado pela produção de dois gluinos que decaem através do canal de decaimento (2.52) com 100% da razão de ramificação, com cada gluino resultando em um quark leve e um squark de mesmo sabor ($\tilde{g} \rightarrow \tilde{q}\bar{q}$), que é seguido por um decaimento do squark em um outro quark leve e um neutralino ($\tilde{q} \rightarrow q\tilde{\chi}_1^0$). Sabendo-se que o squark possui uma massa muito alta nesse modelo simplificado, essa etapa do decaimento é substituída por um decaimento de três corpos ($\tilde{g} \rightarrow q\bar{q}\tilde{\chi}_1^0$), onde um acoplamento efetivo é introduzido. No total, o estado final soma dois quarks, dois antiquarks e dois neutralinos, como pode ser visto na figura 2.9. Fora o gluino e o neutralino mais leve, todas as outras partículas supersimétricas são consideradas dissociadas do espectro. O objetivo do modelo simplificado usado é estudar os limites da sensibilidade da busca com respeito à massa do gluino, $m_{\tilde{g}}$, e à diferença de massa, Δm , entre o gluino e o neutralino LSP.

Os parâmetros do modelo simplificado considerado foram configurados de modo a que fossem produzidos gluinos de vida longa. A figura 2.10 mostra como varia o comprimento de decaimento dos gluinos com relação a $m_{\tilde{g}}$ e Δm . Pode-se observar uma grande variação do comprimento de decaimento entre os valores extremos de Δm . Esse comportamento expõe o quanto a supressão no espaço de fase em cenários de compressão extrema afeta o tempo de vida dos gluinos.

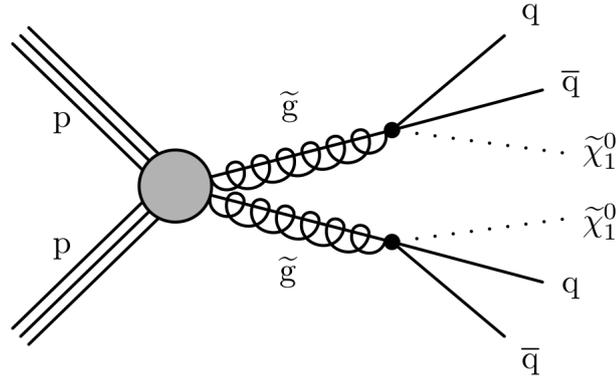


FIGURA 2.9 – Modelo simplificado da produção de dois gluinos na colisão próton-próton decaindo ambos em quark, antiquark e no neutralino LSP. O círculo cinza representa os potenciais canais de produção dos gluinos ($pp \rightarrow \tilde{g}\tilde{g}$).

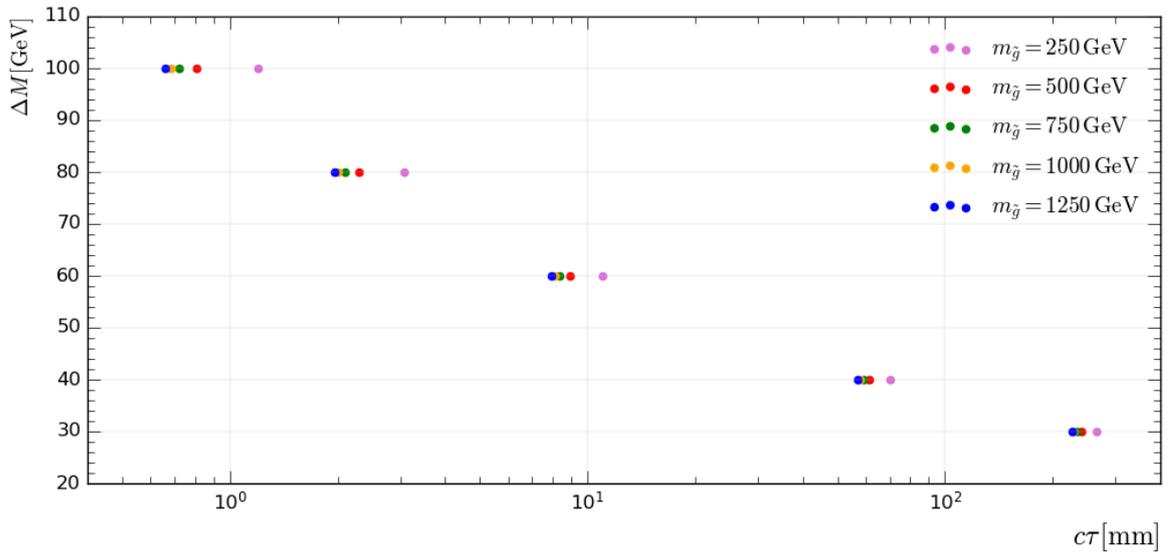


FIGURA 2.10 – Gráfico da dependência do comprimento de decaimento do gluino com respeito à massa do gluino e à diferença de massa entre o gluino e o neutralino. Gerado com o pacote SDECAY 1.5a [60].

2.2.2.2 Caracterização e Cinemática dos Estados Finais

Para categorizar os eventos candidatos de sinal, variáveis de busca com vantagens especiais são introduzidas. Suas distribuições esperadas de potenciais modelos de sinal dependem altamente das massas e das diferenças de massa no modelos de SUSY. Assim, buscas por SUSY frequentemente fazem uso de um elevado número de regiões de busca, definidas por cortes nessas variáveis, a fim de fornecer uma boa cobertura de possíveis realizações de SUSY. Um dos observáveis mais fundamentais é o número de jatos no evento. A multiplicidade de jatos N_{jets} ajuda a distinguir modelos com quarks top no estado final dos modelos com quarks bottom ou apenas quarks leves, já que o primeiro normalmente resulta em um número maior de jatos no estado final. Da mesma forma, a

multiplicidade do jato também depende do cenário do sinal: modelos com produção de pares de gluinos normalmente têm mais jatos no estado final do que modelos com produção de pares de squarks. Uma variável que descreve a atividade hadrônica geral no evento é a soma escalar dos momentos transversos de todos os jatos, H_T . A energia transversa perdida, E_T^{miss} , mede o desequilíbrio no momento transverso de todas as partículas do estado final que foram detectadas. Ambos H_T e E_T^{miss} dependem muito do espectro de massa dos modelos supersimétricos considerados. Outro observável que pode ser usado para classificar ainda mais o evento é o número de jatos de quarks bottom, $N_{b\text{jets}}$.

Existem outros observáveis mais complexos que podem ser usados para caracterizar um candidato a evento de sinal. Estudos envolvendo simulação de amostras de eventos de sinal e *background* identificaram várias variáveis que podem ser usadas para distinguir os eventos de *background* do MP dos possíveis eventos de sinal de modo que regiões de busca enriquecidas de sinal possam ser definidas. Alguns exemplos de outras variáveis amplamente utilizadas e mais sofisticadas incluem M_{T2} [67,68], α_T [69,70] e as chamadas variáveis *raser* [71,72]. Alguns desses observáveis focam em uma topologia de sinal bem distinta e/ou exploram certas propriedades cinemáticas. As variáveis usadas nesta análise são definidas detalhadamente ao longo do capítulo 4.

Em estados finais com pelo menos dois neutralinos LSPs que não serão detectados pelo detector, pois interagem apenas fracamente, os neutralinos carregarão uma energia igual a pelo menos duas vezes a sua massa. Em cenários comprimidos, é possível que uma grande fração do momento dos neutralinos se cancele, o que reduz significativamente a energia transversa perdida, conforme ilustrado na figura 2.11. Os gluinos são produzidos em direções opostas no plano transverso, e como a maior parte do momento é transferida para os neutralinos, os seus momentos transversos também estão aproximadamente alinhados em direções opostas. Isso leva a um valor baixo de E_T^{miss} e é bastante desafiador do ponto de vista experimental, porque um alto número de eventos de *background* do MP são esperados nessa região. Além disso, nesse tipo de evento, dificilmente os jatos provenientes do decaimento do neutralino possuem energia suficiente para serem selecionados pelos triggers, necessitando então de E_T^{miss} suficientemente elevada para que o evento seja selecionado por um *trigger* de MET. Esses eventos possuem MET suficientemente alta nos casos em que são produzidos um ou mais jatos adicionais no estado de radiação inicial (ISR), que impulsionam ambos os neutralinos para a mesma direção, conforme ilustrado na figura 2.11. Apesar disso, as regiões comprimidas permanecem experimentalmente desafiadoras e uma diminuição substancial na sensibilidade das pesquisas para esses modelos é esperada [73].

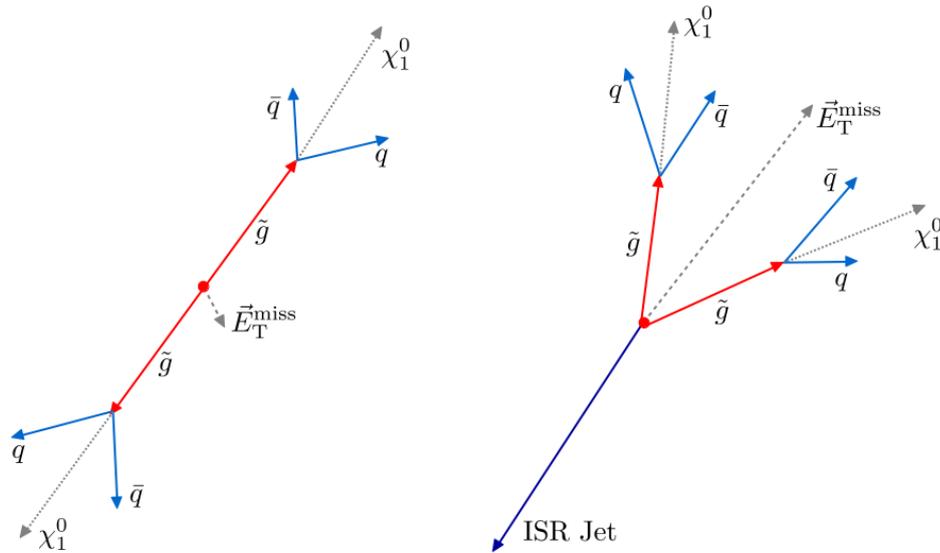


FIGURA 2.11 – Produção de um par de gluínos na colisão próton-próton, os neutralinos são produzidos em direções opostas quase colineares resultando em uma baixa energia transversa perdida (esquerda). O par de gluínos é produzido junto com um jato do estado de radiação inicial que gera um desalinhamento maior entre os neutralinos produzidos, aumentando a quantidade de energia transversa perdida (direita).

2.3 Buscas por Supersimetria no LHC

Antes do LHC, muitas buscas diretas por partículas supersimétricas foram realizadas em experimentos de colisores de partículas como o LEP (*Large Electron-Positron Collider*), com uma energia de centro de massa de até $\sqrt{s} = 209 \text{ GeV}$, e o Tevatron, até $\sqrt{s} = 1,96 \text{ TeV}$. Nenhum desses experimentos observou qualquer tipo de sinal de SUSY. A sensibilidade das buscas por partículas supersimétricas foi grandemente ampliada com os experimentos do LHC no *Run I*, primeiro período de tomada de dados do LHC (2009-2013) com uma energia de centro de massa de até 8 TeV . Entretanto, nenhum sinal de SUSY foi observado, estimando os limites inferiores para partículas supersimétricas coloridas com massas de até 1 TeV e também melhorando os limites de massa para partículas que interagem fracamente. No *Run 2* (2015-2018) com energia de centro de massa de 13 TeV , há um aumento significativo nas seções de choque de produção das partículas supersimétricas, figura 2.12, o que permitiu estender os limites das massas para a ordem de 2 TeV . Mesmo que nenhum sinal óbvio de alguma física além do Modelo Padrão tenha sido encontrado, SUSY na escala dos TeV ainda é altamente motivada, por causa do aumento das seções de choque e sensibilidade das pesquisas, estados finais promissores estão sendo cuidadosamente investigados novamente.

A figura 2.13 mostra uma seleção representativa de buscas por SUSY, envolvendo a produção de pares de gluínos, no experimento CMS. Este resumo serve como um guia qualitativo para o alcance de massa em vários canais de busca, mas qualquer interpretação

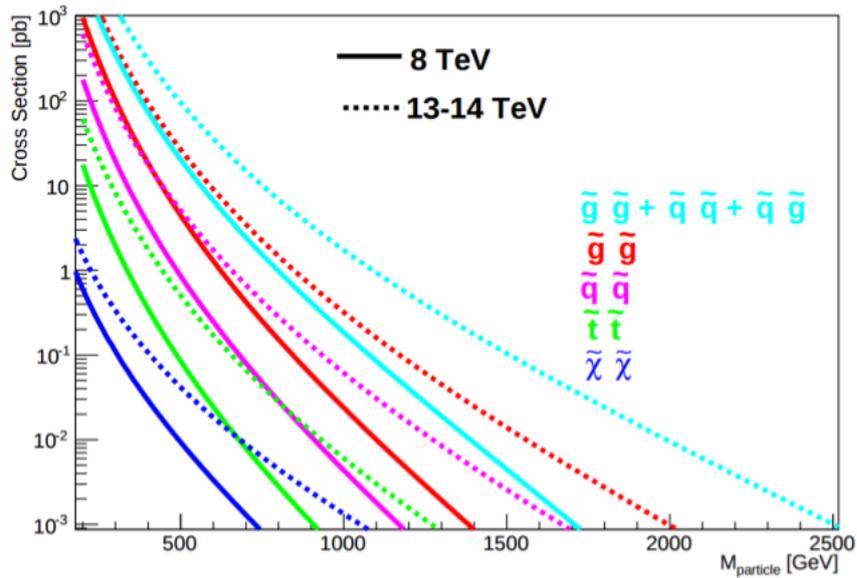


FIGURA 2.12 – Seções de choque de produção de pares de partículas supersimétricas no LHC em energias de centro de massa de $\sqrt{s} = 8 \text{ TeV}$ e $\sqrt{s} = 13 - 14 \text{ TeV}$ [74].

direta deve ser tratada com cuidado. Como típico para modelos simplificados, supõe-se uma fração de decaimento de 100% e as partículas não envolvidas são dissociadas do espectro. Além disso, para o sumário ilustrado, são escolhidos valores arbitrários da massa de neutralino que servem de referência. Esses limites têm sempre que ser interpretados como um limite superior do potencial de exclusão das buscas com respeito ao intervalo de massa. Ainda assim, essa figura fornece uma visão geral dos estados finais e canais de pesquisa, e expõe o trabalho intensivo que a comunidade de física de partículas realiza, a fim de cobrir tanto quanto possível o espaço cinemático de SUSY.

A figura 2.14 mostra o esforço do detector ATLAS na busca por SUSY. Esse sumário é mais abrangente que o do CMS por considerar outros canais de busca além dos originários da produção de pares de gluinos. Temos as categorias de buscas inclusivas realizadas no canal hadrônico que possui a maior taxa de produção do LHC, produção direta de squarks da terceira geração, produção direta de partículas supersimétricas do setor eletrofraco, produção de partículas de vida longa e busca por canais com violação da paridade-R. A definição exata dessas categorias pode variar de um experimento para outro, no CMS, por exemplo, as buscas inclusivas englobam apenas os canais hadrônicos com a produção de no máximo 1 lépton enquanto buscas envolvendo dois ou mais léptons entram na categoria de buscas leptônicas. Em geral, as assinaturas de SUSY observadas no LHC são caracterizadas pela quantidade de léptons, jatos e E_T^{miss} .

O experimento LHCb realiza buscas por Supersimetria, geralmente de maneira indireta, complementares ao ATLAS e CMS. Decaimentos de partículas transmitem informação sobre as propriedades internas e o funcionamento dos processos físicos. Ao estudá-los

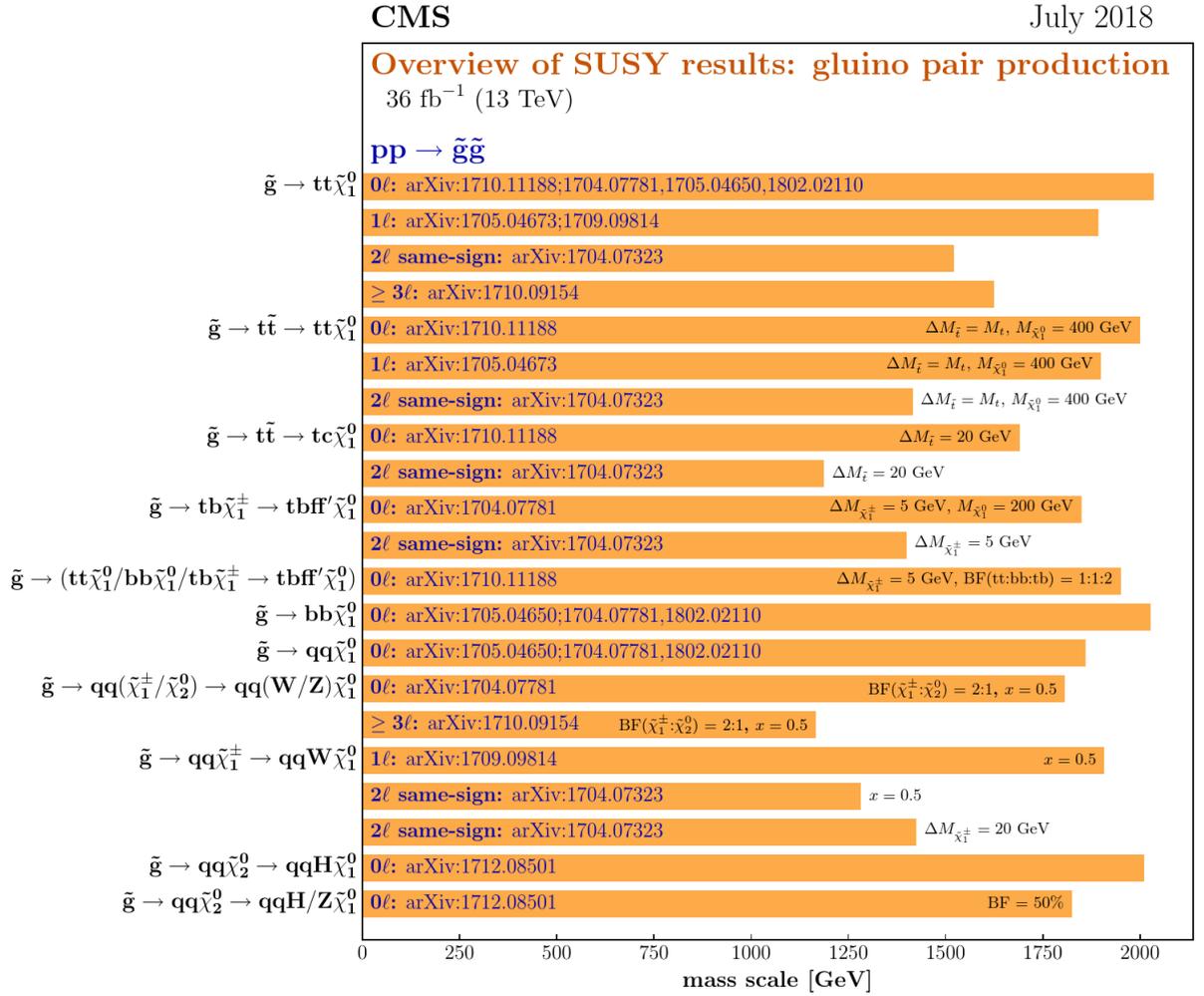


FIGURA 2.13 – Visão geral dos resultados de buscas por SUSY no CMS envolvendo a produção de pares de gluinos. Os limites observados da massa dos gluinos com 95% C.L. são representados pela barra laranja [75].

e sua ocorrência, os físicos podem inferir as regras que os controlam. Frequentemente, verifica-se que alguns decaimentos raros, difíceis de observar, são capazes de revelar a presença de uma física além do Modelo Padrão. Um exemplo, é o caso do decaimento $B_s^0 \rightarrow \mu^+ \mu^-$, sendo B_s^0 uma partícula composta por um anti-quark *bottom* ligado a um quark *strange*. Sua taxa é prevista com muita precisão no Modelo Padrão e foi observada pela primeira vez em 2015 combinando resultados do LHCb e do CMS [76]. Uma grande classe de teorias que estendem o MP, como a Supersimetria, permitem modificações significativas na sua razão de ramificação, portanto, uma observação de qualquer desvio significativo da previsão do MP indicaria uma descoberta de novos efeitos.

Em experimentos com colisões de hádrons, buscas por SUSY são frequentemente realizadas buscando estados finais com multijatos e energia transversa perdida, ou seja, sem a presença de léptons energéticos isolados no evento. Esse canal de busca é considerado

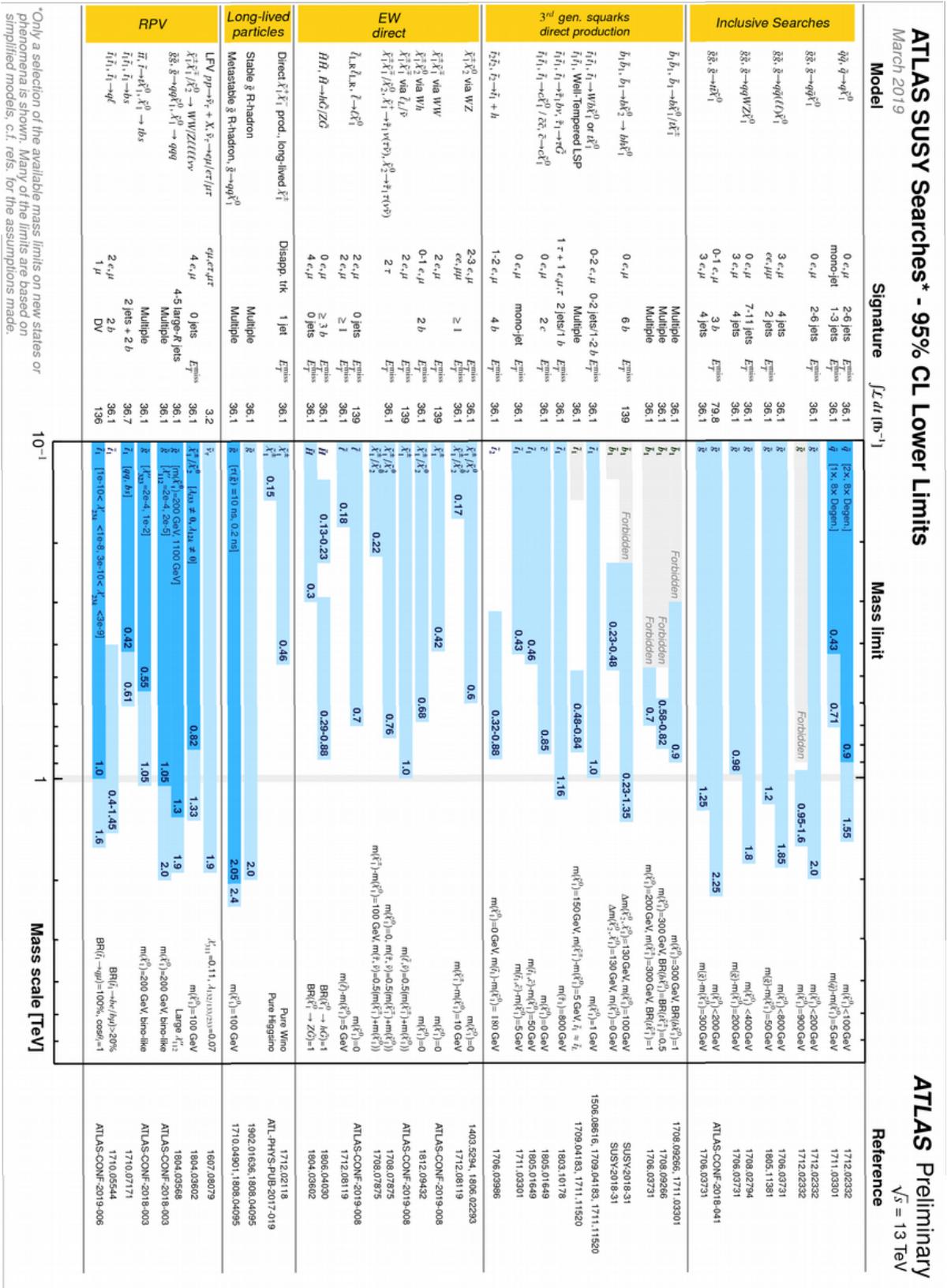


FIGURA 2.14 – Visão geral dos resultados de buscas por SUSY no ATLAS. Os limites observados da massa das partículas produzidas na colisão próton-próton com 95% C.L. são representados pela barra azul [77].

o canal de descoberta mais provável de SUSY no LHC. O canal de busca totalmente hadrônico geralmente tem uma razão de ramificação muito maior do que os canais de busca que incluem léptons, mas também possui mais *backgrounds* do Modelo Padrão [78]. Buscas no canal totalmente hadrônico requerem tipicamente a reconstrução de 2 ou mais jatos com $p_T > 30$ GeV, assim, à medida que considera-se gluino e neutralino com diferença de massa cada vez menor, diminui também a probabilidade de se ter jatos oriundos do decaimento do gluino que ultrapassem o limiar de $p_T > 30$ GeV. Conseqüentemente, modelos de SUSY comprimida são muito mais difíceis de serem detectados no LHC que outros modelos ordinários de SUSY, uma considerável fração dos potenciais eventos de sinal são rejeitados devido a terem jatos muito pouco energéticos (*soft jets*). Nos casos de compressão extrema, apenas eventos com estado de radiação inicial são capazes de passar pela seleção de eventos.

Buscas envolvendo a produção de partículas com vida longa possuem um tratamento especial no CMS e ATLAS, pois essa propriedade é geralmente usada na elaboração da estratégia de busca para melhorar a performance das mesmas. Como será descrito em mais detalhes no capítulo 4, este trabalho usa a propriedade de vida longa dos gluinos para elaborar uma estratégia de busca que faz uso de vértices e jatos deslocados do ponto de colisão para obter informações sobre o decaimento dos gluinos e melhorar a sensibilidade da busca para eventos de SUSY comprimida.

3 Experimento

A física de partículas experimental é fortemente baseada no uso de aceleradores de partículas. Essas máquinas aceleram feixes de partículas carregadas até altas energias e os fazem colidir com um alvo estático ou com outro feixe. Os dados provenientes da colisão são armazenados e utilizados no estudo das propriedades e interações das partículas descritas pelo MP e também na busca de novas partículas postuladas por teorias além do Modelo Padrão. Detectores de partículas são construídos com propósitos específicos e a informação coletada por eles pode variar drasticamente. Nesta tese são analisados dados de colisões próton-próton (pp) coletados pelo experimento CMS, instalado no LHC.

3.1 O LHC

O LHC pertence à Organização Europeia para a Pesquisa Nuclear conhecida como CERN (antigo acrônimo para *Conseil Européen pour la Recherche Nucléaire*), um dos mais respeitados centros de pesquisa do mundo. Medindo 27 quilômetros de circunferência, o LHC é o maior acelerador de partículas e o de maior energia existente atualmente, projetado para produzir colisões entre feixes de prótons até energias de centro de massa de $\sqrt{s} = 14 \text{ TeV}$ ou feixes de íons pesados (chumbo-chumbo) até $\sqrt{s} = 5,5 \text{ TeV}$. Em diferentes pontos do LHC, estão localizados quatro grandes detectores de partículas, CMS, ATLAS, LHCb e ALICE (*A Large Ion Collider Experiment*) [79], responsáveis por coletar os dados das colisões. O CMS e o ATLAS formam a dupla de experimentos destinados a buscas de propósitos gerais, projetados para investigar o maior espectro físico possível. O LHCb é projetado para estudar a violação de CP nas interações eletrofracas e no estudo da assimetria de matéria e antimatéria através da análise de decaimentos raros de hádrons contendo o quark b and c. Por fim, o ALICE tem como objetivo o estudo de colisões de íons pesados, nas quais é produzido um estado especial de matéria, plasma quark-glúon, onde os quarks e glúons estão em um estado quase livre. O LHC está localizado na fronteira da França com a Suíça entre o Lago Geneva e as montanhas Jura, instalado em um túnel subterrâneo que anteriormente hospedava o LEP (*Large Electron Positron Collider*) [80], figura 3.1.

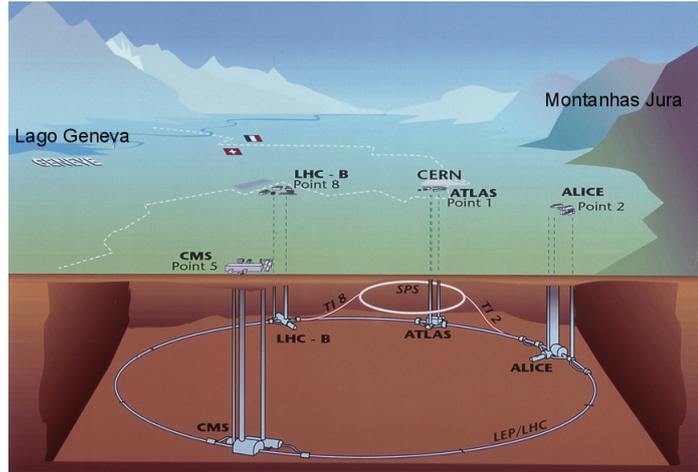


FIGURA 3.1 – Visão abrangente dos experimentos do LHC localizados em um túnel subterrâneo na fronteira da França com a Suíça.

As vantagens de um colisor hadrônico em comparação ao seu antecessor leptônico, LEP, começam primeiramente com a possibilidade de construir um acelerador do mesmo tamanho, podendo ser acomodado dentro do mesmo túnel, que atingisse maiores energias de centro de massa, devido à menor quantidade de energia perdida através da emissão de radiação síncrotron por partículas aceleradas, a qual é proporcional à quarta potência da razão entre a energia e a massa dessas partículas. A outra vantagem é que a estrutura complexa dos prótons permite o acesso a um espectro de energia mais abrangente, o que é uma característica desejada na busca por física além do Modelo Padrão.

O processo de geração e aceleração do feixe de prótons acontece em várias etapas [81]. A fonte de prótons é um Duoplasmatron, cilindro de metal onde é injetado o gás hidrogênio, que tem os prótons separados dos elétrons devido à ação de um campo elétrico. Os prótons são enviados ao LINAC2, um acelerador linear de 36 m que gera um feixe com energia de 50 MeV. O feixe é então enviado ao PSB (*Proton Synchrotron Booster*), um acelerador circular composto por quatro anéis síncrotron sobrepostos com uma circunferência de aproximadamente 160 m, que aumenta a energia dos prótons até 1,4 GeV. Na sequência, os prótons são injetados no PS (*Proton Synchrotron*), um anel síncrotron único com uma circunferência de 600 m onde a energia é aumentada até 25 GeV. Essa combinação sequencial de dois síncrotrons permite a produção de “pacotes” (*bunches*) de prótons, com aproximadamente $1,18 \times 10^{11}$ prótons, no pico da intensidade, intercalados por 25 ns (em 2016), necessários para a correta operação do LHC. O estágio final da injeção do feixe de prótons é o SPS (*Super Proton Synchrotron*), um síncrotron com uma circunferência de aproximadamente 7 km, onde os prótons alcançam uma energia de 450 GeV. Subsequentemente, os prótons são extraídos e injetados no anel do LHC através de duas linhas de transmissão, gerando dois feixes circulando em direções opostas em dois tubos paralelos até alcançarem a energia desejada, figura 3.2. Os pacotes dos feixes são colimados dentro

de um comprimento de 8 cm e um raio transverso de aproximadamente $16 \mu\text{m}$. O LHC pode armazenar até 2808 pacotes, os quais colidem com uma taxa de 40 MHz no espaço nominal de 25 ns entre os pacotes.

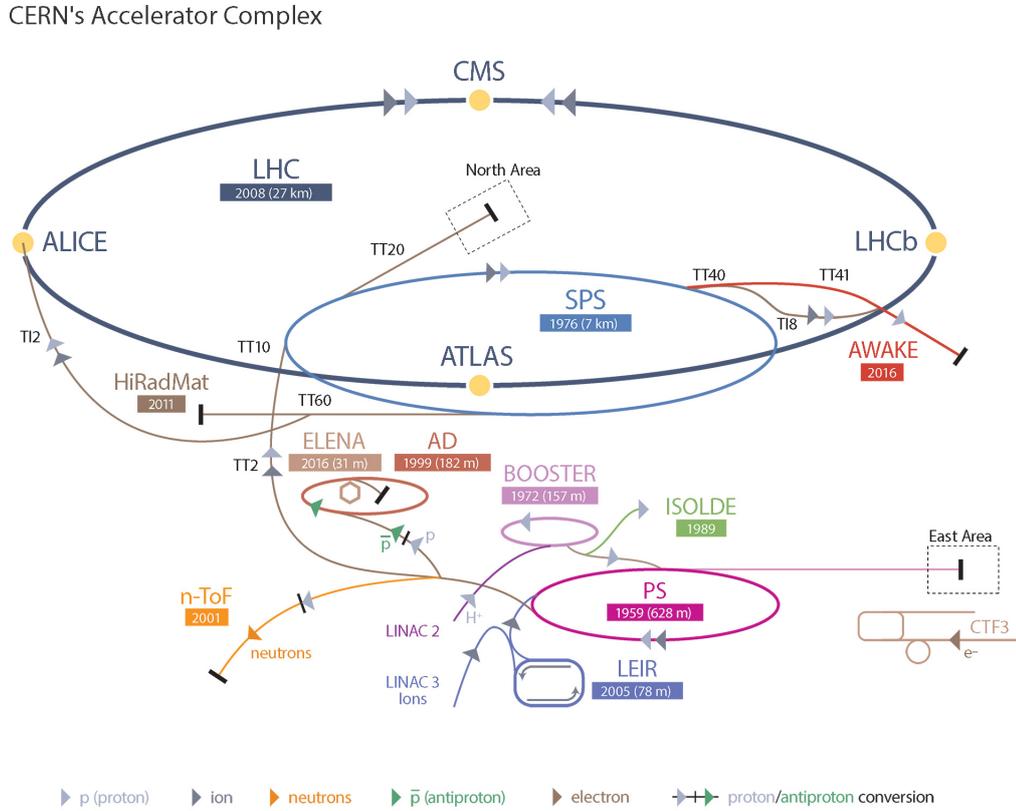


FIGURA 3.2 – Representação esquemática do complexo do LHC [82]. Os feixes de prótons são acelerados, passo a passo, por uma cadeia de aceleradores.

Um dos parâmetros principais de um acelerador é a luminosidade instantânea, \mathcal{L} , que no caso ideal mede o número possível de colisões por unidade de área e unidade tempo. Desse modo, a taxa esperada de eventos de um processo X , dN_X/dt , que possui seção de choque σ_X , é dada por

$$\frac{dN_X}{dt} = \sigma_X \cdot \mathcal{L}. \quad (3.1)$$

A luminosidade instantânea é um parâmetro que depende das características de construção do acelerador e pode ser expressa pela seguinte expressão aproximada:

$$\mathcal{L} = f \frac{n N_b^2}{4\pi \sigma_x \sigma_y} F(\theta_c, \sigma_x, \sigma_y). \quad (3.2)$$

Onde f é a frequência de revolução do feixe, n é o número de pacotes por feixe, N_b é o número de prótons por pacote, σ_x e σ_y caracterizam a seção transversal do feixe e F modela a dependência da luminosidade instantânea com respeito ao ângulo de cruzamento entre os dois feixes, θ_c . O LHC foi projetado para alcançar uma luminosidade instantânea

de $10^{34} \text{ cm}^{-2}\text{s}^{-1}$, que permite aproximadamente 1 bilhão de interações próton-próton por segundo.

A luminosidade integrada é uma medida da quantidade de eventos coletados pelo detector, ela é obtida integrando-se a luminosidade instantânea sobre o período de tomada de dados,

$$L = \int \mathcal{L} dt. \quad (3.3)$$

Em 2016, o LHC entregou uma luminosidade integrada de $L = 41,0 \text{ fb}^{-1}$, da qual $L = 37,8 \text{ fb}^{-1}$ foi coletada pelo experimento CMS [83]. Após os dados serem armazenados, o CMS realiza uma verificação dos dados de modo a selecionar para análise física somente os eventos coletados quando todos os subsistemas do detector estavam completamente operacionais. A luminosidade integrada disponível para análises físicas em 2016 foi de $L = 35,9 \text{ fb}^{-1}$.

Seções de luminosidade são subdivisões da luminosidade integrada, durante as quais espera-se que a luminosidade instantânea não mude significativamente. O *status* do detector é gravado no banco de dados em cada seção de luminosidade, dessa forma, as seções em que parte do detector não está funcionando adequadamente são excluídas de uma posterior análise física. Elas têm aproximadamente 23 segundos de duração, sendo grandes o suficiente para uma determinação precisa da luminosidade e curtas o suficiente para que haja uma perda mínima de dados quando o detector não está completamente funcional.

As primeiras colisões próton-próton do LHC começaram a ser coletadas em 2010 com energia de centro de massa de $\sqrt{s} = 7 \text{ TeV}$. A tomada de dados foi estendida até o final de 2012 e a energia de centro de massa foi aumentada para $\sqrt{s} = 8 \text{ TeV}$, isso ocorreu devido aos resultados promissores das buscas pelo bóson de Higgs que culminaram na sua descoberta divulgada pelos experimentos CMS e ATLAS naquele mesmo ano. O período de 2010 até 2012 é conhecido como *Run 1*. Durante 2013 e 2014, o CMS parou suas operações para realização de um grande período de manutenção e melhoramentos, denotado como LS1 (*long shutdown*). Em 2015 iniciou-se o *Run 2*, com colisões próton-próton a $\sqrt{s} = 13 \text{ TeV}$. Com o começo do *Run 2* uma nova etapa da análise de dados começou, buscas por novos fenômenos, especialmente buscas por Supersimetria, apareceram entre as mais importantes incumbências. Uma nova parada técnica inciou-se no final de 2018, LS2, marcando o fim do *Run 2* e o início das preparações para o *Run 3* esperado para começar em 2021 com energia de centro de massa projetada para $\sqrt{s} = 14 \text{ TeV}$ [84]. A figura 3.3 mostra a luminosidade integrada entregue pelo LHC nos diferentes períodos de tomadas de dados.

A análise exposta nesta tese foi realizada utilizando os eventos coletados durante a tomada de dados de 2016. Após a finalização de todas as etapas da estratégia de busca, a análise será estendida aos outros anos do *Run 2*. Portanto os resultados serão publicados

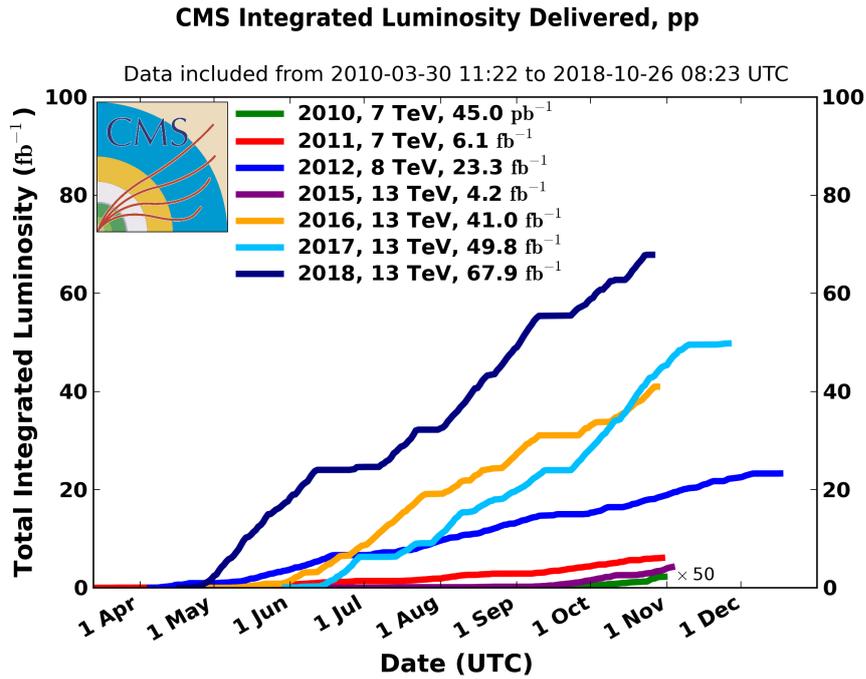


FIGURA 3.3 – Luminosidade integrada entregue pelo LHC ao experimento CMS em colisões pp durante os anos referentes aos *Runs* 1 e 2 [83].

utilizando os dados coletados nos anos de 2016, 2017 e 2018, conforme definido pelos *conveners* de SUSY do CMS.

3.1.1 Colisões próton-próton

As interações próton-próton podem ser divididas em três categorias gerais: interações elásticas que envolvem uma transferência de momento muito pequena, conservando a energia cinética dos prótons e mudando apenas sua direção; interações difrativas que envolvem uma transferência de momento um pouco maior, onde pelo menos um dos prótons se fragmenta produzindo um atividade hadrônica em ângulos muito pequenos com relação ao feixe; e interações inelásticas que são todas as interações hadrônicas não-difrativas em que pártons de cada próton interagem um com o outro. As interações inelásticas são as interações de maior interesse da física do LHC. Os processos principais que ocorrem no momento das colisões próton-próton inelásticas são listados abaixo.

- Espalhamento duro (*hard scattering*): interação entre dois pártons energéticos dos prótons envolvendo uma alta transferência de momento.
- Eventos subjacentes (*underlying events* - UE): interações entre pártons menos energéticos presentes nos prótons que poluem a assinatura real do espalhamento duro no detector, também chamados de interações de viés mínimo (*minimum bias*).

- Radiação de estado inicial (*Initial State Radiation* - ISR): quarks e glúons irradiados pelos dois pártons antes do espalhamento duro.

Sabe-se que os prótons são partículas compostas com uma estrutura interna que pode ser descrita em termos do modelo partônico melhorado pela QCD [85], segundo o qual a colisão inelástica é tratada como um espalhamento quase livre de constituintes dentro do próton, os pártons, compostos pelos quarks de valência e pelo “mar” de glúons e pares virtuais de quark-antiquark constantemente criados e aniquilados. O espalhamento duro, nos colisores hadrônicos, é descrito como uma interação entre pártons carregando uma fração x do quadrimomento do próton, onde x é conhecida como variável de escala de Bjorken. Um típico cálculo da seção de choque de um processo inclusivo $pp \rightarrow X$ consiste de um termo que descreve o espalhamento duro partônico que pode ser calculado usando QCD perturbativa, e fatores, f_i e f_j , que representam a densidade de probabilidade do párton ser encontrado dentro do seu respectivo próton carregando uma fração $x_{1,2}$ do seu momento (*parton distribution functions* - PDF), como mostrado na equação seguinte [86]:

$$\sigma(pp \rightarrow X) = \sum_{i,j} \int dx_1 dx_2 f_i(x_1, \mu_F^2) f_j(x_2, \mu_F^2) \hat{\sigma}_{i,j \rightarrow X}(x_1 x_2 s, \mu_R^2, \mu_F^2). \quad (3.4)$$

A soma é realizada sobre todos os pártons do estado inicial com momento longitudinal x_1 e x_2 , onde 1 e 2 denotam os dois prótons que interagem. O parâmetro μ_F , escala de fatoração, representa a escala na qual ocorre a separação entre a interação perturbativa dura e a evolução não-perturbativa dos pártons produzidos, já μ_R , escala de renormalização, é um parâmetro introduzido em QCD perturbativa para lidar com divergências ultravioletas. O termo $\hat{\sigma}_{i,j \rightarrow X}$ corresponde à seção de choque partônica avaliada nas escalas μ_F e μ_R .

3.2 O Detector CMS

O CMS é um detector de propósitos gerais construído para investigar um vasta gama de física, que envolve a busca pelo bóson de Higgs, medições do MP e buscas por físicas além do MP. O detector fica posicionado na França a uma profundidade de aproximadamente 100 m, perto da comunidade de Cessy. O CMS possui 21,6 m de comprimento, um diâmetro de 14,6 m, massa total de 12500 t. Ele possui uma estrutura cilíndrica, que consiste de várias camadas cilíndricas de subdetectores coaxiais com a direção do feixe (região do barril), fechada em ambas as extremidades por subdetectores em forma de discos perpendiculares ao feixe (região da tampa) de modo a garantir o fechamento hermético do aparato experimental.

A variedade de subdetectores presentes no CMS assegura uma identificação confiável de todas as partículas e uma medição precisa de sua energia e momento. O subdetector

mais interno é o *tracker*, responsável pela reconstrução da trajetória de todas as partículas carregadas. Em seguida estão dispostos os calorímetros eletromagnético e hadrônico, que medem a energia de elétrons, fótons e hádrons. Todos esses subdetectores estão dentro do solenoide supercondutor que aplica um campo magnético de alta intensidade, 3,8 T, ao longo da direção do feixe, que curva a trajetória das partículas carregadas e garante uma alta resolução do momento. A parte exterior do solenoide é composta por uma estrutura formada por camadas alternadas de ferro (*iron yoke*) saturado pelo campo magnético de retorno do solenoide. Entre as camadas de ferro estão acoplados os subdetectores que formam o sistema de múons, que ajuda a identificar e medir o momento dos múons produzidos na colisão. A Figura 3.4 mostra o detector CMS e seus principais componentes.

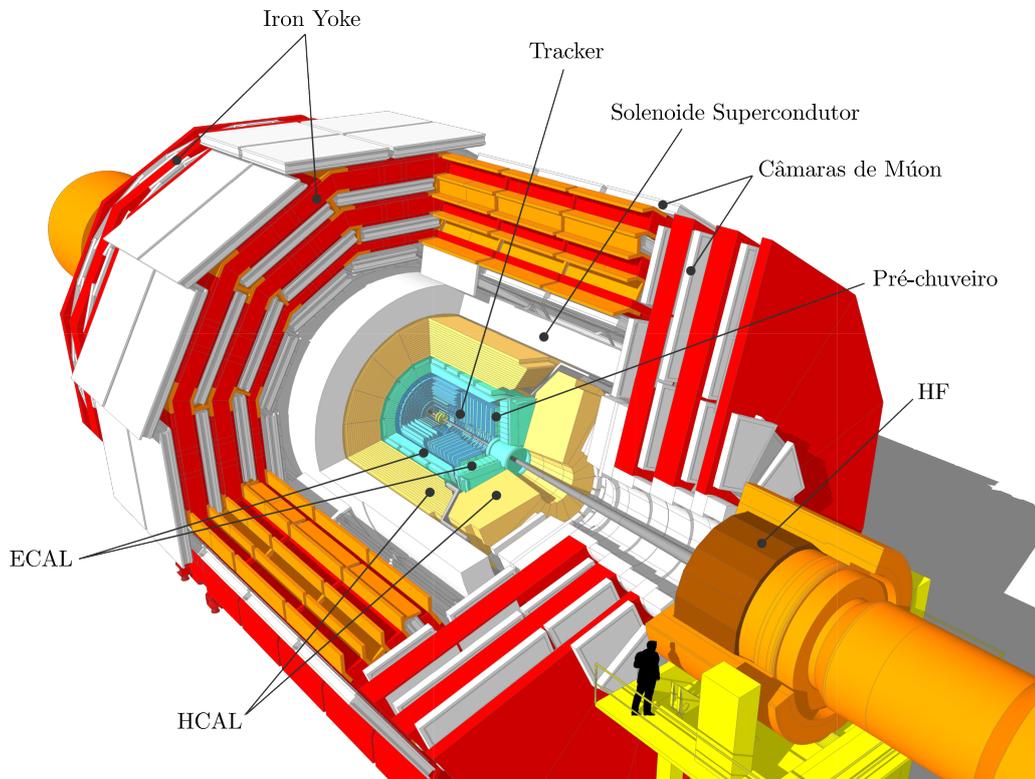


FIGURA 3.4 – Visão seccionada do experimento CMS identificando o posicionamento de seus principais componentes [87].

3.2.1 Coordenadas do Detector

Sendo a fração de momento longitudinal dos pártons que interagem na colisão desconhecida, é útil contar com uma descrição das trajetórias das partículas que é invariante de Lorentz sob *boosts* ao longo do eixo longitudinal. Para esse fim, a grandeza física conhecida como rapidez, y , é introduzida, ela é definida como

$$y = \frac{1}{2} \ln \left(\frac{E + p_z}{E - p_z} \right). \quad (3.5)$$

Com E sendo a energia da partícula e p_z a componente longitudinal do seu momento. No limite ultrarrelativístico, a rapidez converge para a definição da pseudorapidez, η ,

$$\eta = -\ln \left[\tan \left(\frac{\theta}{2} \right) \right], \quad (3.6)$$

sendo θ o ângulo polar. A pseudorapidez tem a vantagem de ser independente da energia da partícula e depende apenas das propriedades geométricas, enquanto ainda fornece diferenças, $\Delta\eta$, invariantes de Lorentz no limite ultrarrelativístico.

O sistema de coordenadas adotado pelo CMS tem a origem no ponto de colisão nominal dentro do experimento, o eixo y aponta para cima, o eixo x aponta radialmente em direção ao centro do LHC e o eixo z aponta ao longo da direção do feixe para as montanhas Jura. Na figura 3.5 estão desenhados o vetor \vec{p} representando o momento de uma partícula e sua projeção \vec{p}_T no plano x - y . Encontra-se também a definição das variáveis usadas para descrever a direção de propagação das partículas no detector, o ângulo azimutal, ϕ , e a pseudorapidez, η . Como o momento longitudinal do estado inicial não é conhecido e o

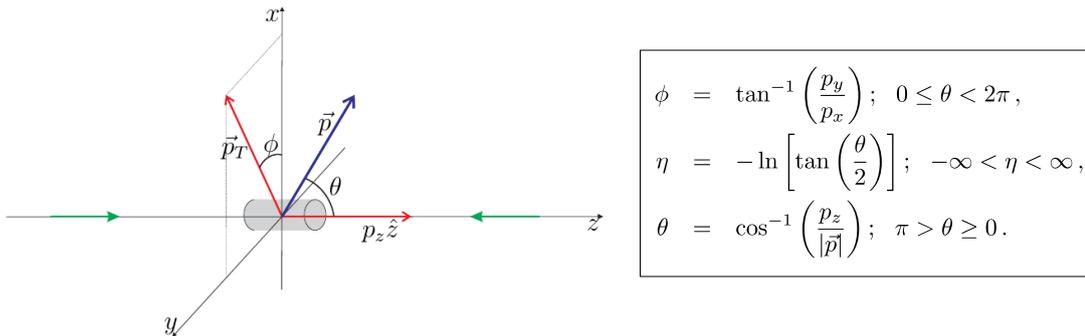


FIGURA 3.5 – Exemplo de um vetor \vec{p} representando o momento de uma partícula e como suas componentes estão relacionadas ao sistema de coordenadas do CMS.

momento transversal dos pártons de estado inicial pode ser desprezado, a conservação de momento implica que a soma vetorial do momento transversal de todas as partículas do estado final seja igual a zero.

3.2.2 Sistema de Trajetórias

O Sistema de Trajetórias (*Tracker*) [88] é o subdetector responsável pela medição das trajetórias de partículas carregadas e pela reconstrução dos vértices primários e secundários. Isso é alcançado usando um algoritmo complexo que ajusta os sinais (*hits*) deixados por essas partículas carregadas ao atravessar os sensores que compõem o *tracker*. Sendo o CMS um detector de propósito geral, o *tracker* é obrigado a ter um bom desempenho em vários domínios, como na resolução do p_T das partículas reconstruídas, reconstrução

eficiente de traços para partículas de baixo p_T e uma excelente resolução espacial. Devido à alta luminosidade alcançada na região de interação, em torno da qual o detector CMS está localizado, o fluxo de partículas produzidas é alto, cerca de 1000 partículas carregadas atravessam o detector a cada 25 ns. Para garantir que um sensor não seja atravessado por múltiplas partículas carregadas ao mesmo tempo, a granularidade do *tracker* deve ser a mais alta possível e os sensores devem ter um tempo de resposta rápido. Isso é especialmente verdadeiro para os sensores próximos ao tubo do feixe, pois a taxa diminui com a distância do feixe.

O *tracker* é baseado exclusivamente em muitas unidades de detectores de silício finamente segmentados. Uma partícula carregada passando por um sensor de silício causa ionização dentro do material. A tensão aplicada no sensor cria um campo elétrico dentro do material e os transportadores de cargas livres se movem para os eletrodos de leitura. A corrente induzida é usada para determinar a posição do *hit* da partícula [89]. Os *hits* das partículas em camadas consecutivas do *tracker* são usados para reconstruir as trajetórias das partículas (traços). O sistema de trajetórias do CMS atinge uma resolução de momento relativo $\sigma(p)/p$ de 1-5% para traços com momento de 1 GeV - 1 TeV.

O cilindro do *tracker* possui 5,8m de comprimento, 2,6 m de diâmetro e cobre uma faixa de pseudorapidez de $|\eta| < 2,5$. O detector de píxeis é instalado perto do tubo do feixe e é cercado pelos detectores de tiras. Ambos os componentes podem ser separados nas regiões central (barril) e dianteira (tampa), que se sobrepõem em $|\eta| \approx 1$.

O detector de píxeis consiste em módulos com píxeis finamente segmentados que podem lidar com o alto fluxo de partículas próximo ao ponto de interação. A descrição a seguir enfoca no *layout* do detector de píxeis da “Fase 0” (antes da atualização no final de 2016), figura 3.6. A região do barril consiste em três camadas cilíndricas em raios de 4,4 cm, 7,3 cm e 10,2 cm. Em cada região dianteira, duas camadas em forma de disco são instaladas em $|z| = 34,5$ cm e $|z| = 46,5$ cm, cada uma cobrindo a distância radial de 6 cm até 16 cm. Isso resulta em um total de 1440 módulos e um total de aproximadamente 66 milhões de píxeis. Os píxeis têm um tamanho de $100 \times 150 \mu\text{m}^2$, proporcionando uma resolução única de cerca de $9,4 \mu\text{m}$ nas coordenadas $r - \phi$ e $20 - 45 \mu\text{m}$ na direção z , dependendo do ângulo de incidência da partícula.

O detector de píxeis foi atualizado durante uma parada técnica estendida após a conclusão da tomada de dados em 2016 (*Extended Year-End Technical stop* - EYETS). A atualização foi realizada para lidar principalmente com o aumento da luminosidade instantânea do LHC e para melhorar o desempenho do *tracker* [90]. Uma quarta camada adicional na região do barril e um terceiro disco na região da tampa foram instalados, o que garante um rastreamento robusto e maior resolução de momento. Além disso, a camada mais interna do barril está mais próxima do feixe numa distância de 2,9 cm, proporcionando uma melhor resolução dos vértices e aumentando a eficiência na identificação

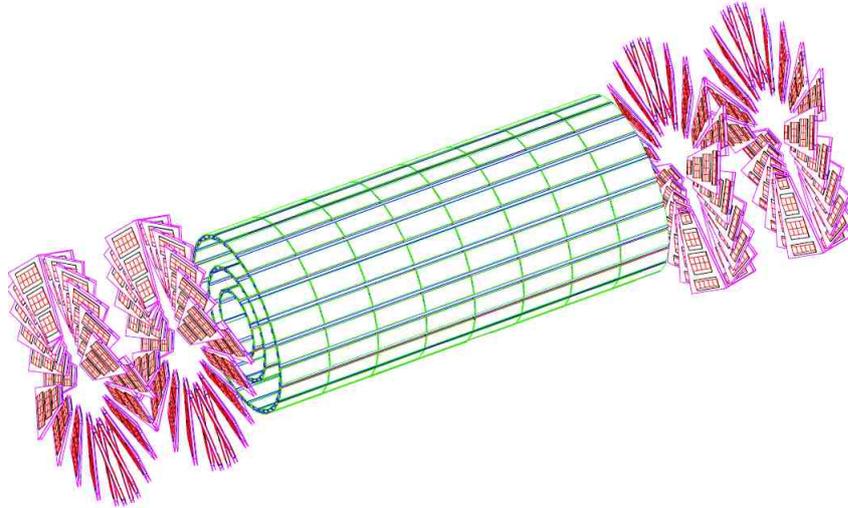


FIGURA 3.6 – Representação esquemática do detector de píxeis da “Fase 0”, composto por três camadas centrais e quatro tampas, duas de cada lado [91].

de b-jatos.

Nas regiões externas do sistema de trajetórias, o detector de tiras com menor granularidade é usado, já que o fluxo de partículas esperado é menor. O detector de tiras é dividido em quatro diferentes subsistemas como pode ser visto na figura 3.7. O barril interno do *tracker* (*tracker inner barrel* - TIB) é composto por quatro camadas concêntricas e cobre a região com raio de até 55 cm. A cobertura em η é prolongada graças ao disco interno do *tracker* (*tracker inner disk* - TID), que é composto de três discos de cada lado até $|z| = 118$ cm. O barril externo do *tracker* (*tracker outer barrel* - TOB), composto por 6 camadas, cobre de $r > 55$ cm até $r < 118$ cm. Finalmente, a tampa do *tracker* (*tracker endcap* - TEC) cobre a região entre $|z| = 124$ cm e $|z| = 282$ cm usando nove discos. No total, o detector de tiras contém 9,2 milhões de tiras, a resolução nas coordenadas $r - \phi$ varia de 23 a $53 \mu\text{m}$ e na direção z , a resolução fica na faixa de 230 – $530 \mu\text{m}$.

3.2.3 Calorímetro Eletromagnético (ECAL)

O Calorímetro Eletromagnético (ECAL) [92] envolve o sistema de trajetórias, estende-se até um raio de 177 cm e abrange a faixa de pseudorapidez $|\eta| < 3,0$. O ECAL consiste de vários elementos, o barril do ECAL (*ECAL barrel* - EB) e a tampa do ECAL (*ECAL endcap* - EE) são os principais componentes da ECAL. Além disso, os subdetectores *preshower* (ES) são instalados no lado interno do subdetector EE. Um esboço no plano $y - z$ do ECAL é mostrado na figura 3.8.

O ECAL tem como objetivo absorver e medir com precisão a energia de partículas com pouca capacidade de penetração na matéria, como elétrons e fótons. Igual a todos os subdetectores CMS, o *design* visa numa resposta rápida, granularidade fina e alta

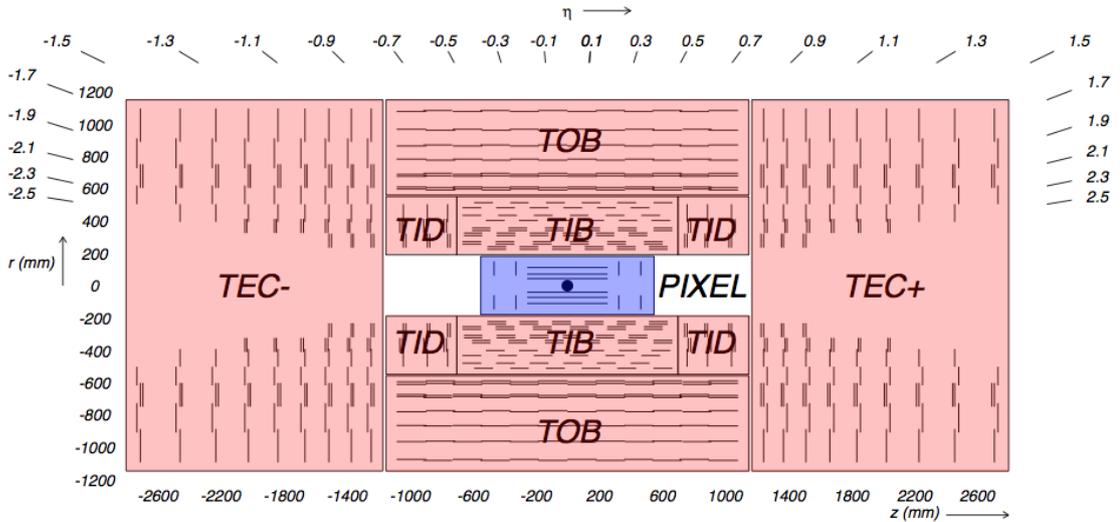


FIGURA 3.7 – Representação esquemática do sistema de trajetórias do CMS no plano $r - z$. A parte roxa representa o detector de píxeis e a parte rosa representa o detector de tiras [93].

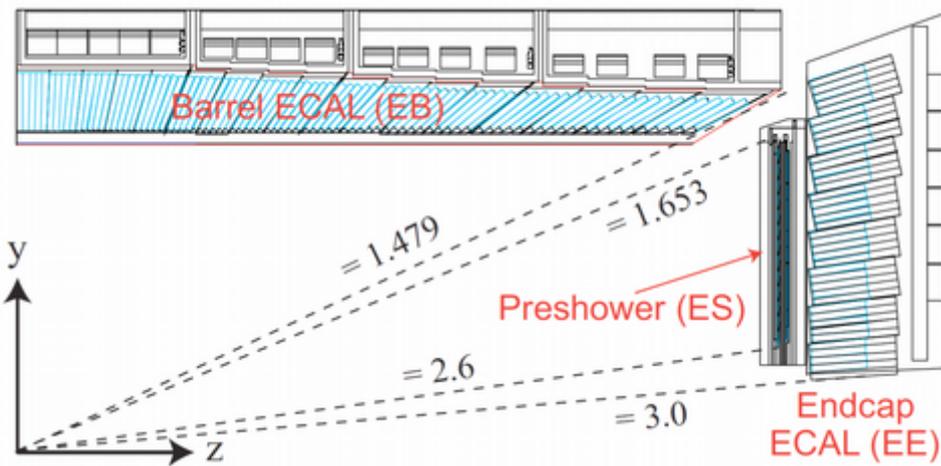


FIGURA 3.8 – Representação esquemática do Calorímetro Eletromagnético no plano $y - z$. As linhas tracejadas denotam a cobertura em pseudorapidez para a região de barril (EB) e região das tampas (EE). [94].

resistência à radiação. Isso motiva a escolha de um calorímetro hermético e homogêneo feito de cristais de tungstato de chumbo (PbWO_4), que serve como absorvedor e material ativo. Tem uma alta densidade de $8,3 \text{ g cm}^{-3}$, um curto raio de Molière $R_M = 2,2 \text{ cm}$ e um pequeno comprimento de radiação de $X_0 = 0,89 \text{ cm}$. O raio de Molière é uma medida da dimensão transversal de um chuvaireto eletromagnético. Da mesma forma, o comprimento de radiação serve como uma escala para descrever a escala longitudinal das cascatas eletromagnéticas. Ademais, tungstato de chumbo são cristais transparentes e emitem uma cintilação de luz azul-esverdeada dentro de um tempo de decaimento curto, o que permite coletar 80% da luz num intervalo de 25 ns. Estas propriedades do tungstato

de chumbo fazem dele um material ideal para um compacto e altamente granular ECAL. A luz de cintilação é detectada e amplificada por fotodetectores. Estes componentes também devem ser rápidos, tolerantes à radiação e insensíveis ao alto campo magnético. Estes requisitos levaram à escolha de fotodiodos de avalanche no barril (*avalanche photodiodes* - APD) e fototriodos a vácuo (*vacuum phototriodes* - VPT) na região da tampa.

O EB é instalado em um raio de 129 cm e cobre a região central do detector até $|\eta| < 1,479$. Contém 61200 cristais com uma seção transversal de aproximadamente $22 \times 22 \text{ mm}^2$, que estão dispostos em um arranjo nas coordenadas $\eta \times \phi$. Os cristais possuem comprimento de 230 mm. Os EE's são montados a uma distância de $|z| = 315,4 \text{ cm}$ e cobrem a faixa de $1,653 < |\eta| < 3,0$, o que deixa uma pequena lacuna entre o EB e o EE. Cada disco da capa contém 7324 cristais, dispostos em um arranjo nas coordenadas $x \times y$. Estes cristais têm uma seção transversal maior que a do EB ($28,6 \times 28,6 \text{ mm}^2$) e são ligeiramente mais curtos (220 mm). Todos os cristais no EB e EE estão instalados de forma que quase apontam para o ponto de interação, mas são inclinadas por um pequeno ângulo de modo a garantir que as partículas não atravessem exatamente entre dois cristais.

Os subdetectores *preshower* cobrem a região de pseudorapidez $1,653 < |\eta| < 2,6$ e principalmente visam identificar píons neutros, que quase exclusivamente decaem em dois fótons. Além disso, eles também melhoram a resolução espacial de elétrons e fótons. O ES é um calorímetro de amostragem com duas camadas ativas. O chumbo é usado como material passivo de alta densidade, que inicia chuveiros eletromagnéticos. Sensores de tira de silicone são colocados após o material passivo para medir a energia depositada e o perfil transversal do chuveiro. A espessura total do material do ES é de 20 cm, o que corresponde a cerca de 3 comprimentos de radiação.

3.2.4 Calorímetro Hadrônico (HCAL)

Entre o ECAL e o solenoide supercondutor fica o Calorímetro Hadrônico (HCAL) [95] que tem como objetivo medir a energia e a direção de jatos hadrônicos e garantir o máximo de hermeticidade do detector para a medição da energia transversa perdida. Seus componentes são o barril (HCAL *barrel* - HB), a tampa (HCAL *endcap* - HE), a parte externa (HCAL *outer* - HO) e a dianteira (HCAL *forward* - HF), cobrindo uma faixa de pseudorapidez até $|\eta| < 5,2$. Um esboço do *layout* completo é mostrado na figura 3.9.

O HB cobre a região central do detector até $|\eta| < 1,3$. Suas dimensões radiais são estritamente limitados pelo ECAL e o solenoide, então o foco está em um material com um comprimento de interação nuclear curto, λ , que é uma medida para a escala longitudinal de um chuveiro hadrônico. O material passivo é feito de latão, que também tem o benefício de não ser magnético. Para as camadas interna e externa, o aço inoxidável é usado para

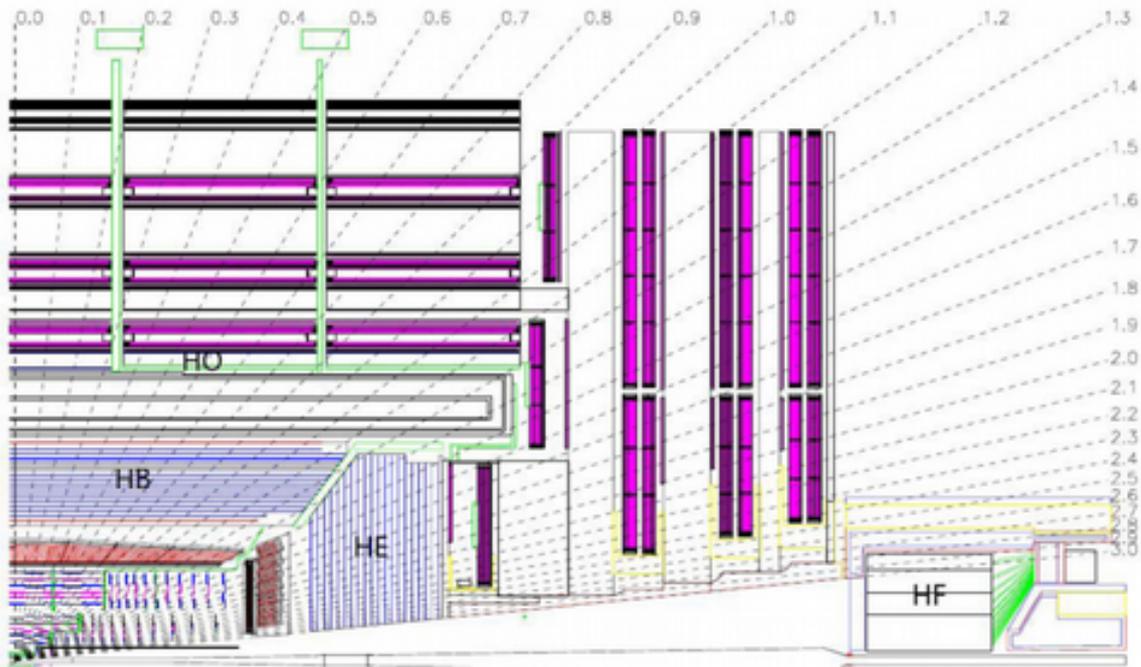


FIGURA 3.9 – Representação esquemática do Calorímetro Hadrônico no plano $r - z$ e seus subdetectores: HB, HE, HO e HF. Os segmentos roxos são partes do sistema de múons [96].

gerar estabilidade estrutural adicional. O material ativo consiste em telhas de plástico cintilador, que são segmentadas em $0,087 \times 0,087$ nas coordenadas $\eta \times \phi$. A luz emitida é processada por fibras de mudança de comprimento de onda, que conduzem a luz aos fotodiodos híbridos para leitura. A espessura de absorção total depende do ângulo de incidência e varia entre $5,82 \lambda$ em $|\eta| = 0$ e $10,6 \lambda$ para $|\eta| = 1,3$.

Como a espessura do HB não é suficiente para conter chuveiros de hádrons de alta energia, o solenoide é usado como um absorvedor adicional. O HO é montado fora do solenoide e contém outra camada de cintiladores para identificar chuveiros de início tardio e para medir a energia do chuveiro depositado após o HB. Na região mais central, onde a espessura é menor devido ao pequeno ângulo de incidência, outra camada de ferro e uma segunda camada de material cintilante é adicionada. Isso aumenta a profundidade total do sistema calorimétrico para um mínimo de $11,8 \lambda$ na região do barril do detector.

O HE segue um desenho semelhante ao HB, e amplia a cobertura do HCAL até $|\eta| < 3,0$. Como antes, são usadas camadas alternadas de latão e cintilador, mas a granularidade do cintilador aumenta com o aumento de $|\eta|$. A luz cintilante é processada de maneira semelhante ao HB. A espessura total do HE é de cerca de 10λ .

O HF é instalado na região dianteira do experimento CMS a uma grande distância, $|z| = 11,2 \text{ m}$ e abrange a faixa de pseudorapidez $3,0 < |\eta| < 5,2$. Isso significa que o HF está fora da cobertura do ECAL. O design se concentra em uma extrema resistência à

radiação do material ativo, necessária devido ao alto fluxo de partículas nessa região. O HF consiste em uma estrutura de absorvedor de aço com fibras de quartzo servindo como material ativo. Chuveiros de partículas carregadas que atravessam essas fibras produzem luz Cherenkov, que é detectada e amplificada por tubos fotomultiplicadores. Essa escolha tem a desvantagem de que o HF é principalmente sensível ao componente eletromagnético dos chuveiros, que leva a uma diminuição da resolução de energia devido a maiores flutuações estatísticas. A fim de distinguir chuveiros eletromagnéticos e hadrônicos, fibras de dois comprimentos diferentes são usadas já que chuveiros de fótons e elétrons geralmente têm uma profundidade menor que a de chuveiros de hádrons.

3.2.5 Solenoide Supercondutor

O solenoide supercondutor [97] é composto por quatro camadas de Nióbio-Titânio e possui 12,5 m de comprimento e 6 m de diâmetro interno. Possui um sistema de resfriamento utilizando hélio líquido com temperatura em torno de 4 K. Ele envolve o *tracker* e os calorímetros eletromagnético e hadrônico, gerando um campo magnético de 3,8 T no seu diâmetro interno, forte o suficiente para medir precisamente o momento de partículas carregadas e altamente energéticas. Seu campo de retorno é grande o suficiente para saturar uma “coleira” de ferro (*iron yoke*) com 1,5 m de espessura. A coleira consiste de uma estrutura de 12 lados, especialmente projetada para conter as estações de detecção de múons. A coleira na região do barril, é dividida ao longo do eixo do feixe em cinco anéis, onde cada anel é composto por três camadas de ferro integradas a quatro estações de múons. A espessura da camada mais interna é de 295 mm e a das outras duas é de 630 mm. Na região da tampa, a coleira de ferro é composta por três discos, aos quais são integradas as estações de múons frontais. Os dois discos internos possuem uma espessura de 600 mm para poderem suportar a força de atração gerada pelo campo magnético axial, e o disco externo possui espessura de 250 mm.

3.2.6 Sistema de Múons

O arranjo de subdetectores dedicados a detecção de múons, normalmente chamados de sistema de múons [98], é instalado na parte mais externa do detector. Os subdetectores são colocados entre as componentes da coleira de ferro. Isso é possível apenas porque os múons depositam uma quantidade mínima de energia no material do detector. O *design* fornece uma assinatura muito limpa dos múons no detector. Além disso, a informação do *tracker* é combinada com o múon do sistema múons, fornecendo uma excelente resolução do momento. Finalmente, o sistema de múons fornece uma maneira rápida e confiável de selecionar (*trigger*) eventos com múons, que podem surgir de muitos processos físicos

interessantes.

O sistema de múons consiste em cerca de 25000 m² de planos de detecção. Então, o foco do *design* está no baixo custo, confiabilidade e robustez. Estes critérios são cumpridos por detectores gasosos. Os detectores gasosos são baseados no conceito de que uma partícula carregada, ao passar pelo gás, ioniza seus átomos. Os elétrons e íons resultantes se movem ao longo do campo elétrico criado pelos cátodos e ânodos instalados e originam uma corrente que pode ser medida. São usados diferentes *designs* de detectores de gasosos que usam diferentes formas e alinhamentos dos eletrodos e operam sob diferentes tensões.

O sistema de múons cobre o intervalo de pseudorapidez $|\eta| < 2,4$ e consiste de um conjunto de subdetectores na região no barril e nas regiões das tampas. No total, três tipos diferentes de detectores gasosos são usados: as câmaras de tubos de arrasto (*Drift Tubes* - DTs), câmaras de tiras catódicas (*Cathode Strip Chambers* - CSCs) e câmaras de placas resistivas (*Resistive Plate Chambers* - RPCs). O tipo de detector é selecionado com base no fluxo esperado de múons e na força e homogeneidade do campo magnético campo em uma determinada posição. Uma visão geral do sistema múons do CMS é mostrada na figura 3.10.

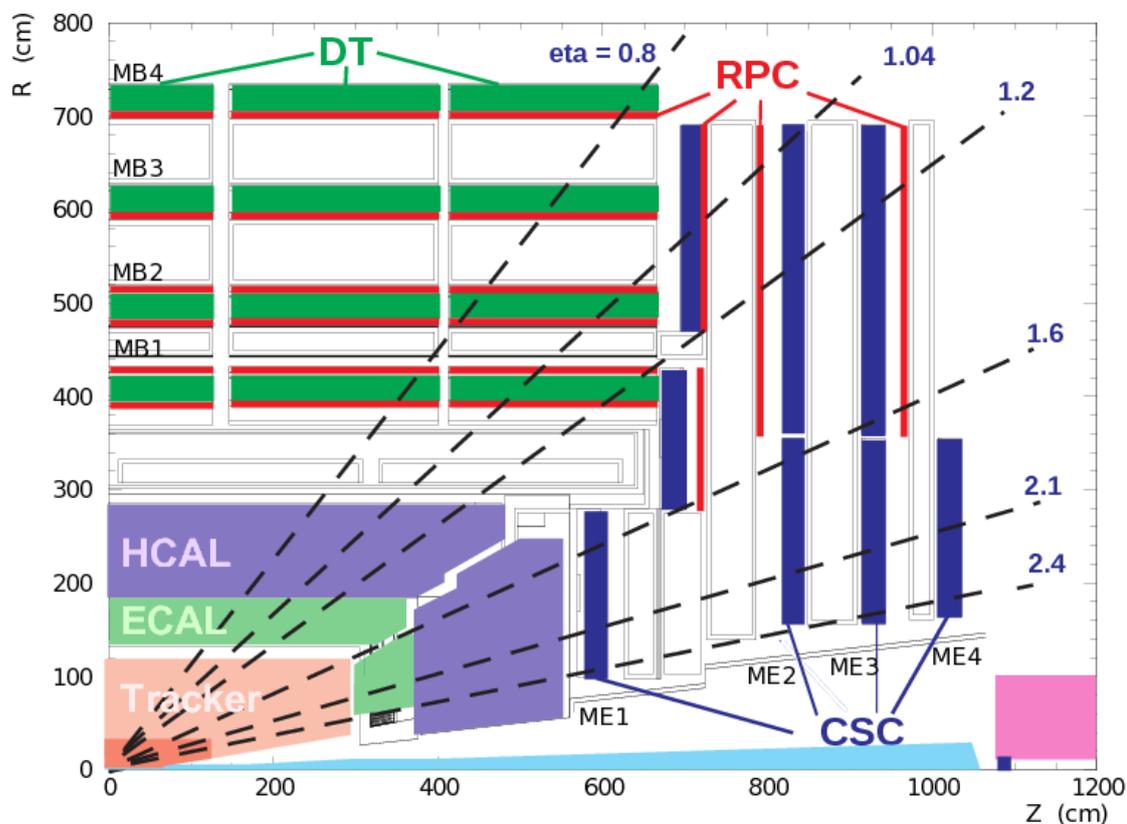


FIGURA 3.10 – Representação esquemática do Sistema de Múons no plano $r - z$ e seus subdetectores: DT, RPC e CSC [99].

Na região do barril, onde a taxa de múons é baixa e o campo magnético é uniforme,

DTs são instalados, cobrindo a faixa de pseudorapidez $|\eta| < 1,2$. Os DTs consistem em um tubo longo com quatro tiras de eletrodos nas paredes interiores e um fio (ânodo) no centro. Sua seção transversal é de $42 \times 13 \text{ mm}^2$ e possui $1,9 - 4,1 \text{ m}$ de comprimento, dependendo da posição. Os DTs estão organizados em quatro estações arranjadas dentro da coleira de ferro. As primeiras três estações contêm DTs que medem as coordenadas do múon no plano $r - \phi$, bem como DTs adicionais, que fornecem uma medida na direção z . A quarta camada não fornece uma medida da coordenada z . A distribuição e orientação dos DTs são escolhidas para oferecer uma boa eficiência no ajuste dos *hits* na formação dos traços de múons, enquanto mantém uma alta eficiência para rejeitar *hits* de *background* [100].

Na região da tampa do sistema de múons, onde temos altas taxas de múons e um grande e não uniforme campo magnético está presente, CSCs são usadas, cobrindo a faixa de $0,9 < |\eta| < 2,4$. A CSC contém tiras de cátodo que correm radialmente para fora e fornecem uma medição no plano $r - \phi$. Os fios de ânodos são alinhados perpendicularmente às tiras e determinam a coordenada η de um *hit*. O tamanho das CSCs é de 10° a 20° em ϕ , enquanto o comprimento é de $1,7 - 3,5 \text{ m}$. As CSCs também estão organizadas em quatro estações, o que proporciona uma reconstrução eficiente de múons e rejeição de *backgrounds* [101].

Os subsistemas DT e CSC podem selecionar eventos (*trigger*) baseados no momento de múons com uma alta eficiência, independente do resto do detector. No entanto, os tempos de resposta desses sistemas não são necessariamente rápidos o suficiente para atribuir um sinal ao cruzamento de pacotes correto (*bunch crossing*). Para esse fim, um segundo sistema de *trigger* complementar, consistindo de RPCs, é instalado em ambas as regiões de barril e tampa na faixa de pseudorapidez $|\eta| < 1,6$. Os RPCs consistem em duas placas paralelas com tiras comuns de leitura no meio e têm uma boa resolução de tempo, mas uma resolução de posição pior que as DTs ou CSCs. A redundância de RPCs melhora ainda mais as resoluções de tempo e momento, enquanto suprime o *background*. Além disso, as RPCs são instalados de maneira que os múons com $p_T \geq 5 \text{ GeV}$, que não atingem as camadas externas do sistema de múons ainda tenham uma alta eficiência de *trigger*.

A resolução de um *hit* é de aproximadamente $80 - 120 \mu\text{m}$ e $40 - 150 \mu\text{m}$ nos subsistemas DT e CSC, respectivamente. O subsistema RPC apresenta um desempenho significativamente pior com uma resolução de $800 - 1200 \mu\text{m}$, mas aumenta a resolução do tempo combinado para menos de 3 ns . A resolução do momento do sistema múons depende fortemente do momento do múon. Para múons com $p_T \leq 200 \text{ GeV}$ o *tracker* fornece a melhor resolução devido aos efeitos do espalhamento múltiplo, especialmente na coleira de ferro. No entanto, o sistema de múons melhora muito a resolução do momento para altos momentos, cerca de 5% para $p_T = 1 \text{ TeV}$.

3.2.7 *Triggers*

Nas colisões próton-próton que ocorrem no LHC, o intervalo em que os feixes se cruzam é de 25 ns. Considerando a alta luminosidade projetada de $10^{34} \text{ cm}^{-2}\text{s}^{-1}$, armazenar os dados associados a essa taxa enorme de eventos ($\sim 1 \text{ GHz}$) é inviável, além de que apenas uma pequena parte desses eventos é de interesse do programa de física do CMS, logo, é feita uma redução drástica na quantidade de eventos a fim de armazená-los. O trabalho do sistema de *triggers* é selecionar esses eventos interessantes para armazenamento, visto que os recursos computacionais limitam a capacidade de armazenamento a um máximo de aproximadamente 1000 eventos por segundo. Esta seção fornece uma visão geral do princípio de funcionamento dos *triggers* do CMS que selecionam os eventos em dois estágios, referidos como Nível Um (*Level One* - L1) e *Trigger* de Alto Nível (*High-Level Trigger* - HLT) [102,103,104].

O L1 *trigger* consiste em um hardware personalizado que decide dentro de $4 \mu\text{s}$ se um evento é aceito ou rejeitado. Essa curta latência é necessária, já que apenas um número limitado de eventos pode ser armazenado nos *buffers* de leitura dos componentes do detector. Para isso, apenas as informações do sistema de calorímetros e múons são consideradas, uma vez que a reconstrução de traços consome muito tempo. O *trigger* L1 começa criando os chamados *triggers* primitivos, separadamente dos dados oriundos dos calorímetros e do sistema de múons. Os *triggers* do calorímetro recebem informações brutas das células calorimétricas, processam as informações em paralelo, e retornam candidatos a elétrons, fótons e jatos, bem como observáveis globais como a energia transversa perdida. O *trigger* de múons processa a informação de todos os subsistemas de múons. Uma variedade de algoritmos de reconhecimento de padrões é usada para identificar os candidatos a múons e medir seu momento a partir da flexão das trajetórias no campo magnético da coleira de retorno. A etapa final do *trigger* L1 contém uma lista de 128 requisitos de seleção que são aplicados nos objetos identificados anteriormente. Se pelo menos um desses requisitos for atendido, o evento é passado para o HLT. O *trigger* L1 reduz a taxa de eventos para cerca de 100 kHz.

O HLT é um sistema de programas que realiza a seleção baseando-se em sofisticados algoritmos de reconstrução *offline*. O foco da reconstrução de eventos está na velocidade, assim, eventos desinteressantes são rejeitados o mais rápido possível sem a necessidade de reconstruir o evento completamente. O conceito principal do processamento de dados é o de “caminhos do HLT” (HLT *paths*), que são executados em uma ordem pré-definida. Cada HLT *path* consiste em uma série de algoritmos que reconstróem objetos físicos e colocam requisitos neles. Estes requisitos de seleção podem ser baseados em um único objeto físico como um múon de alto p_T , mas também observáveis mais complicados, como o isolamento leptônico, energia transversa perdida ou marcadores de b-jatos podem ser avaliados. Se

nenhum objeto físico ou observável potencialmente interessante for encontrado, o evento será rejeitado. A reconstrução do evento é refinada nas etapas subsequentes. Depois de cada passo, todos os eventos que não são de interesse para a análise de dados são rejeitados. Desse modo, algoritmos demorados como a reconstrução de traços e vértices são executados em apenas um pequeno subconjunto dos eventos selecionados pelo *trigger* L1. O HLT processa um evento em uma média de 175 ms e reduz a taxa de dados para cerca de 400 Hz. Para atender a taxa total do HLT, a cada *trigger path* é atribuído uma taxa máxima. Se um *trigger path* exceder essa taxa, os requisitos de seleção correspondentes são aumentados ou apenas o n-ésimo evento que passa os requisitos do *trigger* é selecionado, onde n é referido como a pré-escala. Todos os eventos selecionados são armazenados em disco e agrupados em fluxos de dados não exclusivos que contêm todos os eventos que passaram por *triggers* semelhantes.

3.3 Reconstrução de Eventos

3.3.1 Pileup

No LHC, que é um colisor de alta luminosidade instantânea, são produzidas várias colisões próton-próton dentro de um mesmo cruzamento de pacotes (*bunch crossing*). Se pelo menos uma dessas colisões resulta em uma interação interessante ao detector, o sistema de triggers é disparado e o evento é gravado. Em média, aproximadamente 23 colisões pp foram reconstruídas por evento em 2016. Assim, os sinais das colisões pp interessantes ao detector são poluídos com sinais de outras colisões que ocorreram em intervalos de tempo próximos, tais como depósitos de energia nos calorímetros e traços adicionais. Essa poluição é chamada de contaminação por *pileup* e demanda um grande esforço na reconstrução e análise dos eventos para ser mitigada. Ao analisar-se uma determinada interação, pode-se classificar as colisões de *pileup* em duas categorias:

- *Pileup* no tempo - colisões pp que ocorreram no mesmo cruzamento de pacotes;
- *Pileup* fora do tempo - colisões pp oriundas de outro cruzamento de pacotes.

O *tracker* é sensível apenas ao *pileup* no tempo, já os calorímetros e o sistema de múons também são sensíveis ao *pileup* fora do tempo.

Quando lidamos com efeitos de *pileup* em eventos simulados e reais [105,106,107,108], nos deparamos com dois tipos de distribuições, a distribuição de interações verdadeiras (*true interactions*) e a distribuição de interações observadas (*observed interactions*). A distribuição de interações verdadeiras para os dados é obtida multiplicando a média da luminosidade instantânea de todos os cruzamentos de pacotes dentro de cada seção de

luminosidade pela seção de choque inelástica total, e ponderando esse valor pela luminosidade integrada da seção de luminosidade. A distribuição de interações observadas, como o nome já diz, reflete o número de interações realmente observadas em cada cruzamento de pacotes, ela é obtida atribuindo a cada cruzamento de pacotes um valor aleatório obtido de uma distribuição de Poisson com média dada pelo número médio de interações da seção de luminosidade correspondente àquele cruzamento de pacotes.

As amostras de eventos simulados são geradas no CMS a partir de uma distribuição de interações verdadeiras pré-definida, que representa aproximadamente a distribuição esperada para aquele período de coleta de dados. Dessa forma, um evento simulado, que possua 22 interações verdadeiras, tem seu número de interações observadas obtido aleatoriamente de uma distribuição de Poisson com média igual a 22. Da mesma forma, o número de interações observadas nos cruzamentos de pacotes fora do tempo que contribuem para o *pileup* do evento, são obtidos dessa mesma distribuição. Esse procedimento tenta imitar o que acontece nos dados reais, onde tanto o evento de interesse quanto os eventos de *pileup* estão tipicamente dentro da mesma seção de luminosidade, logo, possuem praticamente a mesma luminosidade instantânea.

Uma boa modelagem dos efeitos de *pileup* nos eventos simulados necessita que a distribuição de interações verdadeiras na amostra da simulação seja igual à dos dados, e para isso é necessário fazer a reponderação de *pileup* dos eventos simulados. O procedimento para a reponderação de *pileup* dos eventos simulados divide bin por bin a distribuição de interações verdadeiras normalizada dos dados coletados pela distribuição normalizada dos eventos simulados. Isso resultará em uma distribuição de pesos de *pileup* dependente do número de interações verdadeiras, de modo que o peso de *pileup* atribuído a um determinado evento simulado dependa do número de interações verdadeiras usado na sua produção. Esse procedimento preserva o número total de eventos simulados. A figura 3.11 mostra um exemplo da reponderação de *pileup* em uma amostra de eventos de $t\bar{t}$ +jatos.

3.3.2 Fluxo de Partículas

A reconstrução do fluxo de partículas [109] de um evento visa reconstruir e identificar todas as partículas estáveis, por exemplo elétrons, múons, fótons, hádrons carregados e neutros. Ela é realizada através da combinação de informações de todos os subdetectores do experimento de modo a obter uma determinação acurada da direção, energia e tipo dessas partículas. Uma dada partícula pode interagir em um ou mais subdetectores como na figura 3.12 em que vemos exemplos da interação de diferentes tipos de partículas em uma seção transversal do detector CMS.

Os sinais deixados por uma partícula em diferentes detectores, chamados de elementos do fluxo de partícula, são conectados uns aos outros por um algoritmo de ligação que é

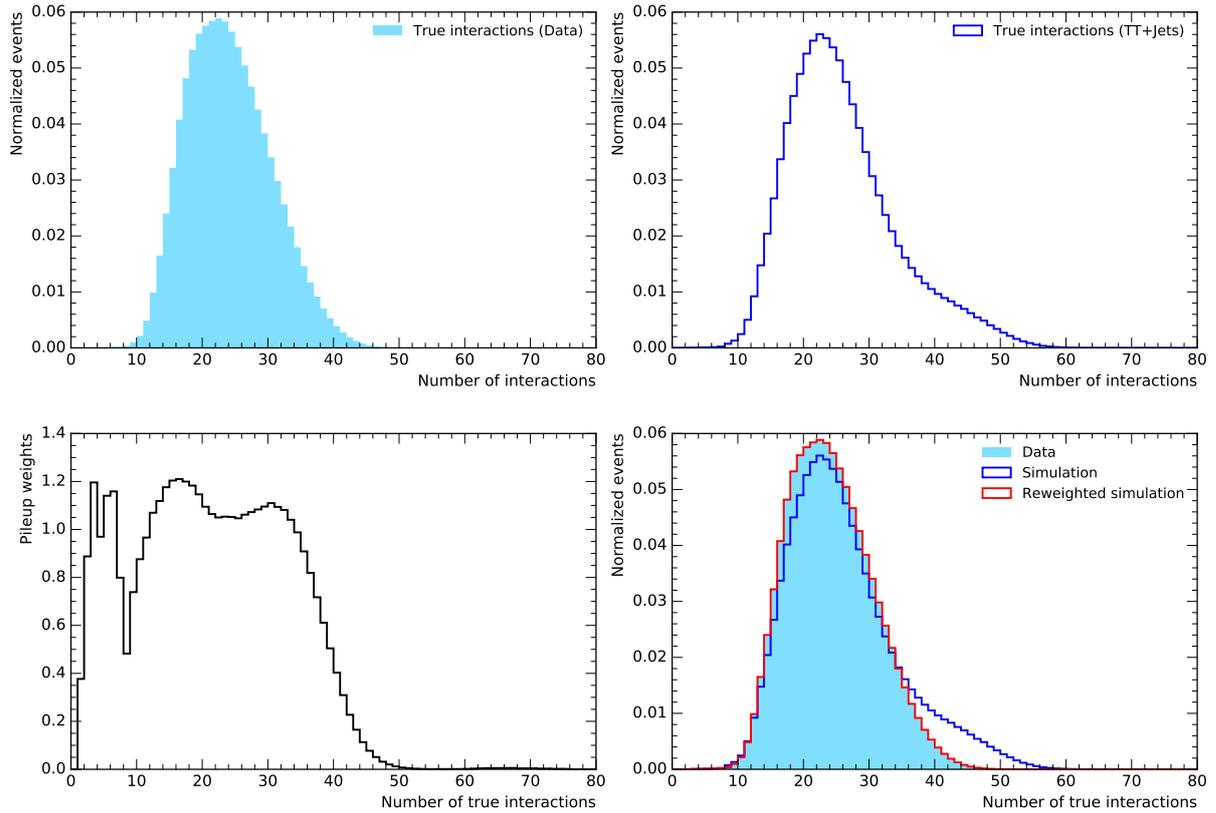


FIGURA 3.11 – Os gráficos acima estão normalizados e mostram as distribuições de interações verdadeiras dos dados coletados em 2016 (esquerda) e da amostra de $t\bar{t}$ +jatos produzida para 2016 (direita). O gráfico abaixo à esquerda mostra os pesos de *pileup*, que serão aplicados a cada evento simulado, em função do número de interações verdadeiras do evento. Ele é obtido pela divisão das distribuições acima. No gráfico abaixo à direita, estão reunidas as distribuições anteriores dos dados e da amostra de simulação, e também a nova distribuição da amostra de $t\bar{t}$ +jatos corrigida (vermelho) após serem aplicados os pesos de *pileup* nos eventos.

executado iterativamente para cada par de elementos no evento e define uma distância para quantificar a qualidade da ligação. O algoritmo de ligação, dessa forma, produz blocos de elementos ligados direta ou indiretamente. Entre traços e *clusters* (ou entre dois *clusters*) a distância de ligação é definida como a distância no plano (η, φ) entre a posição extrapolada do traço e a posição do *cluster* (ou entre a posição dos *clusters*). A distância de ligação entre um traço deixado no *tracker* e um traço deixado no sistema de múons é definida como o χ^2 do ajuste entre os dois traços, e esse par de traços é chamado de múon global. Para cada traço no sistema de múons, o único múon global retido é aquele formado pela ligação com o traço do *tracker* que retorna o menor valor de χ^2 .

A reconstrução e identificação de cada conjunto de partículas originárias de cada bloco de elementos é realizada pelo algoritmo de fluxo de partículas (*particle-flow algorithm*). Primeiramente, os múons globais são considerados “*particle-flow* múons” se o momento combinado é compatível com o momento determinado apenas pelo *tracker* dentro de três

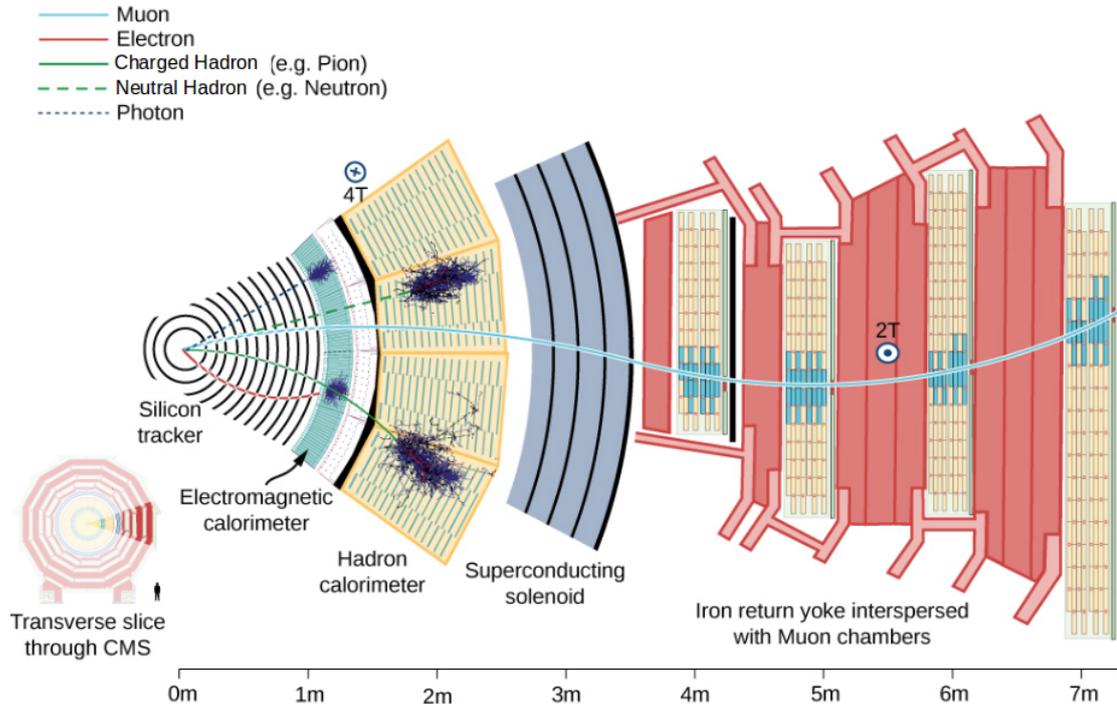


FIGURA 3.12 – Figura esquemática de uma fatia transversal do detector CMS, da região de interação do feixe ao subdetector de múons, mostrando a interação de partículas específicas com os subdetectores do CMS. O múon e o pión são carregados positivamente e o elétron é carregado negativamente. (crédito: David Barney/CERN)

sigmas de desvio. Os traços dos múons são então removidos e o algoritmo procede com a identificação de elétrons que têm como assinatura experimental a produção de traços curtos e a perda de energia por *Bremsstrahlung*¹. Os traços dos elétrons são reajustados na tentativa de seguir a trajetória até o ECAL, onde são combinados com os *clusters* de energia deixados no calorímetro de modo que cada elétron identificado dá origem a um “*particle-flow* elétron”. Após a reconstrução dos elétrons, todos os traços e ECAL *clusters* (incluindo os identificados como fótons da radiação *Bremsstrahlung*) são removidos dos processamentos seguintes. Critérios de qualidade mais rígidos são aplicados aos traços remanescentes que associados aos depósitos de energia nos calorímetros dão origem aos hádrons carregados. Os excessos de energia observados são calibrados nos calorímetros para uma correta estimativa da energia hadrônica depositada no ECAL e HCAL, e se o excesso for maior que a resolução de energia do calorímetro, ele é identificado como um fóton e/ou um hádron neutro. Ao final, os depósitos de energia não ligados a um traço são identificados como fótons no ECAL e hádrons neutros no HCAL.

¹*Bremsstrahlung* é a radiação eletromagnética produzida pela desaceleração ou deflexão de partículas carregadas (elétrons em geral) ao passar pelo forte campo elétrico de outra partícula (frequentemente prótons ou núcleos)

3.3.3 Jatos

Jatos são as assinaturas experimentais de pártons produzidos em processos de altas energias. Devido ao fato dos quarks e glúons não poderem existir livres na natureza por causa do confinamento de cor, uma vez produzidos eles iniciam um processo de formação de hádrons neutros de cor, chamado hadronização, que dá origem a um *spray* colimado de partículas estáveis chamado jato. As partículas que formam esse jato, à medida que se propagam no detector, deixam sinais nos subdetectores, como o *tracker* e os calorímetros. A figura 3.13 mostra um esquema simples de um jato produzido em uma colisão de prótons até a sua interação com os calorímetros.

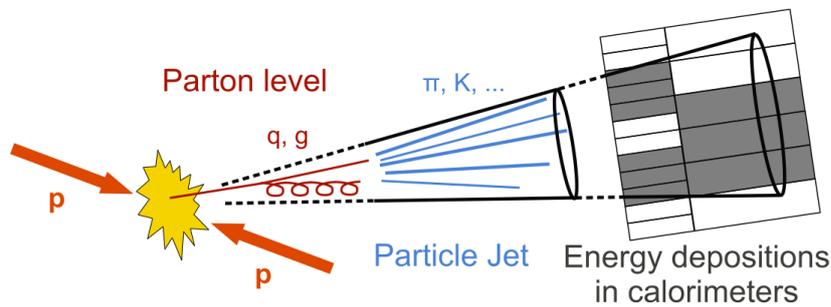


FIGURA 3.13 – Esquema de uma colisão entre prótons resultando na produção de um jato composto por partículas estáveis que deixam sinais nos subdetectores do CMS [110].

3.3.3.1 Algoritmos de reconstrução de jatos

Jatos são objetos compostos definidos por algoritmos de reconstrução de jatos [111,112] que recebem como entrada as partículas previamente reconstruídas pelo detector. Usualmente requer-se aos algoritmos de reconstrução que o conteúdo e o número de jatos não sejam alterados devido à fragmentação de um párton em outros dois (colinearmente seguro) e devido à adição de um glúon de baixa energia ao sistema (infravermelho seguro). Na literatura, os algoritmos que satisfazem esses requerimentos são chamados de IRC seguros (*Infra-Red and Collinear safe*).

A classe de algoritmos IRC seguros mais usada é a dos algoritmos de agrupamento sequencial que compreende os algoritmos K_t , Anti- K_t e Cambridge/Aachen. Esses algoritmos assumem que as partículas no interior do jato possuem pequenas diferenças no momento transversal, desse modo, eles agrupam as partículas baseando-se no espaço de momentos, resultando em jatos que têm flutuações em suas áreas no plano $(\eta - \phi)$. Os algoritmos possuem um método de funcionamento que baseia-se em duas variáveis de distância no espaço de momentos: a distância entre duas partículas i e j , d_{ij} , e a distância entre uma partícula i e o eixo do feixe (*beam axis*), d_{iB} ,

$$d_{ij} \equiv \min(k_{Ti}^a, k_{Tj}^a) \frac{R_{ij}^2}{R} \quad \text{e} \quad d_{iB} \equiv k_{Ti}^a. \quad (3.7)$$

Em que $R_{ij}^2 = (\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2$, ou seja, é a distância entre as duas partículas no plano $(\eta - \phi)$, e R é o raio, parâmetro que determina o tamanho dos jatos nesse plano. O expoente a é o parâmetro que diferencia os três diferentes algoritmos,

- K_t : $a = 2$;
- Anti- K_t : $a = -2$;
- Cambridge/Aachen: $a = 0$.

Os algoritmos de agrupamento sequencial buscam encontrar o valor mínimo no conjunto $\{d_{ij}, d_{iB}\}$, em que i e j são índices que varrem as partículas reconstruídas. Se d_{ij} é o mínimo, as partículas i e j são combinadas em uma nova “partícula” (ij) somando seus quadrimomentos, (ij) é inserida na lista e i e j são removidas. Se d_{iB} é o valor mínimo, a partícula i é removida da lista. O processo é repetido até que todas as partículas na lista façam parte de um jato.

Um raio grande permite ao jato capturar partículas hadrônicas suficientes para um cálculo acurado da massa e energia do jato, por outro lado, um raio pequeno é útil na redução da quantidade de eventos subjacentes (UE) e de *pileup* (PU), prevenindo a superestimação das massas e energias dos jatos. A contribuição de eventos de *pileup* pode ser reduzida removendo os hádrons carregados, que não são associados ao vértice primário do evento, da sequência de agrupamento. Esse método é denominado Subtração de Hádrons Carregados (*Charged Hadron Subtraction - CHS*).

O algoritmo K_t , $a = 2$, prefere agrupar partículas pouco energéticas primeiro, resultando em uma área que flutua consideravelmente durante o processo. O algoritmo Cambridge/Aachen, $a = 0$, tem variáveis de distância independentes do momento, de modo que sua área flutua um pouco. Por fim, o algoritmo Anti- K_t , $a = -2$, prefere agrupar partículas muito energéticas primeiro, logo, a área flutua muito pouco durante o agrupamento. Ele é o algoritmo que produz jatos com formas mais bem definidas, como pode ser visto na figura 3.14, e além disso é o menos suscetível a UE e PU. A reconstrução padrão de jatos realizada pelo CMS, e utilizada na análise, faz uso do algoritmo Anti- K_t com $R = 0,4$.

3.3.3.2 Correções à energia dos jatos

A relação entre o momento reconstruído de um jato com o do párton inicial é chamada de resposta da energia do jato que é afetada por uma soma de efeitos experimentais nas medições. Entre estes efeitos estão a resposta não linear do detector, a perda de energia dos constituintes dos jatos devido às interações com os materiais do detector, ineficiências na reconstrução das partículas e contribuições de *pileup*. A compreensão detalhada da

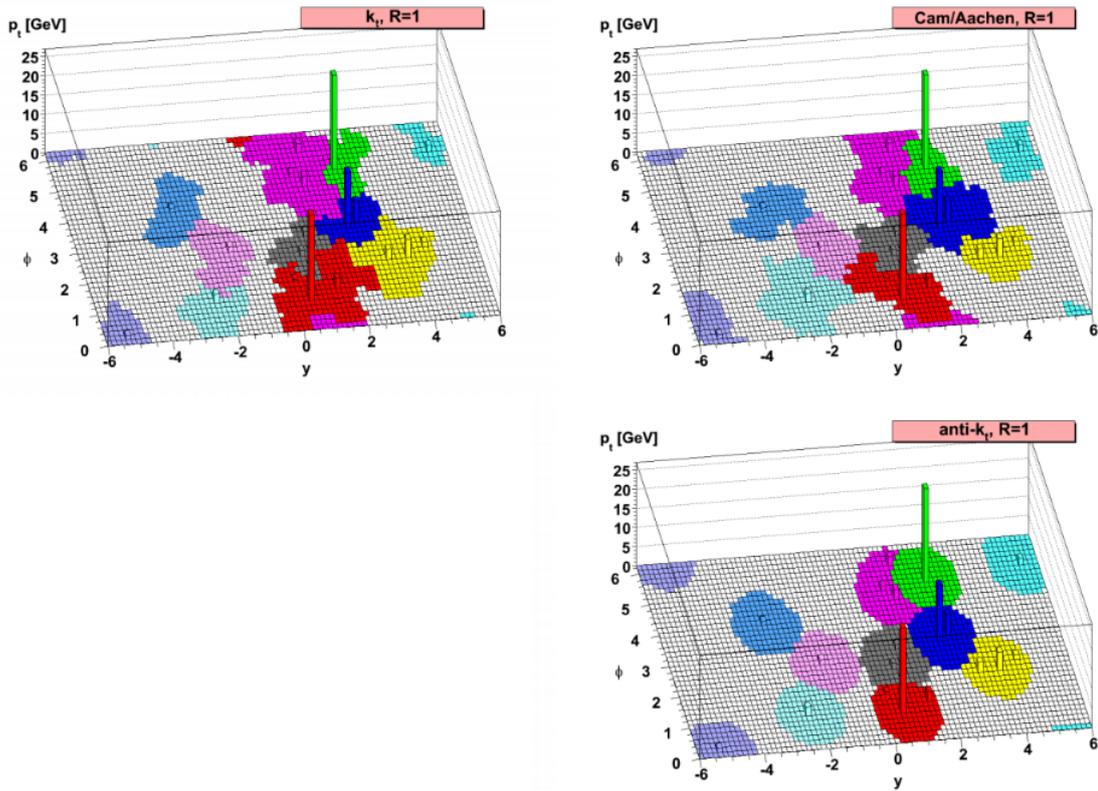


FIGURA 3.14 – Exemplo do desempenho dos três algoritmos de agrupamento sequencial, com $R = 1$, na resolução de jatos [112].

escala de energia do jato é de importância crucial para muitas análises físicas, e muitas vezes é uma componente importante de sua incerteza sistemática.

Na figura 3.15 temos o comportamento, para eventos simulados, da resposta dos jatos PF² com subtração de hádrons carregados. Para jatos centrais com $p_T \geq 60$ GeV, a resposta é quase constante em função de p_T com uma média de cerca de 0,95. O deslocamento da unidade é causada por hádrons neutros, que têm uma baixa resposta de apenas cerca de 0,6 e contribuem em média com 15% da energia total dos jatos [113]. Para jatos com $p_T \leq 30$ GeV, a resposta é significativamente menor devido à aceitação do HCAL. Na capa e na região frontal próxima ao feixe, uma dependência mais forte com respeito ao momento transversal pode ser observada.

Para corrigir esses desvios observados, correções à energia dos jatos (*jet energy corrections* - JEC) são aplicadas de modo que a escala de energia dos jatos (*jet energy scale* - JES) é idêntica para os jatos reconstruídos e a nível de partículas. Essas correções são derivadas em função de p_T e/ou η , e são determinadas em uma abordagem fatorizada. A seguir, os passos desse procedimento são explicados [114,115].

1. **Correção de *Pileup* (L1):** é a primeira correção da cadeia de correções e é destinada a remover a energia oriunda da contribuição de *pileup*. As correções de

²PF denota que o objeto foi reconstruído utilizando algoritmo de fluxo de partículas.

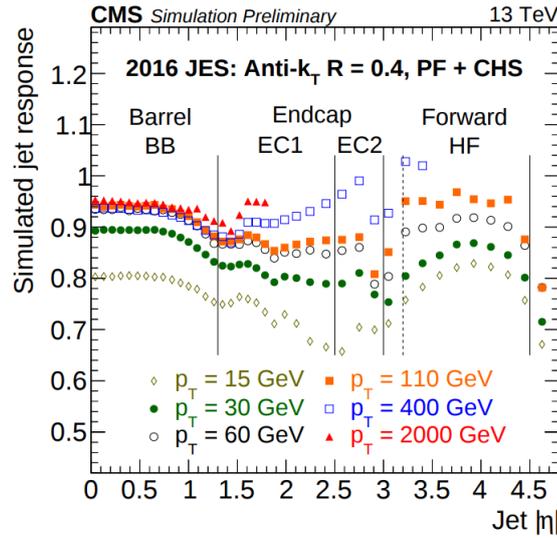


FIGURA 3.15 – Resposta da energia dos jatos no detector CMS em função de $|\eta|$ para jatos com diferentes momentos transversos [113].

pileup são determinadas a partir da simulação de uma amostra de eventos de QCD com dois jatos, processados com e sem sobreposição de *pileup*. O fator de correção derivado é aplicado nos dados e na simulação.

2. **Correções relativas (L2) e absolutas (L3):** estas correções são determinadas após as correções de L1 serem aplicadas. Os alvos da correção são a escala absoluta da resposta média dos jatos. As correções são derivadas de eventos QCD simulados com dois jatos, que têm a vantagem de uma alta precisão estatística para todas as propriedades cinemáticas dos jatos e são aplicados em dados e simulação. As correções são derivadas em função do p_T e η do jato e tornam a resposta uniforme sobre essas duas variáveis.
3. **Correções residuais (L2 e L3):** as correções residuais de L2 e L3 são feitas para eliminar quaisquer diferenças remanescentes na escala de energia do jato entre os dados e a simulação. Na correção residual (L2), eventos com dois jatos são utilizados onde um dos jatos deve estar na região central do detector ($|\eta| < 1, 3$). As correções residuais (L3) também visam a escala absoluta de energia dos jatos e são derivados de uma abordagem combinada usando eventos $Z(\mu\mu) + \text{jatos}$, $Z(ee) + \text{jatos}$ e $\gamma + \text{jatos}$. Ambas as correções são derivadas com respeito ao objeto de referência, ou seja, o jato central ou o bóson Z reconstruído ou o fóton, usando os métodos de balanço de p_T e a fração da projeção da energia transversa perdida (*missing transverse energy projection fraction* - MPF).

As correções de energia do jato são determinadas com alta precisão. Na região do barril do detector, a incerteza total das correções obtidas é inferior a 2%, porém pode ser de até 5% para jatos na região de transição entre o *endcap* e o calorímetro *forward* [113].

A resolução de energia do detector é inferior para jatos em comparação com outros objetos como léptons ou fótons. A resolução de energia do jato (*jet energy resolution* - JER) é definida como a largura Gaussiana da resposta da energia do jato. O JER é determinado usando a abordagem do equilíbrio de p_T [116] em eventos semelhantes aos eventos de dijetos, onde a atividade hadrônica adicional de um terceiro jato é extrapolada para zero. Esse método tem a vantagem de uma alta precisão estatística. Como a simulação tende a subestimar a largura da resposta da energia do jato, uma correção adicional é derivada e aplicada aos eventos simulados. A magnitude do fator de correção é menor que 20% na região central do detector, mas pode chegar a 80% na região de transição [117].

3.3.4 Energia Transversa Perdida

O detector CMS pode detectar quase todas as partículas estáveis ou de vida longa que são produzidas nas interações próton-próton. As exceções são neutrinos e potenciais partículas de física além do MP que só interagem fracamente. No entanto, podemos obter informação sobre essas partículas utilizando a conservação de momento, a qual exige que a soma vetorial do momento transversal de todas as partículas que são produzidos na interação pp seja igual a zero. Assim, partículas que não são detectadas geram um desequilíbrio no momento transversal que é dado por

$$\vec{E}_T^{\text{miss}} = \vec{\cancel{E}}_T \equiv \vec{p}_T^{\text{miss}} = - \sum_i \vec{p}_{T,i}, \quad (3.8)$$

onde a soma inclui todas as partículas dos estados finais que foram detectadas. Geralmente, as análises estão interessadas apenas no valor absoluto $E_T^{\text{miss}} = |\vec{E}_T^{\text{miss}}|$ que é a energia transversa perdida (também denotada como MET). Usa-se MET em medições de bósons W, quarks pesados e léptons tau, pois essas partículas podem decair em neutrinos [118]. Ademais, muitos modelos de física além do Modelo Padrão predizem a existência de partículas ou algo mais que são invisíveis e podem carregar momento, como modelos de Matéria Escura e supersimétricos. Por esse motivo, MET é também bastante usada para testar esses modelos.

A energia transversa perdida é sensível a qualquer medição incorreta das partículas visíveis, bem como a depósitos de energia adicionais provenientes de *pileup*, ruídos do detector, etc. Correções podem ser aplicadas visando suprimir qualquer efeito que cause desvio com respeito ao valor de E_T^{miss} a nível de partículas. A correção mais comum é denominada correção de tipo 1 [119], que propaga a correção à energia dos jatos (JEC) para E_T^{miss} , de maneira que

$$\vec{E}_T^{\text{miss, corr.}} = \vec{E}_T^{\text{miss}} - \sum_{\text{jatos}} (\vec{p}_{T, \text{jato}}^{\text{corr.}} - \vec{p}_{T, \text{jato}}), \quad (3.9)$$

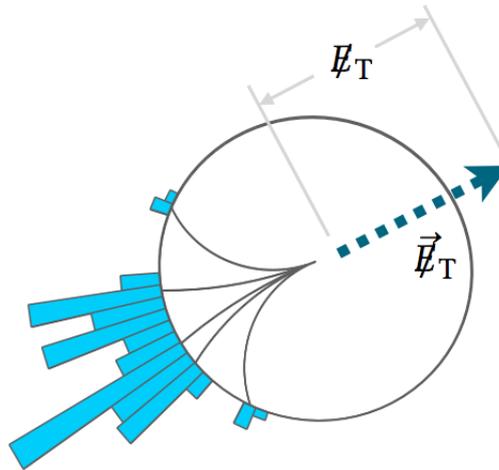


FIGURA 3.16 – Representação esquemática da definição de \vec{E}_T^{miss} e E_T^{miss} [112].

em que “corr” denota o valor corrigido. A soma se estende por todos os jatos com uma fração de energia eletromagnética abaixo de 0,9 e $p_T > 15$ GeV.

3.3.5 Filtros de MET

Existe uma variedade de efeitos que mesmo sendo raros afetam a qualidade da reconstrução da energia transversa perdida e podem gerar “caudas” artificiais na distribuição E_T^{miss} . A identificação e rejeição de eventos afetados por esses efeitos é assegurada por uma sequência de filtros dedicados que são tipicamente sensíveis a certas detecções ou mau comportamentos relacionados à reconstrução de eventos [120,121].

- **Filtro da aréola do feixe:** a aréola do feixe é um conjunto de partículas produzidas por interações do feixe de prótons com gás residual ou com o tubo do feixe, viajando junto com o feixe com um raio de até 5 metros.
- **Filtro de ruído do HBHE:** as telhas do cintilador do barril (HB) e as tampas de extremidade do calorímetro hadrônico (HE) gravam sinais anômalos esporádicos (ruído) a uma taxa fixa independente das condições do feixe. Esses depósitos de energia artificial têm um padrão geométrico característico, multiplicidade de canais e forma de pulso.
- **Filtro do *trigger* primitivo do Ecal:** alguns canais do Ecal podem parar de funcionar e são mascarados durante a reconstrução. Em cerca de 70% desses canais, a energia pode ser estimada a partir do *readout* do L1 *trigger* primitivo. No entanto, esses *trigger* primitivos podem ficar saturados, de modo que a energia gravada é provavelmente subestimada.

- **Filtro do vértice primário:** os eventos devem ter pelo menos um vértice reconstruído que satisfaça determinados critérios de qualidade, por exemplo ter que estar na região central do detector.
- **Filtro de supercristais ruins do Ecal:** foram identificadas três regiões ruins dos supercristais nas extremidades do calorímetro eletromagnético, que geram anormalmente altas energias, por isso os eventos com depósitos de energia em uma dessas regiões são removidos.
- **Filtro de múons ruins:** em alguns eventos, a qualidade do traço do múon é boa o suficiente para que o múon seja reconstruído como um candidato PF, mas o p_T do múon é amplamente superestimado devido a uma má reconstrução.

3.3.6 Vértices Primários e Secundários

A existência de vértices secundários é uma característica importante da assinatura dos eventos de sinal buscados. O algoritmo de associação de jatos a vértices, implementado nesta análise, faz uso dos vértices primários (*primary vertices* - PV's) e vértices secundários (*secondary vertices* - SVs) reconstruídos centralmente pelo CMS. A figura esquemática 3.17 mostra a visão no plano transversal de um evento que possui um vértice secundário reconstruído. Os algoritmos utilizados na reconstrução dos vértices são descritos a seguir [94].

Os vértices primários refletem as posições correspondentes aos pontos das colisões próton-próton de onde os traços se originam, tanto da colisão principal quanto das colisões de *pileup*. O localizador de vértices primários usa todos os traços reconstruídos no evento. Seus quatro passos principais são [123]:

- seleção de traços, onde requer-se que a significância do parâmetro de impacto transversal do traço (razão entre o parâmetro de impacto e sua incerteza) seja menor que 3, e que seu p_T seja maior que 1,5 GeV;
- formação de agrupamentos de traços, baseando-se na coordenada z dos seus pontos de maior proximidade com relação à linha do feixe, de modo que os traços distanciados por menos de 1 mm são agrupados;
- ajuste de vértices primários para cada um dos agrupamentos, removendo recursivamente traços incompatíveis com o vértice candidato, sendo o corte de compatibilidade padrão igual a 5%;
- exclusão de maus ajustes em que a probabilidade χ^2 seja $P(\chi^2_{\text{ajuste}}) < 1\%$ e de vértices incompatíveis com a linha do feixe com probabilidade menor que 1%.

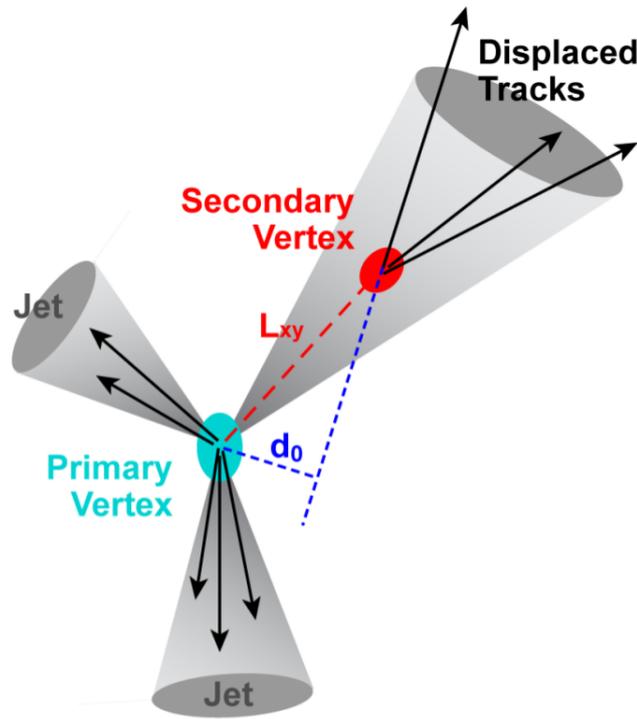


FIGURA 3.17 – Exemplo de um evento com a reconstrução de um vértice secundário. L_{xy} é a distância entre o vértice secundário e o vértice primário no plano transversal e d_0 é o parâmetro de impacto transversal do traço, definido como a distância de maior proximidade entre um traço e o vértice primário no plano transversal [122].

Todas as probabilidades de compatibilidade são computadas supondo resoluções Gaussianas. A compatibilidade com o eixo do feixe é calculada considerando a seção transversal do feixe com uma largura de $15 \mu\text{m}$ nas coordenadas x e y . Após serem localizados, os vértices são ordenados em uma ordem decrescente da soma dos p_T^2 dos traços associados, dessa forma, o vértice da primeira posição é geralmente o mais provável de refletir a posição da colisão principal. A resolução na coordenada- x dos vértices primários reconstruídos está tipicamente em torno de $10 - 25 \mu\text{m}$.

Os vértices secundários refletem as posições correspondentes ao decaimento de partículas com tempo de vida razoavelmente longo, por exemplo o decaimento dos quarks b . O localizador de vértices TKVF (*Trimmed Kalman Vertex Finder*) busca iterativamente os vértices candidatos entre os traços do evento [124]. Na primeira interação, o algoritmo TKF (*Trimmed Kalman Fitter*) [125] é aplicado no conjunto completo de traços gerando como saída um candidato a vértice e o conjunto de traços incompatíveis com aquele vértice. Durante as interações subsequentes, o mesmo procedimento é aplicado sobre o conjunto de traços incompatíveis identificados na interação anterior. O TKVF é sensível a vértices primários e secundários, então a seguinte seleção é usada para filtrar os candidatos a vértice secundário,

- A distância no plano transverso entre o vértice primário (identificado pelo algoritmo de vértices primários) e o candidato a vértice secundário, L_{xy} , deve ser maior que $100 \mu\text{m}$ e menor que 2 cm ;
- A distância no plano transverso, entre o vértice primário e o candidato a vértice secundário, dividida por sua incerteza tem que ser maior que 3, $L_{xy}/\sigma_{L_{xy}} > 3$;
- A massa invariante total dos traços associados ao vértice deve ser menor que $6,5 \text{ GeV}$.

A resolução típica do decaimento de quarks b é aproximadamente $200 \mu\text{m}$ quando está na direção do barril e $580 \mu\text{m}$ na direção da tampa.

3.3.7 Marcadores de b-jatos

Os b-jatos são jatos que surgem do processo de hadronização de quarks *bottom* (quarks b). A identificação de b-jatos desempenha um papel importante nas análises físicas no LHC. Os jatos de quarks b são particularmente úteis para identificar quarks *top* (quarks t), visto que eles quase exclusivamente decaem por $t \rightarrow bW$. Além disso, muitas partículas previstas por modelos além do Modelo Padrão são previstas para decair principalmente em quarks da terceira geração.

Uma variedade de algoritmos têm sido desenvolvidos para identificar b-jatos, cada uma deles resulta em um marcador de b-jatos diferente. Todos esses algoritmos exploram propriedades especiais dos quarks b, que são baseados no fato de que quarks b têm uma vida longa, pois eles só podem decair fracamente em quarks das duas primeiras gerações. Como consequência, os quarks b hadronizam formando hádrons de vida longa. Esses hádrons se propagam através do detector e decaem após alguns milímetros ou até mesmo centímetros dependendo do seu momento. Os produtos do decaimento são observados como traços deslocados (*displaced tracks*) que podem ser usados para reconstruir um vértice secundário. Outras propriedades características são uma alta massa do jato formado pelos produtos do decaimento do hádron b e uma alta probabilidade, de cerca de 20%, que um elétron ou um múon esteja contido no jato produzido no processo de decaimento fraco. Infelizmente, jatos de quarks c têm características semelhantes por isso é especialmente difícil distingui-los de jatos b.

A descrição a seguir enfoca o algoritmo CSVv2 (*Combined Secondary Vertex version 2*) [126], o qual é utilizado na análise apresentada nesta tese. Primeiro, os vértices secundários são reconstruídos pelo algoritmo Localizador de Vértice Inclusivo (*Inclusive Vertex Finder* - FIV). O algoritmo FIV usa todos os traços que têm um momento transverso mínimo de $p_T > 0,8 \text{ GeV}$ e que passam certos requisitos de qualidade. Desses traços, são selecionados “traços sementes” com base na magnitude e importância do parâmetro de impacto. Os

traços de sementes são usados para selecionar mais traços, que têm uma pequena distância e ângulo em relação à semente. Os agrupamentos de traços são então passados para o algoritmo AVF que reconstrói candidatos a vértices. Uma variedade de seleções é aplicada nos candidatos a vértices, para remover os candidatos que têm traços comuns uns com os outros ou com o vértice primário, candidatos a vértices mal reconstruídos, etc. Os vértices candidatos são então atribuídos a jatos reconstruídos, que por sua vez são passados para o algoritmo CSVv2.

O algoritmo CSVv2 é baseado em uma rede neural que combina as informações de variáveis do traço, vértice secundário e jato. Caso nenhum vértice secundário seja atribuído a um jato, o algoritmo CSVv2 pode criar pseudo-vértices candidatos baseados em traços com grande significância do parâmetro de impacto. Finalmente, uma probabilidade é derivada para cada jato, que varia entre zero e um. Esse valor do discriminador tende a ter valores altos para jatos de quarks b e valores baixos para jatos leves oriundos de quarks (u , d , s) ou glúons. Jatos de quarks c preenchem a região intermediária.

4 Seleção e Tratamento de Dados

Neste capítulo são descritos os aspectos gerais da análise: amostras utilizadas, seleção de dados reais feita pelos *triggers* de busca, características dos eventos de sinal, variáveis usadas na discriminação entre sinal e *background*, definição da região de busca e modelagem das simulações de *background*.

4.1 Variáveis Globais

Serão chamadas de variáveis globais todas as variáveis relevantes na análise que descrevem aspectos gerais do evento. Estas variáveis serão usadas no estudo dos *triggers* de busca, definição da região de busca e na verificação da qualidade das simulações de *background*. Posteriormente, elas se juntarão às variáveis regionais¹ para compor o conjunto total de variáveis a serem usadas na análise multivariada do capítulo 5. A seguir temos a descrição de cada uma das variáveis globais.

Os eventos de sinal não possuem léptons, sendo caracterizados por sua informação hadrônica, energia transversa perdida e pelos vértices reconstruídos. Os jatos usados na análise são selecionados seguindo critérios de qualidade definidos pelo experimento CMS [130]. Em seguida, são eliminados os jatos reconstruídos erroneamente, oriundos de elétrons e múons, e por fim são aplicados cortes nas grandezas η e p_T dos jatos para que estejam em intervalos onde é verificado um bom acordo entre os dados reais e as simulações. Os jatos considerados na análise possuem $|\eta| < 2,4$, cortes em limiares inferiores do p_T dos jatos, dependentes da variável considerada, e satisfazem o critério de qualidade “*loose*” para jatos reconstruídos no CMS.

- Energia hadrônica transversa, H_T , é definida como a soma escalar da magnitude do momento transverso, p_T , de todos os jatos do evento com $p_T \geq 20$ GeV.

$$H_T \equiv \sum_{\text{jatos}} p_T. \quad (4.1)$$

¹Variáveis dependentes da região de busca que serão introduzidas na seção 4.7.

- Energia hadrônica transversa perdida, H_T^{miss} , é a soma vetorial do momento transverso, \vec{p}_T , de todos os jatos do evento com $p_T \geq 20 \text{ GeV}$.

$$H_T^{\text{miss}} = |\vec{H}_T^{\text{miss}}| \equiv \left| - \sum_{\text{jatos}} \vec{p}_T \right|. \quad (4.2)$$

Durante a realização dos estudos de *triggers*, a variável H_T^{miss} terá uma definição especial, onde serão considerados na soma todos os jatos que passaram pelo critério de qualidade forte do experimento com $p_T \geq 20 \text{ GeV}$. Essa definição está mais próxima da adotada pelos *triggers* na seleção *online* e será a definição adotada em todos os gráficos do estudo de *triggers*, sendo denotada explicitamente como “*Offline* H_T^{miss} ”, que é a forma usual de informar que a variável foi reconstruída posteriormente à tomada de dados.

- Energia transversa perdida, E_T^{miss} , é a variável reconstruída definida na equação 3.9. Nos gráficos de estudo de *triggers*, ela também é usualmente denotada por “*Offline* E_T^{miss} ” para informar explicitamente que essa versão é diferente da definição adotada pelos *triggers* durante a tomada de dados.
- Valor absoluto da diferença entre as energias transversais perdidas, Δ_T^{miss} , definido como o módulo da diferença entre E_T^{miss} e H_T^{miss} ,

$$\Delta_T^{\text{miss}} \equiv |E_T^{\text{miss}} - H_T^{\text{miss}}|. \quad (4.3)$$

- Número de jatos do evento, N_{jets} . São considerados apenas jatos com $p_T \geq 30 \text{ GeV}$.
- Número de b-jatos do evento, N_{bjets} . São selecionados através do algoritmo CSVv2 [126] e possuem $p_T \geq 20 \text{ GeV}$.

4.1.1 Variáveis dos Vértices

Como já foi mencionado, os principais objetos de interesse desta análise são os jatos e os vértices reconstruídos. Nos eventos de sinal, é interessante saber quais os jatos provêm da radiação de estado inicial, quais são provenientes dos decaimentos dos gluinos e quais são originários de ruídos de *pileup*. Essa informação é obtida através de um algoritmo dedicado a associar os jatos reconstruídos aos vértices primários e secundários, o qual será referido como algoritmo de associação. O algoritmo de associação usa os constituintes carregados dos jatos, em que cada constituinte tem seu traço reconstruído e é associado ao vértice com menor distância do seu traço. O conjunto total de constituintes que compõem um jato pode ter subconjuntos associados a diferentes vértices. Para decidir a qual vértice

será associado o jato, é assinalada uma pontuação, P_{vtx} , a cada vértice candidato, que com relação a um vértice hipotético A, é definida como

$$P_{\text{vtx}}^A \equiv \sum_i \frac{p_{\text{T}}^i}{L_{\text{track}_i}^{\text{min}}}. \quad (4.4)$$

O índice i representa cada constituinte do jato que possui seu traço mais próximo do vértice A, a uma distância $L_{\text{track}_i}^{\text{min}}$ e possuindo o momento transversal p_{T}^i . Assim, o jato será associado ao vértice que possuir maior pontuação.

Na figura 4.1 temos o exemplo de um jato composto por três constituintes e os vértices mais próximos a ele foram chamados de A e B. Foi usado um esquema de cores para mostrar que os dois traços azuis estão mais próximos do vértice A e o traço vermelho está mais próximo do vértice B. A cada traço está assinalado o p_{T} do constituinte. Considerando que os traços com p_{T} iguais a 200 GeV, 70 GeV e 40 GeV têm, respectivamente, $L_{\text{track}_i}^{\text{min}}$ iguais a 5 mm, 10 mm e 8 mm, as pontuações dos vértices A e B são

$$P_{\text{vtx}}^A = \frac{200 \text{ GeV}}{5 \text{ mm}} + \frac{70 \text{ GeV}}{10 \text{ mm}} = 47 \text{ GeV/mm} \quad \text{e} \quad P_{\text{vtx}}^B = \frac{40 \text{ GeV}}{8 \text{ mm}} = 5 \text{ GeV/mm}. \quad (4.5)$$

Portanto o jato é associado ao vértice A.

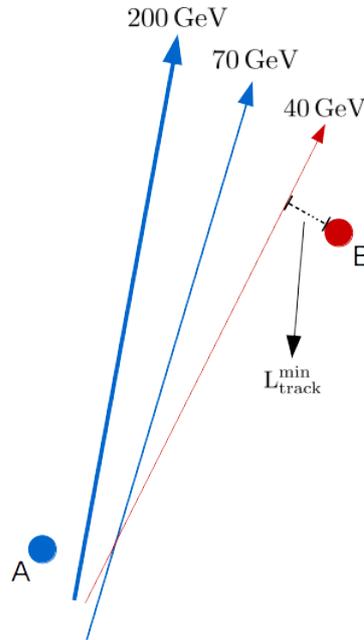


FIGURA 4.1 – Jato composto por três constituintes, cujos traços estão representados pelas setas, localizados próximos aos vértices A e B. Os valores relativos ao p_{T} dos constituintes estão especificados acima das setas.

Fazendo uso do algoritmo de associação, os jatos do evento que são associados ao vértice primário transmitem informação sobre a radiação de estado inicial. Dentre os vértices candidatos, é definido como PV, o vértice da primeira posição que apresenta

maior soma dos p_T^2 dos traços associados à ele pelo algoritmo de reconstrução de vértices. Na definição das variáveis relacionadas ao PV são considerados apenas jatos com $p_T \geq 30 \text{ GeV}$.

- Energia hadrônica transversa do PV, H_T^{PV} , é definida como a soma escalar da magnitude do momento transverso, p_T , de todos os jatos do evento associados ao PV.

$$H_T^{\text{PV}} \equiv \sum_{\text{PV jatos}} p_T. \quad (4.6)$$

- Momento transverso do jato líder do PV, ou seja, do jato de maior p_T , $\text{Jet}_L^{\text{PV}} p_T$.
- Momento transverso do jato sub-líder do PV, ou seja, do jato com o segundo maior p_T , $\text{Jet}_S^{\text{PV}} p_T$. Se não houver um jato sub-líder, o valor dessa variável será zero.
- Número de jatos associados ao PV, $N_{\text{jets}}^{\text{PV}}$.

A informação sobre o decaimento dos gluinos é obtida da construção de objetos compostos formados por um vértice secundário e jatos associados a ele, usando o algoritmo de associação. Os objetos específicos de interesse são o “SV1Jet”, objeto composto por um SV associado a um jato, e o “SV2Jets”, objeto composto por um SV associado a dois jatos. O SV2Jets é o caso de uma reconstrução completa do decaimento do gluino que gera dois quarks, já o SV1Jet representa a reconstrução incompleta do decaimento que é o caso mais comum devido à baixa energia que tipicamente têm os quarks provenientes desse decaimento. Na reconstrução desses objetos são aplicados cortes mais baixos no p_T dos jatos, visando aumentar a eficiência do sinal, preservando um bom acordo entre os dados e as simulações. O número desses objetos em cada evento são variáveis globais que serão usadas para definir as regiões de busca da análise.

- Número de SV1Jet’s, N_{SV1Jet} . São considerados apenas jatos com $p_T \geq 14 \text{ GeV}$.
- Número de SV2Jets’s, N_{SV2Jets} . São considerados apenas jatos com $p_T \geq 12 \text{ GeV}$.

4.1.2 Variáveis Angulares

O tipo de evento buscado nesta análise, como um típico evento de SUSY com conservação de paridade-R, tem uma grande quantidade de energia transversa perdida representada pelo vetor \vec{E}_T^{miss} , cuja magnitude é denotada por E_T^{miss} . Tratando-se de uma busca puramente hadrônica, \vec{E}_T^{miss} geralmente não é muito diferente de \vec{H}_T^{miss} . Um inconveniente nesse tipo de análise é a presença de uma grande quantidade de eventos de QCD na região de busca. Contudo, um artifício comumente usado para reduzir esse tipo

de *background* é o uso de variáveis angulares. O ângulo $\Delta\varphi_i$, ângulo azimutal entre o momento transverso do jato i e \vec{E}_T^{miss} (ou \vec{H}_T^{miss}), é largamente usado para essa tarefa, sendo comum algumas análises requererem que o $\Delta\varphi_i$ dos jatos mais energéticos sejam maiores que certos limiares inferiores [131,2]. Outra variável bastante usada é o ângulo $\Delta\varphi_i^*$ que representa o ângulo azimutal entre o momento transverso de um jato do evento e \vec{E}_T^{miss} (ou \vec{H}_T^{miss}) computado sem levar em consideração esse jato. Os eventos de QCD nesse caso são suprimidos requerendo-se que o menor ângulo $\Delta\varphi_i^*$ entre todos os jatos, $\Delta\varphi_{\text{min}}^*$, seja maior que um limiar definido.

Tipicamente, eventos de QCD possuem grande H_T^{miss} devido a medições incorretas do p_T dos jatos ou por causa da produção de neutrinos em decaimentos hadrônicos dentro dos jatos. A medição do p_T de um jato pode ser superestimada devido a uma medição incorreta, resultando em \vec{H}_T^{miss} em direção oposta a esse jato, (figura 4.2-C). Uma medição de p_T subestimada pode ser resultado tanto de uma medição incorreta quanto de neutrinos produzidos no decaimento hadrônico, o vetor \vec{H}_T^{miss} , nesses eventos, está quase alinhado com o jato subestimado, (figura 4.2-A e B). Na figura 4.2-A, temos o caso em que o jato 1, medido incorretamente, possui $p_{T1} > H_T^{\text{miss}}$. Já na figura 4.2-B, o jato 5, medido incorretamente, possui $p_{T5} < H_T^{\text{miss}}$. Pode-se caracterizar o jato medido incorretamente nessas duas classe de eventos, definindo a variável f_i ,

$$f_i \equiv \frac{p_{Ti}}{H_T^{\text{miss}}}. \quad (4.7)$$

Assim, com respeito ao jato subestimado i , na figura 4.2-A, $f_i > 1$, e na figura 4.2-B, $f_i < 1$. No caso de eventos de SUSY com conservação de paridade-R, o grande valor de H_T^{miss} é decorrente da produção de partículas supersimétricas invisíveis, como o neutralino nos eventos de sinal (figura 4.2-D e E).

O corte em $\Delta\varphi$ nos jatos mais energéticos é eficiente na seleção de eventos de SUSY, uma vez que, geralmente os jatos mais energéticos são os que recebem o recuo das partículas invisíveis, formando grandes ângulos $\Delta\varphi$ com respeito a \vec{H}_T^{miss} . Entretanto, na rejeição dos eventos de QCD, esse critério só é eficiente na rejeição dos eventos da figura 4.2-A. O critério de seleção utilizando o ângulo $\Delta\varphi^*$ melhora a rejeição de eventos de QCD, rejeitando eficientemente eventos como os das figuras 4.2-A e B, porém rejeita os eventos de SUSY na situação da figura 4.2-E, que são mais frequentes em casos com muitos jatos, pois aumenta a possibilidade de que um deles esteja quase alinhado com \vec{H}_T^{miss} . A seguir, são introduzidas novas variáveis angulares, propostas em [133] a partir de um estudo das propriedades geométricas dos eventos na figura 4.2, que visam melhorar o poder de discriminação entre eventos de SUSY e de QCD, focando principalmente na diminuição da rejeição dos eventos de SUSY da figura 4.2-E.

As novas variáveis angulares são $\hat{\omega}_i$ e χ_i que, como $\Delta\varphi_i^*$, suprimem eventos de QCD

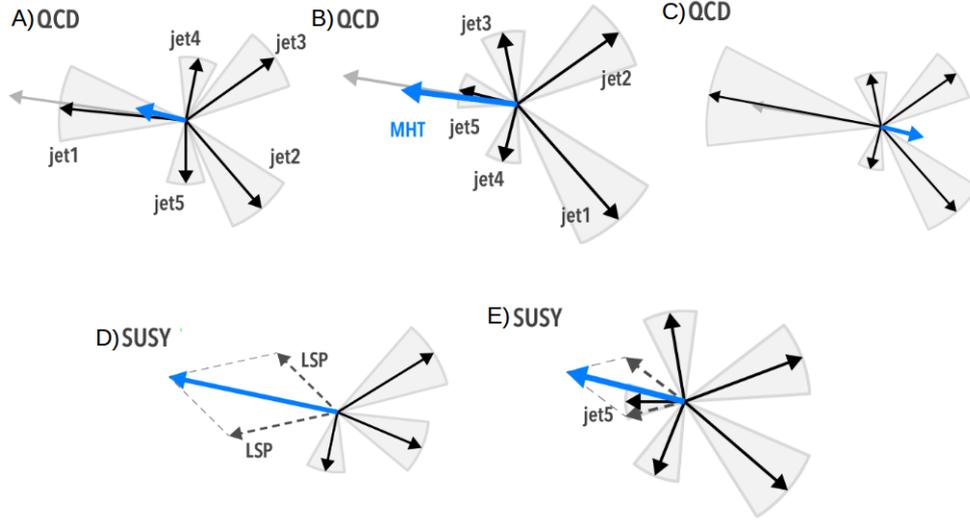


FIGURA 4.2 – Exemplos de eventos de QCD e SUSY com energia transversa perdida [132]. O vetor cinza representa o p_T correto do jato e o vetor preto o p_T reconstruído. O vetor azul representa \vec{H}_T^{miss} (ou \vec{E}_T^{miss}) e o vetor tracejado preto representa o neutralino LSP nos eventos de sinal.

ao requererem que seu menor valor entre todos os jatos, $\hat{\omega}_{\min}$, χ_{\min} e $\Delta\varphi_{\min}^*$, seja maior que um limiar inferior. Chamando o limiar inferior de γ_0 , o critério de seleção de eventos das diferentes variáveis angulares pode ser escrito como

$$\Delta\varphi_{\min}^* > \gamma_0, \quad \hat{\omega}_{\min} > \gamma_0 \quad \text{e} \quad \chi_{\min} > \gamma_0. \quad (4.8)$$

O comportamento dessas variáveis pode ser melhor compreendido analisando-as em termos de f_i e $\Delta\varphi_i$, em que suas expressões resultam em

$$\Delta\varphi_i^* = \arccos\left(\frac{f_i + \cos\Delta\varphi_i}{\sqrt{1 + f_i^2 + 2f_i \cos\Delta\varphi_i}}\right), \quad (4.9)$$

$$\hat{\omega}_i = \arctan\left(\frac{\sin(\min(\Delta\varphi_i, \pi/2))}{f_i}\right), \quad (4.10)$$

$$\chi_i = \arctan\left(\frac{\sqrt{1 + (\min(f_i, -\cos\Delta\varphi_i))^2 + 2\min(f_i, -\cos\Delta\varphi_i)\cos\Delta\varphi_i}}{\min(f_i, \max(f_i + \cos\Delta\varphi_i, 0))}\right). \quad (4.11)$$

Na figura 4.3, é mostrado o comportamento das variáveis angulares em função de $\Delta\varphi_i$ para vários valores diferentes de f_i . As variáveis $\hat{\omega}_i$ e χ_i são iguais para $\Delta\varphi_i < \pi/2$, e na região em que $\Delta\varphi_i \geq \pi/2$, $\hat{\omega}_i$ é uma função constante enquanto χ_i é igual a $\Delta\varphi_i^*$ mas com um valor máximo de $\pi/2$. Todas elas rejeitam eventos que só contenham jatos com altos valores de f_i , isso é uma característica positiva, já que tipicamente a distribuição de

f_i tem seu pico em valores mais altos em eventos de QCD do que em eventos de SUSY. Essa propriedade faz que com que f_{\max} tenha um poder de discriminação razoável entre essas duas classes de eventos.

As novas variáveis têm uma melhora na performance de discriminação entre eventos de QCD e SUSY na região em que $\Delta\varphi_i \rightarrow 0$. Como observado anteriormente, em eventos de SUSY com muitos jatos, há uma boa chance de termos alguns dos jatos quase alinhados com \vec{H}_T^{miss} , como na figura 4.2-E. Usando a variável $\Delta\varphi_i^*$, todos os eventos com pequenos valores de $\Delta\varphi_i$ são rejeitados, porém, se forem usadas as variáveis $\hat{\omega}_i$ e χ_i , o corte realizado no limiar inferior permite que uma quantidade maior de eventos sejam selecionados, mesmo aqueles com jatos possuindo valores de $\Delta\varphi_i$ muito pequenos, aumentando a aceitância de eventos de SUSY.

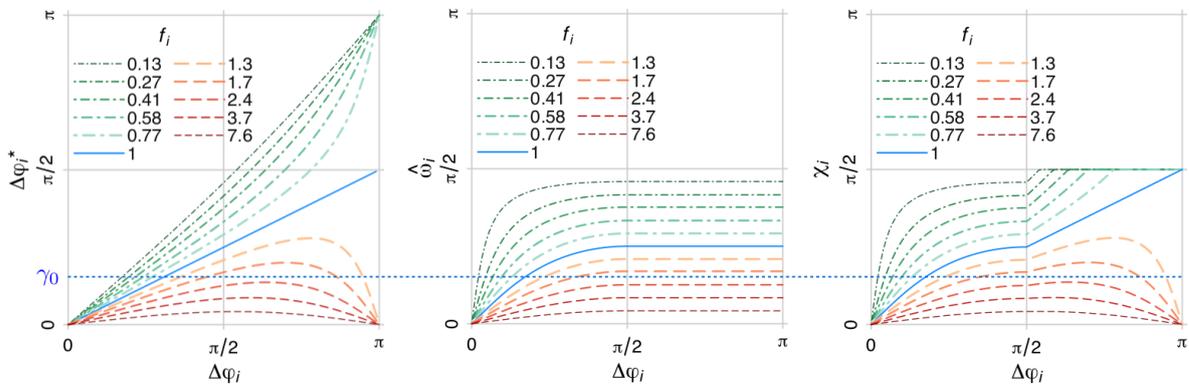


FIGURA 4.3 – Comportamento das variáveis angulares com respeito a $\Delta\varphi_i$ para diversos valores de f_i [132]. A linha pontilhada azul representa um típico corte no limiar inferior γ_0 aplicado para suprimir eventos de QCD.

A performance das variáveis angulares com respeito aos eventos de sinal buscados nesta análise é verificada na figura 4.4, onde são mostradas as distribuições dos eventos de QCD e do sinal(1000,100)² nas variáveis $\Delta\varphi_{\min}^*$, $\hat{\omega}_{\min}$ e χ_{\min} . A vantagem das novas variáveis fica evidente ao observarmos que o pico da distribuição encontra-se em altos valores da variável, isso resulta em uma aceitância muito maior dos eventos de SUSY quando for aplicada seleção para suprimir os eventos de QCD. A curva ROC³ correspondente a essas distribuições também é mostrada para os casos em que são considerados apenas os jatos com $p_T \geq 10 \text{ GeV}$ e $p_T \geq 30 \text{ GeV}$, em que fica evidente a melhor performance discriminatória de $\hat{\omega}_{\min}$ com limiar de 30 GeV para o p_T dos jatos. Portanto essa é a

²sinal(1000,100) é uma notação simplificada dos eventos de sinal em que $m_{\tilde{g}} = 1000 \text{ GeV}$ e $\Delta m = 100 \text{ GeV}$.

³A curva ROC (*receiver operating characteristic curve*) mede a performance da variável na discriminação entre duas classes de eventos. Ela é calculada varrendo todo o intervalo da variável da esquerda para a direita com um limiar inferior e, para cada valor do limiar, calcula na região à direita do limiar a eficiência na seleção de eventos de sinal e a eficiência na rejeição de eventos de QCD. Quanto maior a área sob a curva (*area under the curve* - AUC) maior o poder de discriminação da variável. Quando as duas classes de eventos são completamente separáveis, $AUC = 1$, e quando as duas classes de eventos estão completamente sobrepostas com distribuições idênticas, $AUC = 0,5$.

variável escolhida para aplicar a seleção que suprime os eventos de QCD na busca. As variáveis $\hat{\omega}_{\min}$, χ_{\min} e f_{\max} , definidas utilizando jatos com $p_T \geq 30$ GeV, fecham o grupo de variáveis globais.

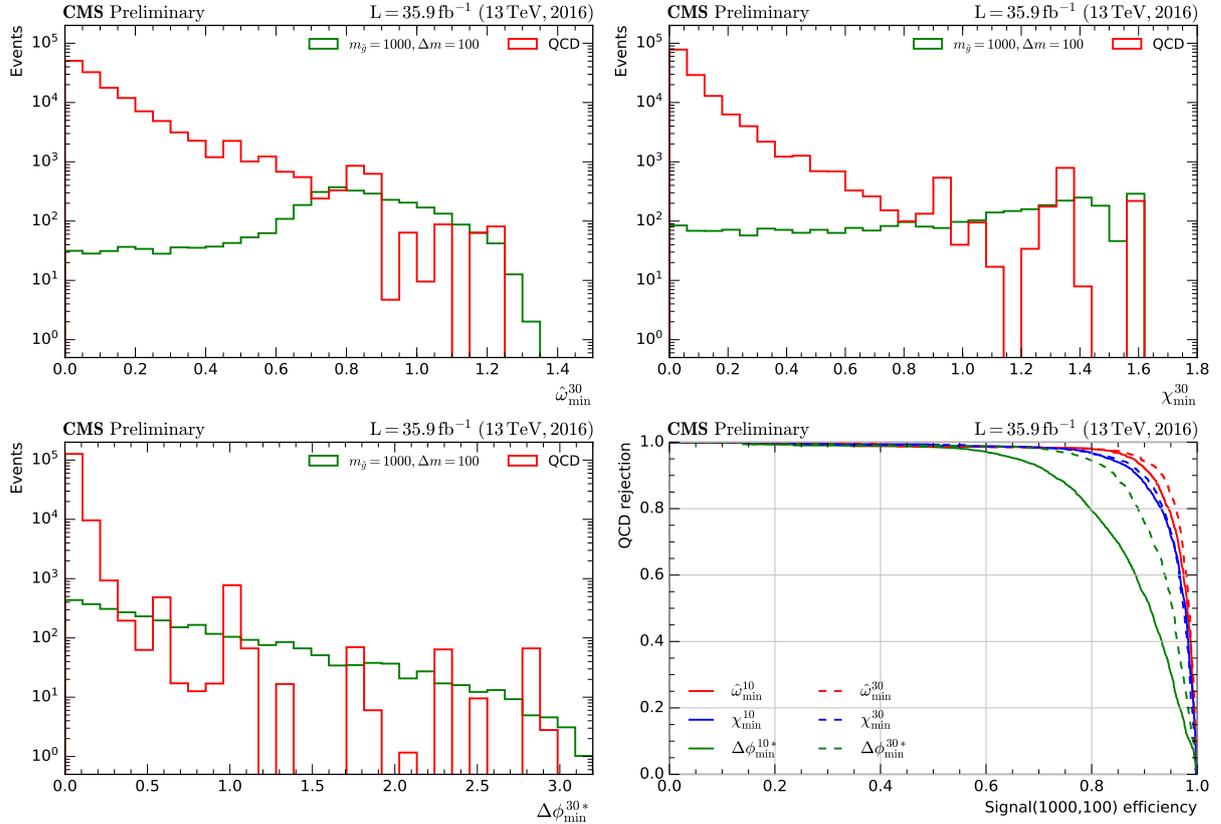


FIGURA 4.4 – Distribuições dos eventos de QCD e do sinal(1000,100) nas variáveis $\Delta\varphi_{\min}^*$, $\hat{\omega}_{\min}$ e χ_{\min} , e a curva ROC medindo a performance das variáveis na discriminação entre as duas classes de eventos.

4.2 Amostras de dados

Diferentes etapas da análise de dados são realizadas sobre eventos reais e(ou) eventos simulados (também chamados de eventos de Monte Carlo ou simplesmente, eventos de MC). Nesta seção, é feita uma breve descrição das amostras de dados reais e simulados utilizadas na análise. As tabelas com os nomes das amostras e informações adicionais estão reunidas no apêndice B.

- **Dados:** dados coletados pelo CMS em 2016 com energia de centro de massa de $\sqrt{s} = 13$ TeV e luminosidade integrada de $L = 35,9 \text{ fb}^{-1}$, tabela B.4. Os dados estão separados com respeito aos diferentes intervalos de funcionamento do LHC durante o ano, chamados de “eras”. Sendo assim, as amostras estão identificadas pelas letras B, C, D, E, F, G e H, correspondentes as diferentes eras. Dois grupos

de amostras de dados são utilizados, o primeiro é o grupo de amostras selecionado por *triggers* de elétrons que requerem a existência de um elétron no evento. As amostras de dados com a presença de um elétron são usadas no estudo de eficiência dos *triggers* de busca. O segundo grupo de amostras é selecionado por *triggers* de MET, que selecionam os eventos em que a variável E_T^{miss} (reconstruída *online*) seja maior que o limiar inferior de seleção para o qual foi configurado o *trigger*. As amostras de MET são usadas em todas as outras etapas da análise para a estimação do número de eventos nas regiões de busca e regiões de controle.

- **Simulação do background:** dados de MC para os diferentes processos físicos do MP que ocorrem nas colisões pp, simulando as condições do detector CMS durante o período de tomada de dados de 2016, tabelas B.1 e B.2. As amostras foram produzidas centralmente pelo grupo do CMS responsável por essa tarefa. O nome das amostras começa com a identificação do processo físico simulado e termina com a identificação dos geradores e configurações usadas na sua produção. As mostras de cada processo podem ser produzidas em intervalos específicos de certas variáveis cinemáticas (como p_T e H_T) com o intuito de garantir uma quantidade razoável de eventos simulados em cada um desses intervalos, permitindo uma boa descrição do processo físico, mesmo em regiões do espaço de fase com seções de choque pequenas. São utilizadas amostras de MC para os processos do MP dominantes na região de busca da análise: $Z(\nu\nu)+\text{Jatos}$, $W(l\nu)+\text{Jatos}$, $t\bar{t}+\text{Jatos}$ e Multijatos (QCD) . Também são utilizadas amostras de MC para os processos residuais: $\gamma+\text{jatos}$, $\text{DY}+\text{Jatos}$ (Drell-Yan), produção de top único e dibósons. Não são usadas amostras de MC de outros processos do MP por serem considerados desprezíveis nas regiões cinemáticas da análise. As seções de choque consideradas para cada processo físico são obtidas de diversas fontes na colaboração do CMS. Elas são fornecidas com diferentes precisões (no cálculo perturbativo) entre os diferentes processos físicos, sendo denotadas por LO (*leading order*), NLO (*next-to-leading order*) e NNLO (*next-to-next-to-leading order*).
- **Simulação do sinal:** dados de MC para o modelo simplificado de SUSY descrito na seção 2.2.2.1, produzidos privadamente para esta análise⁴, tabela B.3. As amostras são identificadas pela massa do gluino, $m_{\tilde{g}}$, e a diferença de massa entre o gluino e o neutralino, Δm , especificadas no espectro de massa do modelo. O *grid* de sinal é composto por 25 amostras de sinal, varrendo 5 valores diferentes de $m_{\tilde{g}}$, (250, 500, 750, 1000 e 1250) GeV, e 5 valores de Δm , (30, 40, 60, 80 e 100) GeV. As buscas usuais de SUSY em estados finais totalmente hadrônicos, na região em que $\Delta m < 100$ GeV, são sensíveis até valores de $m_{\tilde{g}} \sim 1000$ GeV. Por esse motivo, o

⁴O grupo de SUSY do CMS, responsável por produzir amostras de MC para modelos supersimétricos, está trabalhando na produção de amostras de sinal para serem usadas nesta análise futuramente.

valor mais alto considerado para a massa do gluino foi de 1250 GeV, o qual é usado para avaliar a necessidade de uma possível extensão do *grid*.

4.3 Estudo do Sinal

Os eventos de SUSY comprimida estudados nesta análise possuem propriedades únicas que resultam em um tipo de sinal particular no detector CMS. O par de gluinos produzidos na colisão pp formam R-hádrons e posteriormente decaem cada um em quark, anti-quark e no neutralino LSP. O gluino possui vida longa, devido à isso, os traços reconstruídos dos quarks e anti-quarks produzidos no decaimento resultam na reconstrução de vértices secundários nos pontos dos decaimentos. Além disso, os quarks e anti-quarks produzidos geram jatos descolados do ponto da colisão pp. Um desenho esquemático descrevendo os principais elementos dos eventos de SUSY buscados é mostrado na figura 4.5.

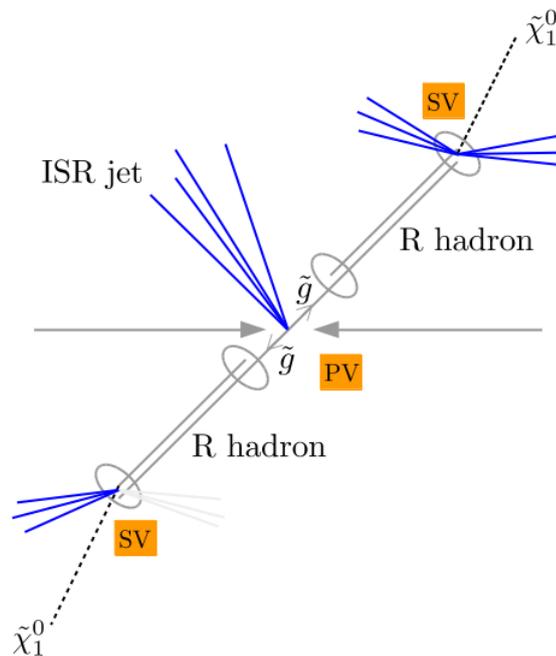


FIGURA 4.5 – Assinatura do sinal - um jato energético originário do vértice primário (ISR jato), dois vértices secundários com jatos pouco energéticos associados a eles e energia transversa perdida.

A compressão existente entre a massa do gluino e do neutralino deixa pouca energia disponível para os quarks e antiquarks produzidos no decaimento. Os jatos reconstruídos desses quarks são tipicamente pouco energéticos e uma grande parte nem mesmo é reconstruída por estarem associados a quarks⁵ com valores de p_T muito baixos. Lembrando

⁵Os quarks e antiquarks produzidos no decaimento dos gluinos serão denotados, em muitos momentos, apenas como quarks, por questão de simplicidade.

que o CMS reconstrói jatos com $p_T > 10$ GeV, quarks com valores de p_T menores que esse limiar possuem grandes chances de não gerarem jatos reconstruídos, essa característica é representada na figura 4.5, onde o vértice secundário na posição inferior está associado a apenas um jato (traços azuis), visto que o outro não foi reconstruído.

O decaimento do gluino sempre produz um quark com menor valor de p_T e outro com um valor maior. A figura 4.6 mostra as distribuições do momento transversal dos quarks com menor valor de p_T (esquerda) e dos quarks com maior p_T (direita). Observa-se que os quarks com menor momento transversal têm o pico de suas distribuições abaixo do limiar de 10 GeV, portanto, é muito mais provável que os jatos associados a esses quarks não sejam reconstruídos. A situação melhora no caso das distribuições dos jatos de maior p_T , os picos das distribuições estão localizados entre 10 e 30 GeV, logo, possuem maior probabilidade de serem reconstruídos. Fica claro também que o grau de compressão do modelo, expresso pelo parâmetro Δm , é o fator dominante na diferença de comportamento das distribuições. Modelos com valores menores de Δm produzem quarks menos energéticos, consequentemente os picos das distribuições de p_T tendem para valores menores também.

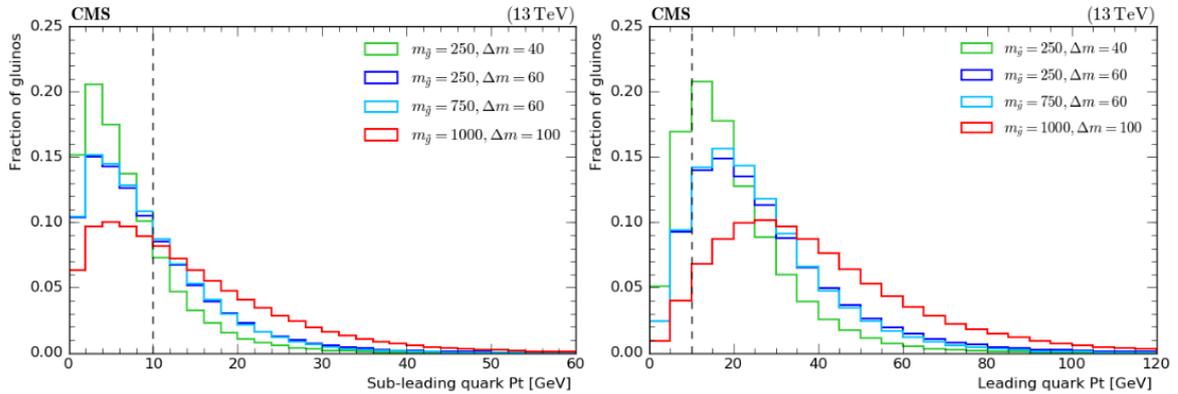


FIGURA 4.6 – Distribuições do momento transversal dos quarks produzidos no decaimento dos gluinos. À esquerda estão as distribuições correspondentes aos quarks com menor p_T e à direita estão as distribuições correspondentes aos quarks de maior p_T .

Sabendo-se que os jatos reconstruídos provenientes do decaimento dos gluinos são pouco energéticos e que não há produção de léptons nesses eventos, as possibilidades de *triggers* para a seleção desse tipo de evento ficam bem restritas. A melhor opção é realizar a seleção utilizando a variável E_T^{miss} , que explora o comportamento descrito na subseção 2.2.2.2 e representado na figura 2.11. Eventos de SUSY, especialmente os modelos comprimidos, possuem tipicamente um valor baixo de E_T^{miss} quando não são produzidos jatos da radiação de estado inicial (ISR jatos). Por outro lado, quando são produzidos um ou mais ISR jatos, a energia transversal perdida desses eventos assume valores elevados. Por tanto, os *triggers* de MET são bons na seleção de eventos com um ou mais ISR jatos, mas são péssimos na seleção de eventos que não possuem ISR jatos.

Isso pode ser visto claramente na figura 4.7, onde temos as distribuições da variável E_T^{miss} dos eventos que produziram 0, 1 ou 2 pártons como radiação de estado inicial⁶ referentes a dois pontos do *grid* de sinal com a mesma massa do gluino, que foi mantida fixa por não produzir efeito no comportamento descrito. Os gráficos mostram os eventos que sobreviveram ao corte $H_T \geq 300$ GeV. O modelo mais comprimido (à direita), com $\Delta m = 30$ GeV, é o caso mais extremo do *grid*, onde o comportamento descrito anteriormente é mais acentuado, repare que os picos das distribuições com 1 e 2 ISR jatos, estão em valores mais elevados de E_T^{miss} que os das distribuições do modelo com $\Delta m = 100$ GeV (à esquerda). Em ambos os modelos, os eventos com 0 ISR jatos estão concentrados em valores baixos de E_T^{miss} , devido ao alinhamento na direção entre os neutralinos. No modelo mais comprimido, $\Delta m = 30$ GeV, pouquíssimos eventos sobreviveram ao corte em H_T , visto que, sem ISR jatos, a atividade hadrônica fica restrita aos jatos do decaimento do gluino, porém, como esse modelo é extremamente comprimido, dificilmente haverá energia hadrônica suficiente para superar o limiar de 300 GeV em H_T . Por fim, a linha tracejada vertical na posição de 240 GeV representa o corte aplicado para se estar na região de maior eficiência do *trigger* de MET, como será visto na seção 4.4. Por tanto, só serão selecionados os eventos acima desse limiar, o que praticamente elimina a contribuição oriunda de eventos sem ISR jatos em todos os pontos do *grid* de sinal. Espera-se então que os eventos de sinal possuam pelo menos um ISR jato, como representado na figura 4.5.

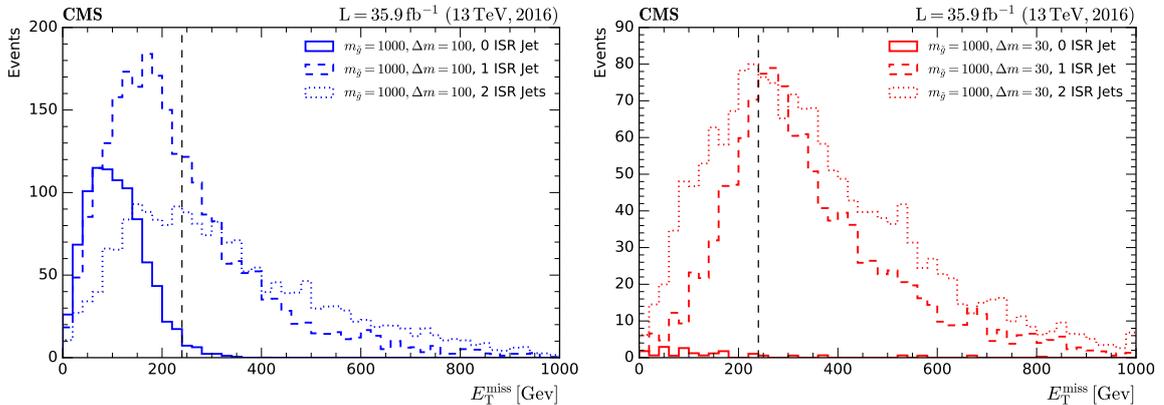


FIGURA 4.7 – Distribuições de E_T^{miss} correspondentes a eventos com 0, 1 ou 2 pártons na radiação de estado inicial.

Focando agora na propriedade de vida longa dos gluinos, sabe-se que valores menores de $m_{\tilde{g}}$ e Δm aumentam o tempo de vida dos gluinos, todavia com relação aos pontos do *grid* de sinal, as variações em Δm produzem um impacto maior, como visto na figura 2.10. A tabela 4.1 mostra a fração de gluinos que decaem em diferentes intervalos de

⁶Os pártons da radiação de estado inicial são denotados na figura 4.7 e no decorrer desse parágrafo como ISR jatos, mas tratam-se de informações a nível de gerador.

comprimento para alguns pontos do *grid* de sinal, ela informa sobre a posição aproximada dos decaimentos com relação ao *tracker*. O limite superior de 1,2 m corresponde as dimensões do *tracker* na plano transversa, e observa-se que quase a totalidade dos eventos decaem dentro desse limite. Uma fração grande dos gluínos ($\sim 70 - 100\%$) decaem após 0,2 mm, resolução aproximada da posição dos SVs reconstruídos. O intervalo entre 3 cm e 20 cm delimita aproximadamente os limites do subdetector de píxeis, que possui maior precisão e tem praticamente todos os decaimentos ocorrendo no seu interior em modelos com $\Delta m \gtrsim 60$ GeV.

Fração de gluínos (%) decaindo no intervalo:				
Sinal	> 0,2 mm	> 3 cm	> 20 cm	> 1,2 m
(250, 40)	99,8	76,0	24,6	0,9
(250, 60)	98,6	24,9	0,7	0
(750, 60)	97,5	8,8	0	0
(1000, 100)	72,1	0,0	0	0

TABELA 4.1 – Fração de gluínos decaindo em intervalos de comprimento distintos para diferentes pontos do *grid* de sinal.

A eficiência na reconstrução de SVs originários de decaimentos de gluínos é afetada por fatores como, posição do decaimento, direção do decaimento e energia das partículas visíveis produzidas. O último sendo o fator dominante que tem grande impacto no número de traços reconstruídos no decaimento. A figura 4.8 mostra as distribuições do número de SV1Jet's e SV2Jets's reconstruídos por evento para alguns pontos do *grid* de sinal. Variações significativas no comportamento das distribuições estão relacionadas a modelos com diferentes valores de Δm . Enquanto o modelo com $\Delta m = 40$ GeV possui eficiências baixas na reconstrução de eventos com SV1Jet's ($\approx 35\%$) e SV2Jets's ($\approx 6\%$), o modelo com $\Delta m = 100$ GeV apresenta melhora na eficiência com respeito aos SV2Jets's ($\approx 21\%$), embora ainda seja baixa, e tem um aumento significativo na fração de eventos com SV1Jet's ($\approx 73\%$). Os SV2Jets's são difíceis de serem reconstruídos porque, o quark de menor p_T , produzido no decaimento do gluíno, possui tipicamente uma energia muito baixa, tendo alta probabilidade de não ser reconstruído como visto anteriormente. A reconstrução perfeita de dois SV2Jets's acontece em um número extremamente pequeno de eventos.

Na distribuição de SV1Jet's na figura 4.8 existem eventos em que foram reconstruídos mais de dois vértices secundários. Isso ocorre porque os algoritmos de identificação de SVs, apresentado na subseção 3.3.6, e o algoritmo de associação, descrito na subseção 4.1.1, não são perfeitos e podem resultar em más reconstruções de SV1Jet's e SV2Jets's. A principal fonte de ruídos consiste de jatos de *pileup* e para poder avaliar o impacto do ruído de *pileup* na reconstrução de SV1Jet's e SV2Jets's classifica-se os jatos dos eventos de sinal em duas categorias: jatos gerados de quarks produzidos no decaimento do gluíno

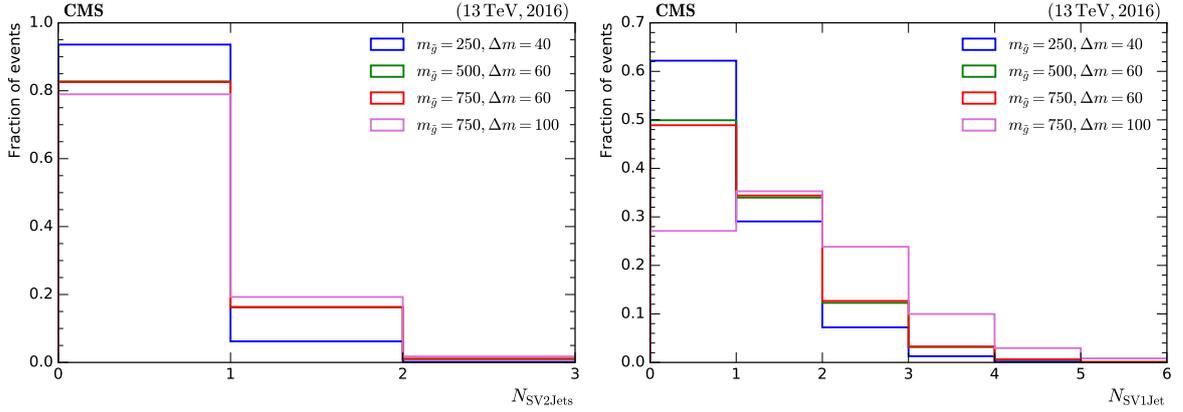


FIGURA 4.8 – Distribuições normalizadas da multiplicidade de SV1Jet’s e SV2Jets’s reconstruídos por evento.

(GD jatos) e jatos oriundos de ruídos de *pileup* (PU jatos). A classificação é feita através de um algoritmo que combina as informações dos jatos reconstruídos com as informações a nível de gerador dos quarks produzidos nos decaimentos. O algoritmo de combinação (*matching*) executado sobre cada evento de sinal é descrito pelas seguintes etapas:

1. Identifica-se os quatro quarks correspondentes aos decaimentos dos dois glúinos;
2. Realiza-se um *loop* sobre todos os jatos e sobre todos os quarks identificados na etapa anterior para encontrar o par (jato, quark) com menor valor de $\Delta R_{j,q} \equiv \sqrt{(\eta_j - \eta_q)^2 + (\phi_j - \phi_q)^2}$, sendo ϕ o ângulo azimutal e η a pseudorapidez do jato j e do quark q que compõem o par. Ao final do *loop*, o jato e o quark do par com menor $\Delta R_{j,p}$ são selecionados e removidos do grupo inicial, repetindo-se o *loop* sobre os jatos e quarks restantes. Esse procedimento é repetido até que todos os quatro quarks tenham formado pares com jatos;
3. Adota-se um valor de corte para a definição dos jatos de sinal, ΔR_{GD} , e um valor de corte para a definição dos jatos de *pileup*, ΔR_{PU} . Os jatos correspondentes aos pares que tiverem $\Delta R_{j,q} < \Delta R_{GD}$ são classificados como jatos de sinal, e todos os jatos que não formaram pares com os quarks, como também todos os jatos dos pares em que $\Delta R_{j,q} > \Delta R_{PU}$, são classificados como jatos de *pileup*.

A definição dos jatos de *pileup* permite a contaminação por jatos da radiação de estado inicial, mas essa contribuição é muito pequena em comparação aos ruídos de *pileup*, não alterando significativamente os resultados.

A avaliação é realizada sobre duas amostras distintas, uma formada por todos os jatos dos eventos pertencentes aos pontos do *grid* com $\Delta m = 30$ GeV e a outra correspondente aos pontos com $\Delta m = 100$ GeV, nas discussões que seguem, elas são referidas, respectivamente, como modelos mais comprimidos e modelos menos comprimidos. Essa combinação

foi feita para aumentar a estatística das amostras utilizadas e baseou-se na consideração de que a massa do gluino não altera significativamente a eficiência na reconstrução de SV1Jet's e SV2Jets's, como pode ser verificado na figura 4.8. Ademais, duas definições diferentes de jatos de sinal e *pileup* são utilizadas, a definição forte com $\Delta R_{\text{GD}} = 0,2$ e $\Delta R_{\text{PU}} = 0,6$ e a definição fraca com $\Delta R_{\text{GD}} = 0,4$ e $\Delta R_{\text{PU}} = 0,4$. A figura 4.9 mostra a distribuição de ΔR de todos os pares quark-jato formados na etapa 2 do algoritmo de combinação. Os jatos utilizados na análise são reconstruídos pelo algoritmo Anti- K_t com $R = 0,4$, por essa razão os pontos de corte da combinação quark-jato são realizados em valores próximos a 0,4. Pode ser visto também na figura 4.9 que o comportamento da distribuição de ΔR muda na região em torno de 0,4. Os pontos de corte da definição forte são representados pelas linhas tracejadas verticais da cor preta, e o ponto de corte da definição fraca é indicado pela linha tracejada verde.

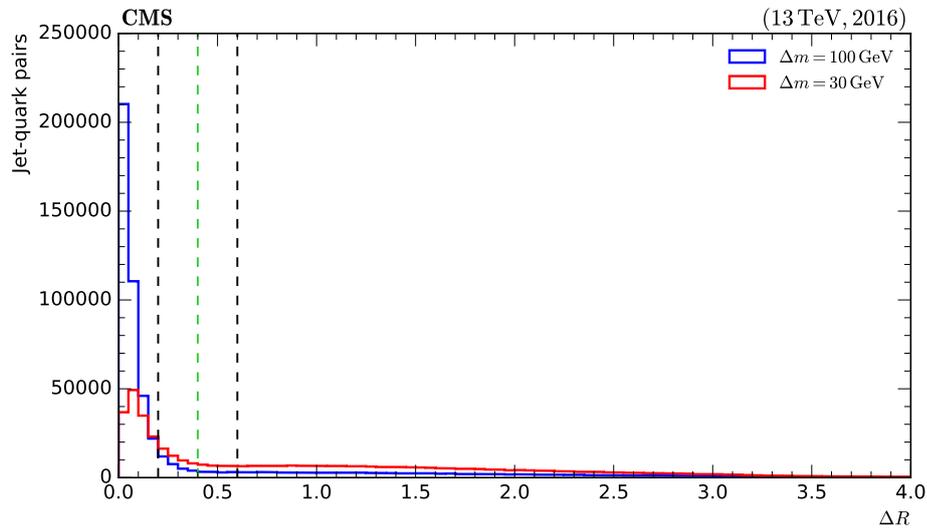


FIGURA 4.9 – Distribuição do ΔR de todos os pares quark-jato formados na etapa dois do algoritmo de combinação.

Os resultados da avaliação são mostrados nos gráficos das figuras 4.10 (realizada com a definição forte) e 4.11 (realizada com a definição fraca). Em cada figura, os três gráficos de cima mostram a eficiência na associação de jatos de sinal a SV1Jet's, SV2Jets's e SVs⁷. Já os gráficos de baixo mostram a eficiência dos jatos de *pileup* em serem associados a SV1Jet's, SV2Jets's e SVs. As eficiências correspondentes às duas amostras são expostas, $\Delta m = 30$ GeV em vermelho e $\Delta m = 100$ GeV em azul. Nos histogramas, os tons mais claros representam a quantidade total de jatos que satisfazem a condição fraca ou forte, e os tons escuros denotam os jatos associados aos SV1Jet's, SV2Jets's ou SVs.

As distribuições totais dos GD jatos com relação ao p_T do jato são bem diferentes entre as duas amostras. O número de jatos reconstruídos cai mais rápido nos modelos

⁷SV nesse contexto denota os elementos do conjunto formado pelos SV1Jet's e SV2Jets's.

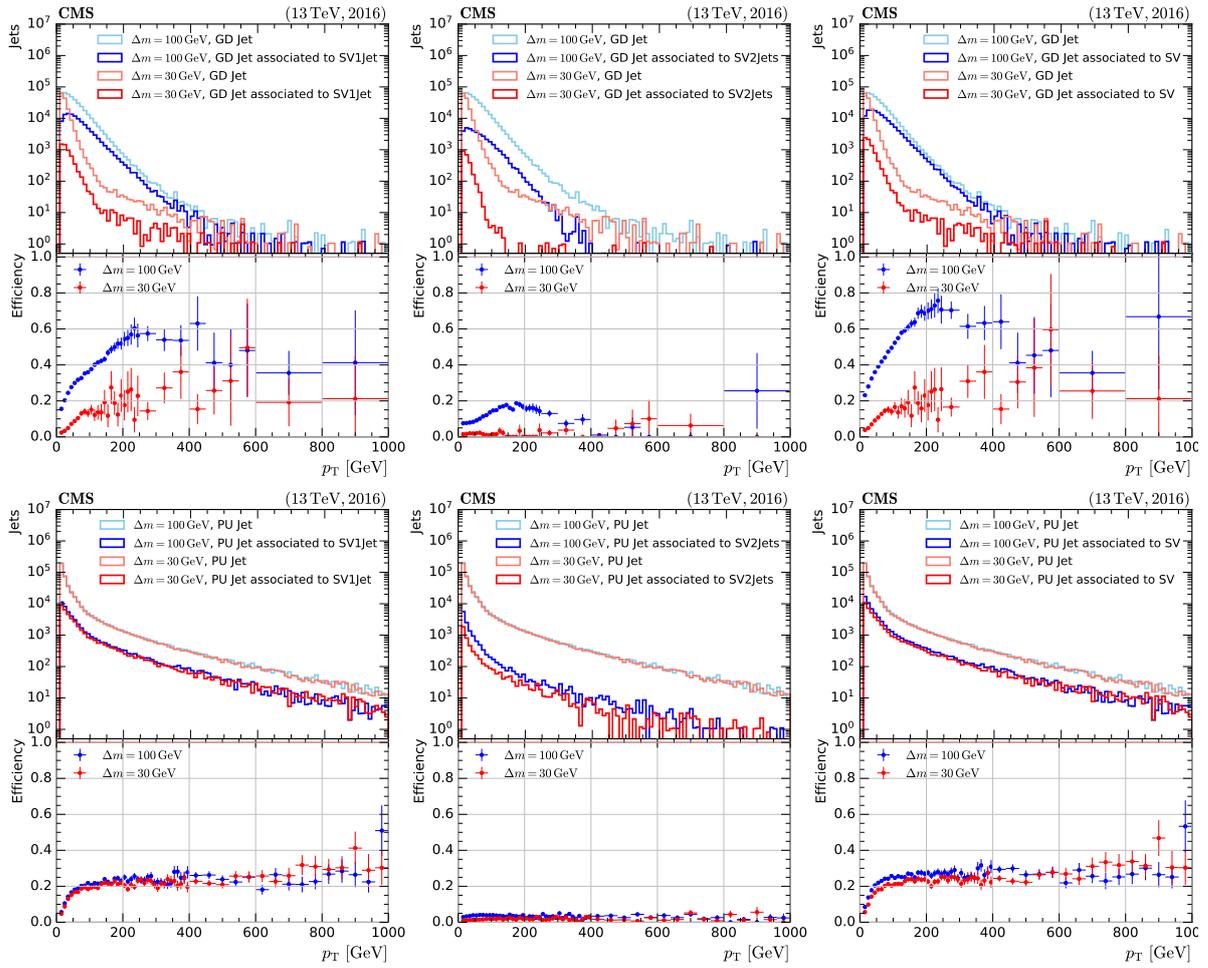


FIGURA 4.10 – Gráficos de eficiência na reconstrução de SV1Jet's, SV2Jets's e SVs, considerando-se a definição forte para GD jatos e PU jatos.

mais comprimidos devido a restrição mais forte na energia dos quarks do decaimento. As eficiências também apresentam comportamentos diferentes, como era esperado, os valores das eficiência para os modelos menos comprimidos são maiores no geral. Nesse modelo, a proporção de quarks com energia suficiente para originarem jatos reconstruídos é maior, consequentemente é maior a probabilidade de serem reconstruídos SVs também. A eficiência na associação a SV1Jet's aumenta acentuadamente com relação ao p_T do jato no intervalo $10 \text{ GeV} \leq p_T < 250 \text{ GeV}$ para os modelos menos comprimidos, partindo de uma eficiência de 15% e alcançando um máximo em torno de 60% na definição forte e 55% na fraca enquanto nos modelos mais comprimidos também existe um crescimento acentuado na região de baixo p_T , mas que não fica tão nítido após 100 GeV, onde sua eficiência começa a oscilar. Na eficiência com relação aos SV2Jets's, a curva dos modelos menos comprimidos apresenta um pico em torno de $p_T = 180 \text{ GeV}$, ficando entre 5% e 20%. Nos modelos mais comprimidos, a eficiência é bem baixa, menor que 5%. A região com valores mais altos do p_T dos jatos fica cada vez mais longe da janela de energia acessível ao p_T dos jatos de sinal, logo, aumenta a probabilidade de serem originários de ruídos de

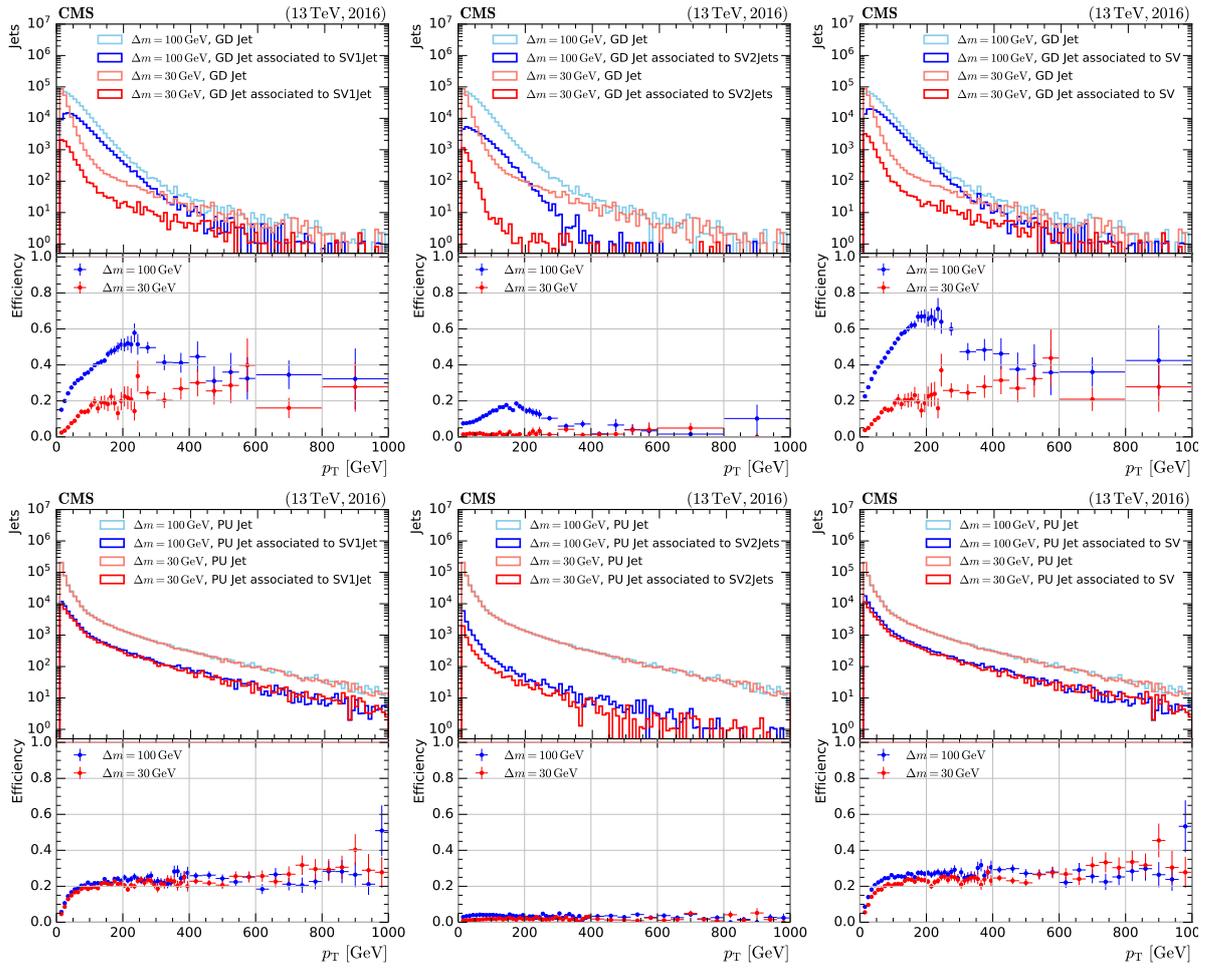


FIGURA 4.11 – Gráficos de eficiência na reconstrução de SV1Jet's, SV2Jets's e SVs, considerando-se a definição fraca para GD jatos e PU jatos.

pileup, e adicionando a baixa estatística nessa região, resulta em um comportamento da eficiência não muito bem definido. Os jatos de sinal dos modelos mais comprimidos, além de possuírem menos jatos reconstruídos, apresentam eficiências similares aos dos jatos de *pileup*, por isso sofrem maior contaminação de *pileup*.

Em consideração aos jatos de *pileup*, as distribuições do total de jatos em ambas as definições apresentam a propriedade esperada de serem independentes da amostra de sinal usada. Esse comportamento é devido ao fato de que a contaminação de *pileup* não depende da física do evento contaminado. O tipo de contaminação de *pileup* dominante é composta por jatos de *pileup* deslocados que deram origem a SV1Jet's. Obviamente a energia dos jatos afeta a eficiência na reconstrução dos SVs. Na região com $10 \text{ GeV} \leq p_T < 100 \text{ GeV}$, o aumento da eficiência de associação a SV1Jet's é mais acentuado, indo de 5% a 22% aproximadamente. Isso ocorre porque jatos com p_T muito baixo possuem tipicamente menos traços e dificultam a reconstrução de SVs. A partir de 100 GeV a eficiência apresenta um crescimento bem mais suave. O comportamento das curvas correspondentes aos modelos

mais e menos comprimidos são similares para esse tipo de contaminação. A eficiência da contaminação de *pileup* em SV2Jets's é suprimida em valores menores que 7%. Ela ocorre em duas situações possíveis, a primeira é independente da física do evento e consiste de um SV reconstruído com traços de dois jatos de *pileup* que foram associados posteriormente a esse vértice pelo algoritmo de associação. A segunda possibilidade é a reconstrução de um SV através de traços de um jato de sinal, com ou sem contaminação por traços de *pileup*, em que são associados ao SV um jato de sinal e um jato de *pileup*. Essa segunda possibilidade depende da física do evento e é a origem da diferença nas distribuições de PU jatos associados a SV2Jets's entre os modelos mais e menos comprimidos. Os modelos menos comprimidos reconstróem mais SVs, dessa forma, são mais susceptíveis a essa segunda possibilidade de contaminação. A variação das eficiências com relação às definições fraca e forte é insignificante nas curvas de eficiência do *pileup*. A contaminação de *pileup* é mais controlada nos SV2Jet's, de modo que a informação contida nesses objetos é mais limpa que a contida nos SV1Jet's.

As eficiências baixas dos jatos de sinal estão associadas à baixa eficiência na reconstrução de SVs em modelos comprimidos. Para se avaliar a eficiência do algoritmo de associação, é necessário examinar jatos correspondentes a quarks originários de decaimentos de gluinos em que foi possível reconstruir o SV. Tendo esse objetivo em mente, para cada evento, é realizado um *loop* sobre a posição do vértice de decaimento dos dois gluinos (GDV) e sobre a posição de todos os SVs. Ao final do *loop*, o par GDV-SV com menor distância entre eles é selecionado e removido do conjunto inicial. Então, realiza-se um *loop* novamente sobre os SVs restantes para selecionar aquele que possui menor distância com relação ao vértice do outro gluino. Caso o evento tenha menos de dois SVs reconstruídos, os gluinos que não formam pares, não são considerados no restante da análise. Após a identificação dos pares GDV-SV, a eficiência do algoritmo de associação é avaliada sobre GD jatos em duas situações distintas:

1. o GD jato candidato é pré-selecionado pela definição forte, em que $\Delta R_{GD} = 0,2$, e são selecionados apenas os GD jatos combinados com quarks correspondentes a GDV's que formam pares com algum SV. Os três gráficos superiores da figura 4.12 mostram as curvas de eficiência para essa situação.
2. O GD jato é pré-selecionado pela definição forte, em que $\Delta R_{GD} = 0,2$, são considerados apenas os GD jatos combinados com quarks correspondentes a GDV's que formam pares com algum SV, e são selecionados apenas os GD jatos correspondentes a pares GDV-SV cuja distância entre eles é menor que 0,25 mm para o modelos menos comprimidos e menor que 12 mm para os modelos mais comprimidos, esses pontos de corte correspondem a seleção de aproximadamente 50% dos pares GDV-SV. Os três gráficos inferiores da figura 4.12 mostram as curvas de eficiência para essa segunda situação.

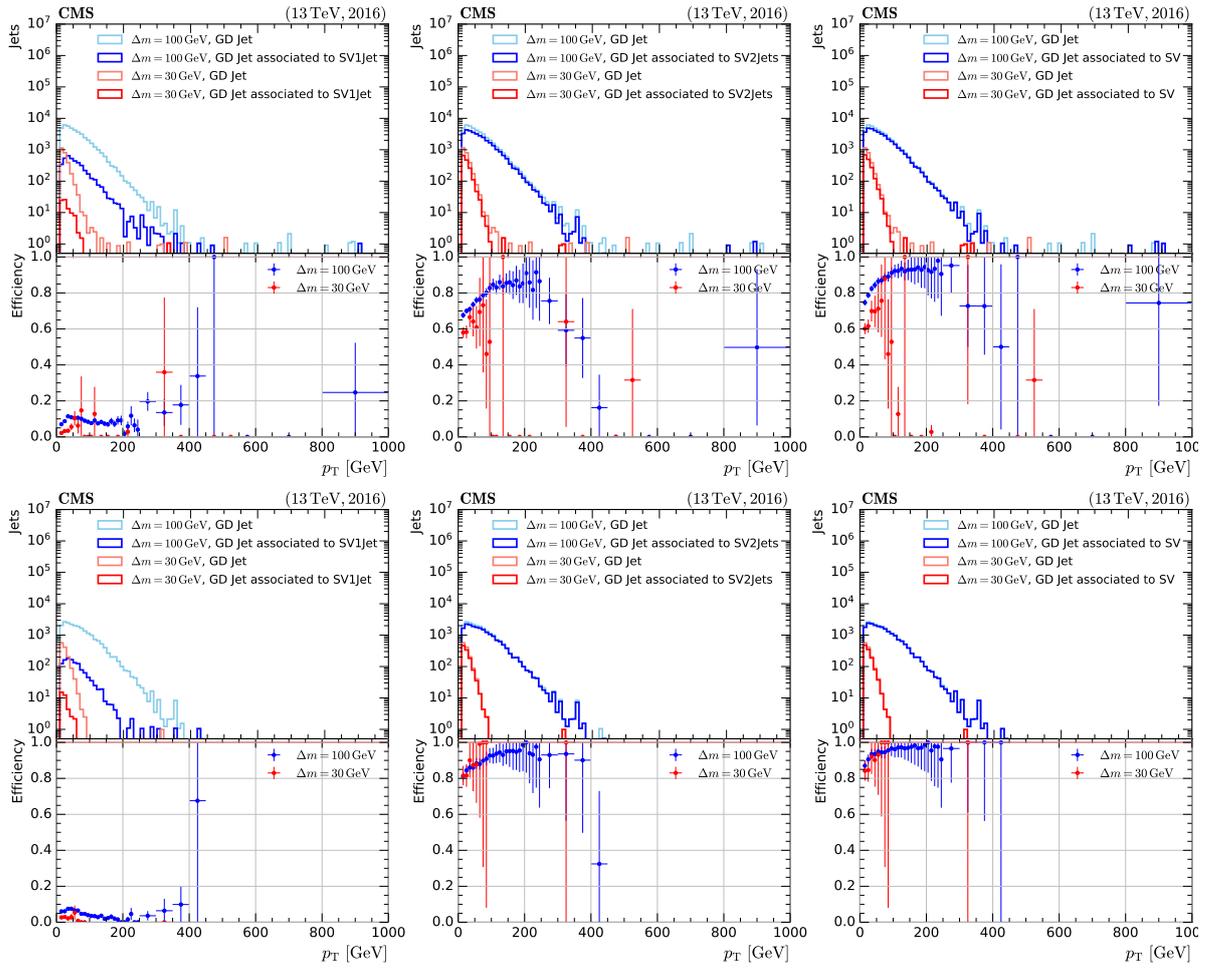


FIGURA 4.12 – Gráficos de eficiência do algoritmo de associação a SV1Jet's, SV2Jets's e SVs. Nos gráficos de cima, os GD jatos são definidos pela situação 1, e nos gráficos de baixo, eles são definidos pela situação 2.

As duas situações propostas, visam garantir que o SV correspondente ao GD jato tenha realmente sido reconstruído. Desse modo, a eficiência da associação dos GD jatos aos SVs nessas situações, depende exclusivamente da performance do algoritmo de associação. A primeira situação só requer que um SV tenha sido reconstruído enquanto a segunda aplica um corte de qualidade, selecionando apenas a metade dos pares GDV-SV que foram reconstruídos mais próximos um do outro. A eficiência total do algoritmo de associação nos modelos comprimidos é alta na região de baixo p_T dos jatos, em torno da janela de energia dos modelos comprimidos. Os modelos menos comprimidos apresentam uma região de eficiência máxima, em torno de $p_T = 180 \text{ GeV}$, de aproximadamente 94% na situação 1 e 98% na situação 2. Já os modelos mais comprimidos não têm uma região de saturação bem definida, mas para valores baixos de p_T , possui eficiência total entre 45-90% na situação 1 e 84-100% na situação 2. A maior parte dos jatos são associados a SV2Jets's e a parte associada a SV1Jet's corresponde a uma porcentagem menor que 15%.

4.4 Eficiência do Trigger

A análise de eficiência do *trigger* deve ser realizada na região mais similar possível da região de busca dos eventos de sinal para reduzir ao máximo um possível viés na medição da eficiência. A região de busca, definida exatamente na seção 4.5, tem como base uma série de filtros e seleções, chamados de pré-seleção, que garantem uma boa qualidade dos dados e excluem as contribuições redutíveis dos eventos de *background*. A série de filtros e seleções que compõem a pré-seleção é descrita abaixo.

1. Os eventos devem passar por todos os filtros de MET, descritos da subseção 3.3.5, que eliminam eventos anômalos afetados por efeitos diversos que resultam em uma má reconstrução da energia transversa perdida.
2. $H_T \geq 300 \text{ GeV}$: o corte em 300 GeV é realizado, devido a um desacordo entre os dados e as simulações abaixo desse limiar. Esse desacordo é devido ao fato de que a maioria das amostras de MC são produzidas com um limiar inferior em H_T a nível de gerador, Gen $H_T \geq 100 \text{ GeV}$.
3. Veto de múons e elétrons: os eventos de sinal buscados não resultam na produção de múons e elétrons isolados, sendo assim, veta-se a existência de múons e elétrons isolados reconstruídos com $p_T \geq 10 \text{ GeV}$.
4. $\text{Jet}_L^{\text{PV}} p_T \geq 50 \text{ GeV}$: como foi visto na figura 4.7, os eventos de sinal que sobrevivem ao corte realizado na variável E_T^{miss} possuem tipicamente um ou mais jatos gerados na radiação de estado inicial. O corte em 50 GeV garante uma boa definição do vértice primário.
5. $\hat{\omega}_{\text{min}} \geq 0.4$: corte aplicado para supressão de eventos de QCD. O corte em 0,4 elimina mais de 90% dos eventos de QCD.
6. $\Delta_T^{\text{miss}} < 120 \text{ GeV}$: corte realizado para que o estudo da eficiência do *trigger* seja realizado em uma região onde ele apresenta uma performance bem comportada.

A primeira etapa na análise dos dados coletados pelo detector CMS é a definição do *trigger* ou conjunto de *triggers* mais apropriados para a busca em questão, ou seja, qual deles seleciona a maior quantidade de eventos correspondentes ao sinal procurado. A seleção deve considerar as características dos eventos de sinal. Os *triggers* geralmente usados em buscas de SUSY, em canais totalmente hadrônicos com conservação da paridade-R, realizam seleções nas variáveis H_T^{miss} e E_T^{miss} [131,2]. Os neutralinos produzidos fazem com que os eventos de SUSY possuam uma maior tendência a valores elevados de H_T^{miss} e E_T^{miss} que eventos do MP. Dentre os HLT *triggers* disponíveis, foram escolhidos três *triggers* que funcionaram durante praticamente toda a tomada de dados de 2016.

- HLT_PFMET170_HBHECleaned;
- HLT_PFMET110_PFMHT110_IDTight;
- HLT_PFMET120_PFMHT120_IDTight;

Portanto são utilizados na análise os eventos selecionados por pelo menos um dos *triggers* de busca acima⁸.

O primeiro é um *trigger* simples de MET, em que PFMET170 informa a quantidade calculada *online* pelo *trigger* e o limiar inferior da seleção. Logo, o *trigger* calcula o MET com objetos físicos reconstruídos pelo algoritmo PF e seleciona os eventos com MET maior que 170 GeV. HBHECleaned informa que foi utilizado um filtro para reduzir os ruídos provenientes dos subdetectores HB e HE do calorímetro hadrônico. Para que sejam analisados por esse *trigger*, os eventos devem ser previamente selecionados por pelo menos um dos seguintes L1 *triggers*⁹: L1_ETM50, L1_ETM60, L1_ETM70, L1_ETM80, L1_ETM90, L1_ETM100. A sigla ETM, como MET, refere-se à energia transversa perdida.

Os outros dois são *triggers* combinados de MET e MHT, calculadas com objetos reconstruídos pelo algoritmo PF com limiares inferiores de 110 e 120 GeV. Desse modo, são selecionados os eventos com MET ou MHT maiores que o limiar inferior. IDTight refere-se ao critério de qualidade dos jatos utilizados no cálculo *online* da variável MHT. Os L1 *triggers* L1_ETM50, L1_ETM60, L1_ETM70, L1_ETM75, L1_ETM80, L1_ETM85, L1_ETM90, L1_ETM95, L1_ETM100, L1_ETM120, L1_ETM75_Jet60_dPhi_Min0p4 e L1_DoubleJetC60_ETM60 formam o conjunto de sementes do *triggers* combinados. O HLT *trigger* com limiar em 110 GeV foi pré-escalado no final da tomada de dados em 2016, de modo que a quantidade de dados selecionados por ele para análise física corresponde a uma luminosidade integrada de $L = 35,3 \text{ fb}^{-1}$, um pouco menor que a luminosidade integrada total de 2016 ($L = 35,9 \text{ fb}^{-1}$).

A eficiência de um *trigger* pode ser definida como a porcentagem de eventos selecionados pelo *trigger* dentre os eventos que ele é esperado selecionar. Ela depende do *hardware* e *software* desenvolvido para o experimento e também das propriedades de determinado tipo de evento. A medição da eficiência dos *triggers* de busca é realizada utilizando o método ortogonal [131,2], no qual a medição é feita sobre uma amostra de dados reais selecionados previamente por um *trigger* (denotado como *trigger* ortogonal) cujo critério de seleção, idealmente, não tenha correlação com o critério utilizado pelos *triggers* de busca. No estudo de *triggers* de MET e MHT, são geralmente utilizados *triggers* de elétrons como *trigger* ortogonal, que nesta análise foi escolhido para ser o HLT

⁸Foi estudada a possibilidade de explorar a propriedade de vida longa dos gluinos e utilizar *triggers* de seleção de jatos deslocados, entretanto, os candidatos existentes possuem o requerimento de dois calojatos (jatos reconstruídos *online* utilizando apenas informações do calorímetros) com $p_T > 40 \text{ GeV}$ e essa restrição elimina muitos eventos de sinal.

⁹Estes L1 *triggers* são chamados de “sementes” (*seeds*) do HLT *trigger* em questão

trigger HLT_Ele27_WPTight_Gsf. Ele seleciona eventos com pelo menos um elétron com $p_T > 27$ GeV, identificado pelo critério de seleção forte (*Working Point Tight*) que possui uma eficiência média de 70%, e foi filtrado pelo algoritmo denominado Gsf (*Gaussian-sum Filter*) [134]. Adicionalmente é requerido um elétron reconstruído com $p_T > 30$ GeV para assegurar uma eficiência constante com respeito ao momento transverso do elétron.

Os *triggers* de busca utilizados na análise selecionam os eventos com base nos valores de E_T^{miss} e H_T^{miss} calculados *online*. Porém, o estudo dos *triggers* é realizado com os valores dessas variáveis calculados *offline*. A variável H_T^{miss} possui uma versão construída para se assemelhar ao critério adotado no cálculo *online*, como descrito na seção 4.1, a qual será utilizada no estudo de eficiência dos *triggers* de busca. A figura 4.13 mostra o comportamento da eficiência dos *triggers* de busca com respeito a E_T^{miss} e H_T^{miss} . A eficiência em cada *bin* é dada pela razão entre o **numerador** e o **denominador** definidos como:

- **denominador** - soma de todos os eventos que passam pela pré-seleção sem a restrição sobre a presença de elétrons, são selecionados pelo *trigger* de elétrons e possuem um elétron reconstruído com $p_T > 30$ GeV;
- **numerador** - soma de todos os eventos que satisfazem os requerimentos do denominador e que também são selecionados pelos *triggers* de busca.

No caso de *triggers* “perfeitos”, que rejeitassem todos os eventos abaixo do limiar e selecionassem todos que estão acima, existiriam no gráfico apenas regiões azuis (com eficiência de 0%) e vermelhas (com eficiência de 100%). Entretanto, na realidade, mesmo acima dos limiares de corte a eficiência dos *triggers* não é 100%, e é justamente esse valor que será calculado no método ortogonal. A região de transição com cores intermediárias caracteriza a diferença entre a definição das variáveis calculadas *online* e *offline*.

Espera-se que os eventos selecionados pelo *trigger* para a análise física estejam na região de máxima eficiência do *trigger* (região do *plateau*) e que a eficiência seja constante nessa região. No gráfico à esquerda na figura 4.13, observa-se que a região de transição é mais comportada dentro da região delimitada pelas duas linhas azuis diagonais, que corresponde aos eventos em que $\Delta_T^{\text{miss}} < 120$ GeV. O requerimento em Δ_T^{miss} faz parte dos cortes da pré-seleção de eventos, e o gráfico à direita na figura 4.13 mostra a exata região de eventos considerada na medição da eficiência. O corte em $\Delta_T^{\text{miss}} < 120$ GeV resulta na exclusão predominantemente de eventos de $W(l\nu)$ +jatos e praticamente não afeta os eventos de sinal que estão tipicamente dentro dessa região.

A região do *plateau* é delimitada no gráfico à direita na figura 4.13 pelas linhas azuis horizontal e vertical que representam, respectivamente, os limiares inferiores das variáveis *offline* $H_T^{\text{miss}} \geq 220$ GeV e $E_T^{\text{miss}} \geq 240$ GeV. Ela é definida ao se analisar o comportamento

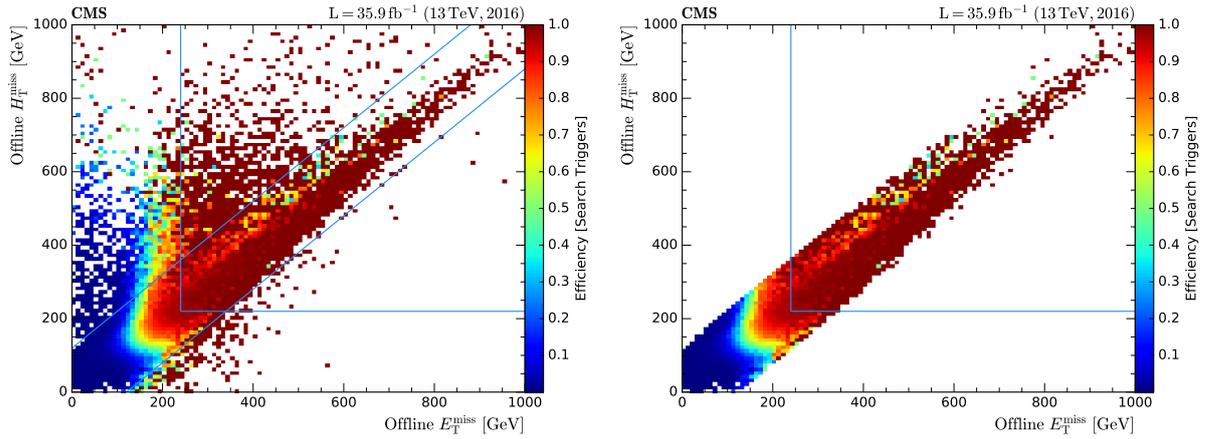


FIGURA 4.13 – Gráficos bidimensionais da eficiência dos *triggers* de busca com respeito as variáveis *offline* H_T^{miss} e E_T^{miss} , sem a aplicação do corte em Δ_T^{miss} (esquerda) e com a aplicação do corte em Δ_T^{miss} (direita).

da curva de eficiência com relação a uma das variáveis após aplicar o corte de limiar inferior na outra, acima do qual essa segunda variável possui eficiência constante. Os limiares inferiores mostrados na figura 4.13 são os valores otimizados (que selecionam mais dados no *plateau*) após a avaliação de diversos cortes possíveis. A figura 4.14 mostra as curvas de eficiência de cada uma das duas variáveis *offline*, após ter sido realizado o corte da região de *plateau* na outra.

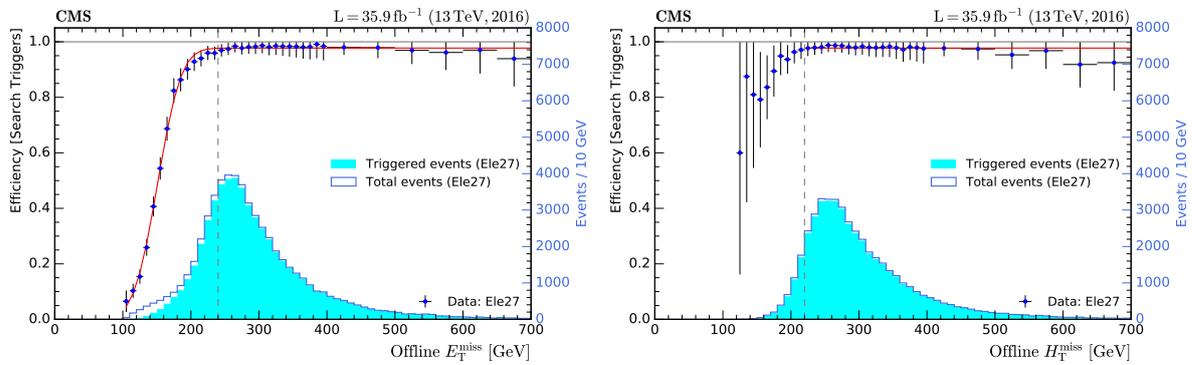


FIGURA 4.14 – Curva de eficiência com respeito à variável E_T^{miss} após ter sido aplicado o corte $H_T^{\text{miss}} \geq 220$ GeV (esquerda). Curva de eficiência com respeito à variável H_T^{miss} após ter sido aplicado o corte $E_T^{\text{miss}} \geq 240$ GeV (direita). As linhas tracejadas verticais denotam os limiares inferiores da região em que se alcança a eficiência máxima.

A eficiência da região de *plateau* é estimada utilizando a curva de eficiência com relação a E_T^{miss} , visto que é a curva de eficiência do *trigger* melhor definida. Uma curva de *trigger* perfeita, teria sua eficiência ajustada por uma função degrau, entretanto, a curva de um *trigger* real possui uma região de transição, entre as regiões de eficiência mínima e máxima. Logo, é usada uma função aproximada da função degrau para realizar o ajuste. A função

utilizada é definida por

$$f(x; A, \mu, \sigma) = \frac{A}{2} \left(1 + \operatorname{erf} \left(\frac{x - \mu}{\sigma} \right) \right), \quad \text{tal que} \quad \operatorname{erf}(z) \equiv \frac{2}{\sqrt{\pi}} \int_0^z e^{-t^2} dt. \quad (4.12)$$

Os parâmetros A , μ e σ caracterizam, respectivamente, a amplitude da função, o centro da região de transição e a inclinação da função na região de transição, e $\operatorname{erf}(z)$ é conhecida como função erro. Os valores obtidos no ajuste são expostos na tabela 4.2 e a curva de ajuste resultante é mostrada em vermelho no gráfico à esquerda na figura 4.14 e o valor de sua amplitude, que denota a eficiência do *trigger*, e representado por um reta vermelha horizontal no gráfico à direita. O valor da amplitude, A , é naturalmente a estimativa da eficiência do *trigger* na região do *plateau*. Um segundo ajuste, utilizando a função logística, foi realizado como forma de verificação e resultou numa curva de eficiência praticamente idêntica à da função anterior, o valor obtido para a amplitude foi de $A = 0,9784 \pm 0,0019$, que é compatível com o resultado anterior em 1σ .

Parâmetro	Valor
A	$0,9772 \pm 0,0023$
μ	$150,9 \pm 0,4$
σ	$38,8 \pm 0,8$

TABELA 4.2 – Valores obtidos para os parâmetros da função f , após ter sido realizado o ajuste com a curva de eficiência dos *triggers* de busca do gráfico à esquerda na figura 4.14.

Na região do *plateau*, onde espera-se um comportamento praticamente constante da eficiência, foi observada uma anomalia. A eficiência do *trigger* de busca aparentemente decresce de forma sistemática à medida que se aumenta o valor de E_T^{miss} acima de 500 GeV, figura 4.15. Esse comportamento estranho ainda será alvo de uma investigação mais profunda, e interações com o grupo de *triggers* de SUSY do CMS serão realizadas caso a anomalia persista. Na prática, a anomalia afeta uma quantidade pequena de eventos na cauda da distribuição de E_T^{miss} , podendo produzir comportamentos similares em variáveis tipicamente correlacionadas, como H_T^{miss} , H_T e $\text{Jet}_L^{\text{PV}} p_T$. A figura 4.16 mostra os fatores de escala aplicados aos eventos de Monte Carlo para que estejam de acordo com a eficiência observada nos dados reais.

O erro sistemático atribuído ao valor nominal da eficiência do *trigger* no *plateau*, obtido anteriormente através da amplitude, A , do ajuste da curva de eficiência, será estimado ajustando-se uma reta paralela à distribuição de eficiência dos *triggers* de busca em um conjunto de variáveis importantes na estratégia de busca da análise. O maior valor absoluto da diferença entre o valor nominal e os resultados dos ajustes $\pm 1\sigma$ é considerado a incerteza sistemática da eficiência. Os valores dos ajustes estão na tabela 4.3, em que pode-se constatar que a maior diferença observada foi com respeito à variável H_T^{miss} , que

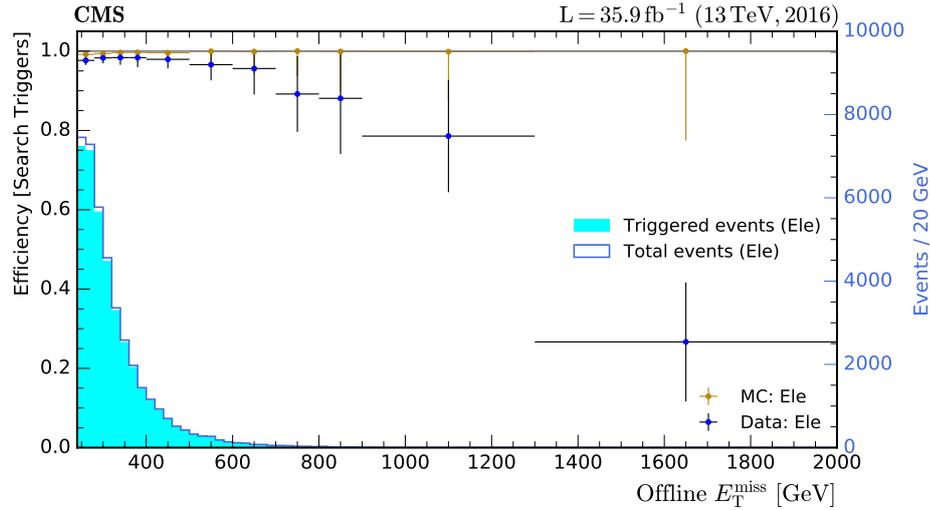


FIGURA 4.15 – Comportamento anômalo do *trigger* de buscas para $E_T^{\text{miss}} > 500 \text{ GeV}$, com respeito aos dados reais, em que a eficiência decresce de forma sistemática. O mesmo comportamento não é observado na amostra de dados de MC em marrom.

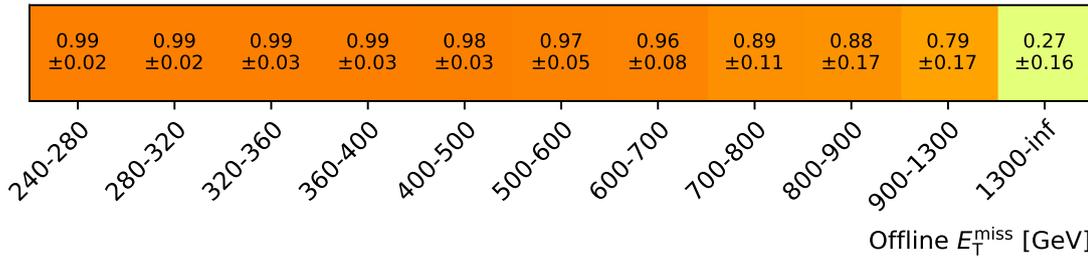


FIGURA 4.16 – Fatores de escala aplicados nos eventos de MC para que fiquem em acordo com a eficiência dos dados reais.

resulta numa diferença de 0,86%. Aproximando para uma casa decimal, a estimativa da eficiência dos *triggers* de busca e seu respectivo erro sistemático é igual a $(97,7 \pm 0,9)\%$. Na figura 4.17 são mostradas as distribuições das variáveis presentes na tabela 4.3, as distribuições de eficiência dos *triggers* de busca em cada uma delas, o valor nominal da eficiência do *trigger* e seu respectivo erro sistemático. A versão da variável H_T^{miss} , referente à primeira distribuição (no topo à esquerda), é a versão padrão descrita na seção 4.1 e não a adotada anteriormente no estudo da eficiência dos *triggers*. A diminuição das eficiências nas caudas das variáveis H_T e $\text{Jet}_L^{\text{PV}} p_T$, embora estatisticamente aceitável, é um efeito sistemático relacionado à anomalia observada na distribuição de E_T^{miss} .

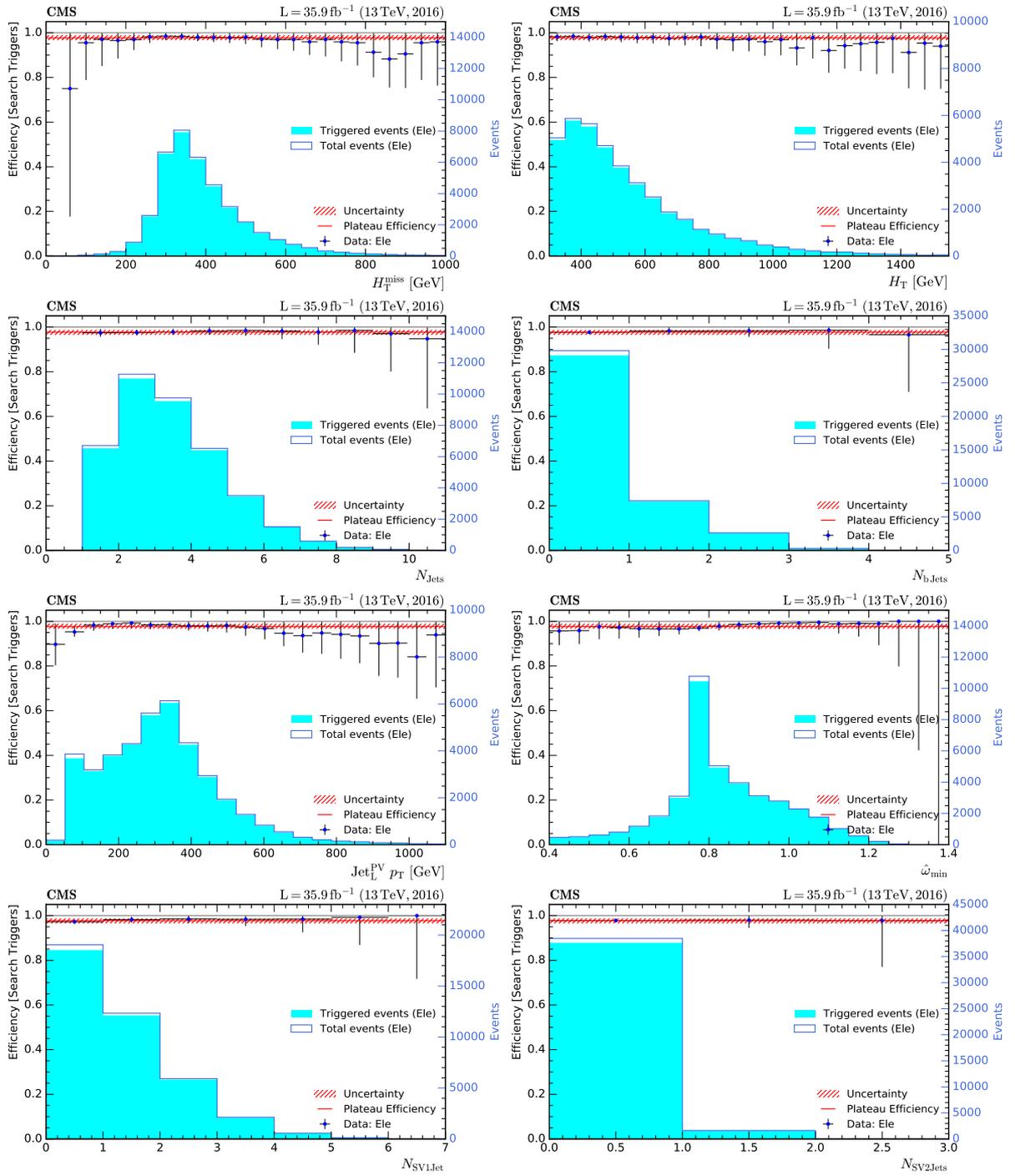


FIGURA 4.17 – Distribuições de eficiência dos *triggers* de busca nas variáveis presentes na tabela 4.3. O valor nominal obtido para a eficiência e seu erro sistemático estão representados em vermelho nos gráficos.

Variável	Valor (%)	Valor + 1σ (%)	Valor - 1σ (%)	Diferença (%)
H_T^{miss}	$97,88 \pm 0,70$	98,58	97,88	0,86
H_T	$97,86 \pm 0,70$	98,56	97,86	0,84
N_{jets}	$97,82 \pm 0,70$	98,52	97,82	0,80
N_{bjets}	$97,82 \pm 0,70$	98,52	97,82	0,80
$\text{Jet}_L^{\text{PV}} p_T$	$97,82 \pm 0,70$	98,52	97,82	0,80
$\hat{\omega}_{\text{min}}$	$97,81 \pm 0,70$	98,51	97,81	0,79
N_{SV1Jet}	$97,82 \pm 0,70$	98,52	97,82	0,80
N_{SV2Jets}	$97,82 \pm 0,70$	98,52	97,82	0,80

TABELA 4.3 – Estimativa do erro sistemático da eficiência dos *triggers* de busca. São apresentados os valores obtidos nos ajustes (coluna 2), referentes a cada variável descrita na coluna 1. Nas colunas 3 e 4 são expostos, respectivamente, os resultados dos ajustes mais 1σ de suas incertezas e menos 1σ de suas incertezas. A coluna 5, mostra os valores absolutos da maior diferença entre os valores das colunas 3 e 4 com relação ao valor nominal, 97,72%.

4.5 Região de busca

A região de busca é projetada para ser uma região com o número de eventos de *background* reduzido, entretanto, sem penalizar muito os eventos de sinal, visto que a análise multivariada realizada posteriormente se encarregará de definir regiões com maior sensibilidade para o sinal buscado. As variáveis utilizadas para reduzir as contribuições de *background* foram o número de elétrons e múons no evento $N_{e,\mu}$ e a variável angular $\hat{\omega}_{\text{min}}$. A proposição de $N_{e,\mu} = 0$ na região de busca elimina um pouco mais de 40% dos eventos de $W(l\nu)+\text{Jatos}$ e $t\bar{t}+\text{Jatos}$ e tem um efeito desprezível nos eventos de sinal, que não envolvem a produção de léptons, onde todos os pontos do *grid* possuem uma eficiência maior que 99% com relação a esse corte. Os eventos de QCD sofrem uma redução drástica com a aplicação do corte $\hat{\omega}_{\text{min}} \geq 0,4$, em que são eliminados mais de 90% dos seus eventos. Além do objetivo principal de eliminar os eventos de QCD, esse corte ainda tem uma razoável performance sobre eventos de $t\bar{t}+\text{Jatos}$, os quais têm em torno de 25% dos seus eventos eliminados enquanto os eventos de sinal respondem ao corte com um eficiência superior a 90%. Na figura 4.18¹⁰, os cortes comentados neste parágrafo são mostrados em sequência sobre os eventos selecionados pelo conjunto de requerimentos a seguir que será referido como **seleção inicial**:

1. Passar pelos itens 1, 2, 4 e 6 da pré-seleção definida na seção 4.4;
2. Ser selecionado pelos *triggers* de busca;

¹⁰Nas figuras, a categoria *residual SM* representa os outros processos de *background*, não citados explicitamente no gráfico por contribuírem com uma quantidade menor de eventos na região.

3. Passar pelos cortes *offline* das variáveis do *trigger*:

$$E_T^{\text{miss}} \geq 240 \text{ GeV e } H_T^{\text{miss}} \geq 220 \text{ GeV;}$$

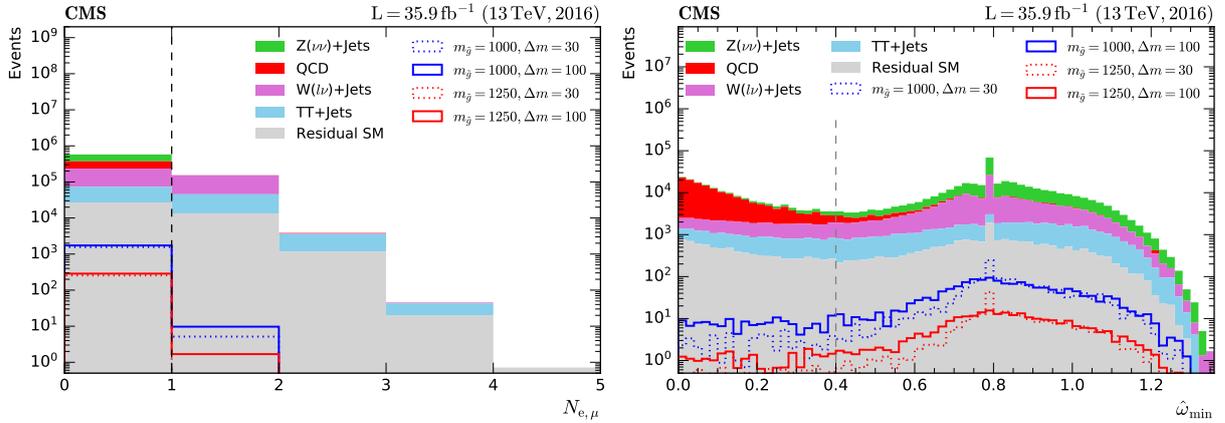


FIGURA 4.18 – O gráfico à esquerda mostra a distribuição dos eventos de *background*, que passaram pela seleção inicial, com respeito ao número de elétrons e múons no evento. A linha vertical tracejada em 1 mostra o ponto de corte, sendo selecionados apenas os eventos com $N_{e,\mu} = 0$. O gráfico à direita mostra a distribuição dos eventos de *background*, que passaram pelo corte em $N_{e,\mu}$, com respeito a variável $\hat{\omega}_{\min}$. A linha vertical tracejada em 0,4 aponta a localização do corte, sendo selecionados apenas os eventos com $\hat{\omega}_{\min} \geq 0,4$.

Após a aplicação dos cortes em $N_{e,\mu}$ e $\hat{\omega}_{\min}$, visando a eliminação de *backgrounds*, um último corte também na variável $\hat{\omega}_{\min}$ é realizado com o objetivo de definir uma região de controle para eventos de $Z(\nu\nu)+\text{Jatos}$. O corte elimina da região de busca o pico observado na distribuição de $\hat{\omega}_{\min}$, no *bin* localizado no intervalo $0,78 \leq \hat{\omega}_{\min} < 0,8$. As características dos eventos nesse *bin* são discutidas na subseção 4.6.4 que descreve a região de controle para eventos com $Z(\nu\nu)+\text{Jatos}$. Esse é o corte mais efetivo sobre os eventos de $Z(\nu\nu)+\text{Jatos}$, onde aproximadamente 22% deles são eliminados da região de busca. O efeito desse corte sobre os eventos de sinal, varia com o valor de Δm , eventos mais compridos tendem a também formar um pico nesse *bin* e perdem mais eventos no corte. No geral, eventos de sinal com $\Delta m = 100 \text{ GeV}$ apresentam eficiência maior que 94% enquanto eventos com $\Delta m = 30 \text{ GeV}$ possuem eficiência em torno de 82-85%. O efeito exato dos cortes em diferentes pontos do *grid* de sinal e em diferentes *backgrounds* é descrito na tabela 4.4, onde são mostrados o número de eventos que sobreviveram após cada corte com seu respectivo erro estatístico, e em azul é mostrado o valor da eficiência relativa com relação ao número de eventos que existiam antes do corte.

A definição exata da região de busca é composta pelos seguintes requerimentos para seleção de eventos:

1. Passar pela pré-seleção definida na seção 4.4;
2. Ser selecionado pelos *triggers* de busca;

Processo físico	Fluxo de cortes			
	Seleção inicial	$N_{e,\mu} = 0$	$\hat{\omega}_{\min} \geq 0,4$	$\hat{\omega}_{\min} < 0,78$ ou $\hat{\omega}_{\min} \geq 0,8$
$m_{\tilde{g}} = 750 \text{ GeV}$ $\Delta m = 30 \text{ GeV}$	10796(218) —	10747(218) 99,5 %	10260(213) 95,5 %	8692(196) 84,7 %
$m_{\tilde{g}} = 750 \text{ GeV}$ $\Delta m = 60 \text{ GeV}$	11268(218) —	11172(217) 99,1 %	10297(208) 92,2 %	9414(199) 91,4 %
$m_{\tilde{g}} = 750 \text{ GeV}$ $\Delta m = 100 \text{ GeV}$	11736(221) —	11679(221) 99,5 %	10547(210) 90,3 %	10085(205) 95,6 %
$m_{\tilde{g}} = 1000 \text{ GeV}$ $\Delta m = 30 \text{ GeV}$	1566(30) —	1561(30) 99,7 %	1508(29) 96,6 %	1264(27) 83,8 %
$m_{\tilde{g}} = 1000 \text{ GeV}$ $\Delta m = 60 \text{ GeV}$	1620(31) —	1608(30) 99,3 %	1497(29) 93,1 %	1371(28) 91,6 %
$m_{\tilde{g}} = 1000 \text{ GeV}$ $\Delta m = 100 \text{ GeV}$	1742(32) —	1732(32) 99,4 %	1581(30) 91,3 %	1487(29) 94,1 %
$m_{\tilde{g}} = 1250 \text{ GeV}$ $\Delta m = 30 \text{ GeV}$	256(5) —	256(5) 100 %	247(5) 96,5 %	204(5) 82,6 %
$m_{\tilde{g}} = 1250 \text{ GeV}$ $\Delta m = 60 \text{ GeV}$	270(5) —	269(5) 99,6 %	253(5) 94,1 %	229(5) 90,5 %
$m_{\tilde{g}} = 1250 \text{ GeV}$ $\Delta m = 100 \text{ GeV}$	288(5) —	287(5) 99,7 %	264(5) 92,0 %	249(5) 94,3 %
Z($\nu\nu$)+Jatos	203549(356) —	203325(355) 99,9 %	191753(351) 94,3 %	149803(303) 78,1 %
$t\bar{t}$ +Jatos	81223(414) —	46780(314) 57,6 %	35310(275) 75,5 %	34191(270) 96,3 %
W($l\nu_l$)+Jatos	267729(400) —	160284(324) 59,9 %	139598(304) 87,1 %	116559(281) 83,5 %
QCD	139530(1507) —	138312(1497) 99,1 %	12526(587) 9,1 %	12006(577) 95,8 %
MP residual	41544(270) —	26966(218) 64,9 %	18936(190) 70,2 %	17007(178) 89,8 %

TABELA 4.4 – Fluxo de cortes sobre alguns pontos do *grid* de sinal e sobre os *backgrounds*, começando com os eventos que passaram pela seleção inicial.

3. Passar pelos cortes *offline* das variáveis do *trigger*:

$$E_{\text{T}}^{\text{miss}} \geq 240 \text{ GeV} \text{ e } H_{\text{T}}^{\text{miss}} \geq 220 \text{ GeV};$$

4. Não possuir $\hat{\omega}_{\min}$ no intervalo $7,8 \leq \hat{\omega}_{\min} < 8$.

A região de busca é dividida em sub-regiões, cujo objetivo principal visa a otimização do uso da informação disponível em cada evento reconstruído. Tratando-se de uma busca por um sinal hadrônico, sem léptons, em que é extremamente difícil ter todos os jatos reconstruídos, devido às limitações do experimento, a estratégia de busca consiste de dividir os eventos em categorias definidas pela quantidade de informação reconstruída do decaimento dos gluinos. Para isto, são usados os objetos definidos na subseção 4.1.1.

O SV1Jet, nos eventos de sinal, contém informação sobre o decaimento do gluino em que só foi possível a reconstrução de um dos jatos resultantes ou o segundo jato foi associado erroneamente a outro vértice; e o SV2Jets contém a informação completa da parte hadrônica do decaimento, ou seja, os dois jatos foram reconstruídos. Eventos de sinal que possuem o objeto SV2Jets têm mais informação útil para ser usada na discriminação entre eventos de sinal e *background*, eles também apresentam menor contaminação de *pileup*, entretanto, poucos eventos possuem esse objeto reconstruído, e os que possuem, geralmente têm apenas um. Eventos que não possuem um SV2Jets, porém têm um ou mais SV1Jet's reconstruídos, ainda têm informações úteis sobre o decaimento do gluino que podem ser usadas na busca. Os SV1Jet's são reconstruídos com muito mais frequência que os SV2Jets's, todavia apresentam maior contaminação por *pileup*. São definidas quatro sub-regiões de busca, utilizando como critério a quantidade de SV1Jet's e SV2Jets's presentes no evento,

- **Região 1 (R1):** O evento possui pelo menos um SV2Jets;
- **Região 2 (R2):** O evento não possui SV2Jets's e possui dois ou mais SV1Jet's;
- **Região 3 (R3):** O evento não possui SV2Jets's e possui um SV1Jet;
- **Região 4 (R4):** O evento não possui SV2Jets's nem SV1Jet's.

Em cada sub-região de busca será realizada uma análise multivariada que usará o máximo de informação relevante disponível nos eventos, para definir regiões com maior sensibilidade aos eventos de sinal, regiões de sinal (SR).

Na figura 4.19, é mostrada uma representação esquemática das sub-regiões de busca. Na região 1, nos poucos casos em que forem reconstruídos mais de um SV2Jets, será utilizado na análise o SV2Jets com maior significância na posição do vértice secundário, S_{SV} , definida como

$$S_{SV} \equiv \frac{L_{PV-SV}}{\sigma_{SV}}, \quad (4.13)$$

tal que L_{PV-SV} é a distância entre o vértice primário e o secundário e σ_{SV} é a incerteza na posição do vértice secundário. Similarmente na região 2, nos casos em que forem reconstruídos mais de dois SV1Jet's, serão utilizados na análise os dois SV1Jet's com maior S_{SV} .

A definição das sub-regiões de busca propicia uma boa distribuição dos eventos de sinal entre elas. Como a definição depende do número de SV1Jet's e SV2Jets's, já era de se esperar que a massa do gluino não afetasse a distribuição de eventos entre as sub-regiões, esse comportamento é visto no gráfico à esquerda da figura 4.20, em que são expostas as frações de eventos em cada sub-região para três pontos distintos do *grid* de sinal com valores diferentes de $m_{\tilde{g}}$, porém com o mesmo valor de $\Delta m = 60 \text{ GeV}$. Os pontos do

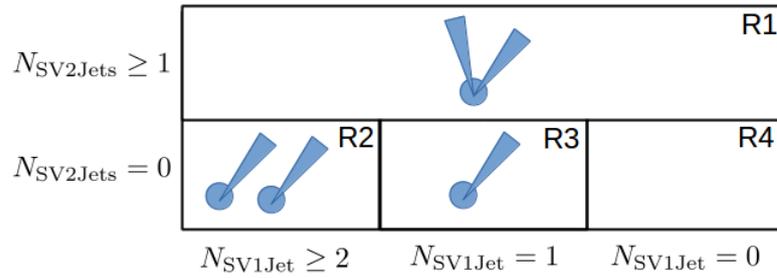


FIGURA 4.19 – Representação esquemática das regiões de busca. Os cones representam os jatos e os círculos representam os vértices secundários do SV2Jets ou SV1Jet’s presentes em cada sub-região.

grid de sinal para $\Delta m = 60$ GeV, apresentam uma distribuição quase igualitária entre as sub-regiões, onde cada uma possui entre 22-30% dos eventos. Contudo, quando consideramos pontos do *grid* com um grau de compressão maior ou menor, as diferenças entre as contribuições para cada sub-região se acentuam. Modelos mais comprimidos apresentam maior dificuldade na reconstrução dos jatos provenientes dos decaimentos dos gluínos, isso tende a gerar uma quantidade maior de eventos nas regiões que requerem menos objetos reconstruídos, como pode ser visto na distribuição correspondente ao $\Delta m = 30$ GeV no gráfico à direita na figura 4.20. Já no caso de modelos com menos compressão, ocorre o efeito inverso, e isso pode ser visto na distribuição do modelo com $\Delta m = 100$ GeV, entre as sub-regiões R2, R3 e R4. A fração de eventos em R1 é praticamente a mesma entre os modelos com $\Delta m = 60$ GeV e $\Delta m = 100$ GeV, embora a fração correspondente a $\Delta m = 30$ GeV seja aproximadamente 3 vezes menor. Isso indica uma certa resistência na reconstrução de SV2Jets’s mesmo em modelos menos comprimidos, provavelmente relacionada à dificuldade de reconstruir tais objetos. Portanto modelos mais e menos comprimidos geram distribuições menos homogêneas com contribuições entre 8-42% para o caso de maior compressão do *grid*, $\Delta m = 30$ GeV, e entre 14-40% para o de menor compressão, $\Delta m = 100$ GeV.

4.6 Modelagem do Background

O próximo estágio da análise será a realização de uma análise multivariada utilizando as simulações de Monte Carlo do sinal e dos backgrounds para definir as regiões de sinal da busca. Logo, é necessário que as amostras de simulação representem bem os dados reais. Por essa razão, são feitas correções nas simulações dos *backgrounds* dominantes para que eles descrevam bem o comportamento dos dados reais em todas as variáveis usadas na análise multivariada dentro da região de busca. As correções são realizadas em regiões de

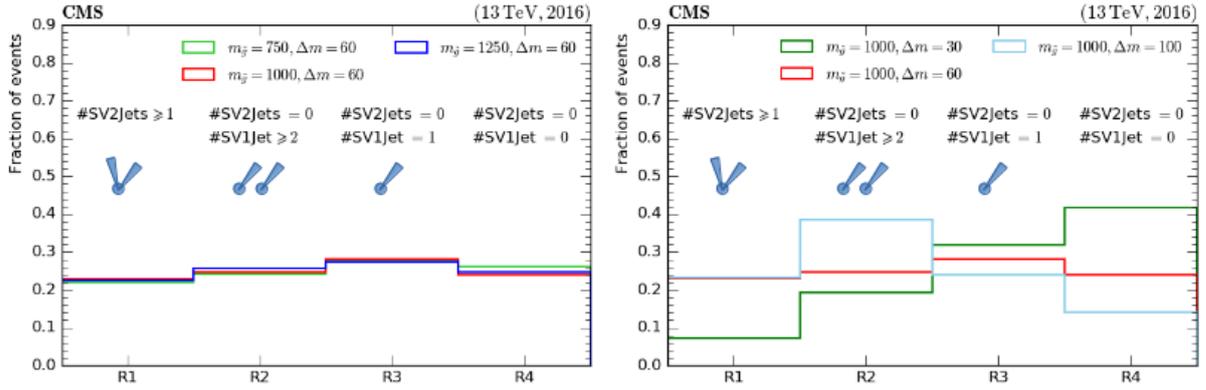


FIGURA 4.20 – Fração do número de eventos de sinal em cada sub-região de busca. No gráfico à esquerda são expostos três pontos do *grid* de sinal com $\Delta m = 60$ GeV e diferentes valores de $m_{\tilde{g}}$, já no gráfico à direita são expostos três pontos do *grid* de sinal com $m_{\tilde{g}} = 1000$ GeV e diferentes valores de Δm .

controle (*control regions* - CR)¹¹ definidas para cada *background* dominante, e que visam estar o mais próximo possível da definição da região de busca. Isso é feito para garantir que os eventos presentes nas regiões de controle e de busca possuam características similares. Os quatro processos físicos do MP que compõem a maior parte do *background* dessa busca são os eventos com $Z(\nu\nu)+\text{Jatos}$, $W(l\nu)+\text{Jatos}$, $t\bar{t}+\text{Jatos}$ e QCD, conforme foi mostrado na tabela 4.4. A porcentagem da contribuição de cada processo no *background* das quatro sub-regiões de busca é exposta na figura 4.21.

As regiões de controle dos backgrounds dominantes são descritas no esquema da figura 4.22, em que os eventos são previamente selecionados pelo conjunto de requerimentos da seleção inicial, definidos na seção 4.5. As regiões de controle dos eventos de $t\bar{t}+\text{Jatos}$ e $W(l\nu)+\text{Jatos}$ diferem da região de busca pela presença de elétrons ou múons no evento, e diferem entre si pelo número de b-jatos. Não são esperados b-jatos nos eventos de $W(l\nu)+\text{Jatos}$, a não ser que os b-jatos sejam produzidos em alguma emissão de radiação, por isso exige-se que na região de controle dos eventos $W(l\nu)+\text{Jatos}$ não haja b-jatos. Já nos eventos de $t\bar{t}+\text{Jatos}$ são esperados dois b-jatos, considerando que o quark top quase sempre decai¹² no bóson W e um quark b. A região de controle dos eventos de QCD difere da região de busca apenas pelo intervalo na variável $\hat{\omega}_{\min}$, a qual deve ser menor que 0,15. Por fim, temos os eventos com $Z(\nu\nu)+\text{Jatos}$, para os quais é tipicamente difícil definir uma região de controle em buscas totalmente hadrônicas, visto que $Z(\nu\nu)+\text{Jatos}$ possui uma topologia muito semelhante aos eventos de sinal. Entretanto, uma característica da variável $\hat{\omega}_{\min}$, que resulta em um pico na distribuição dos backgrounds dentro do intervalo $7,8 \leq \hat{\omega}_{\min} < 8$, é usada para definir a região de controle dos eventos com $Z(\nu\nu)+\text{Jatos}$,

¹¹Região onde determinado processo físico do MP é dominante, possuindo a maior fração de eventos. A pureza é a quantidade que expressa esse domínio, sendo definida pela razão entre o número de eventos do processo e o número total de eventos.

¹²A razão de ramificação dos outros canais de decaimento é menor que 10^{-2} [8].

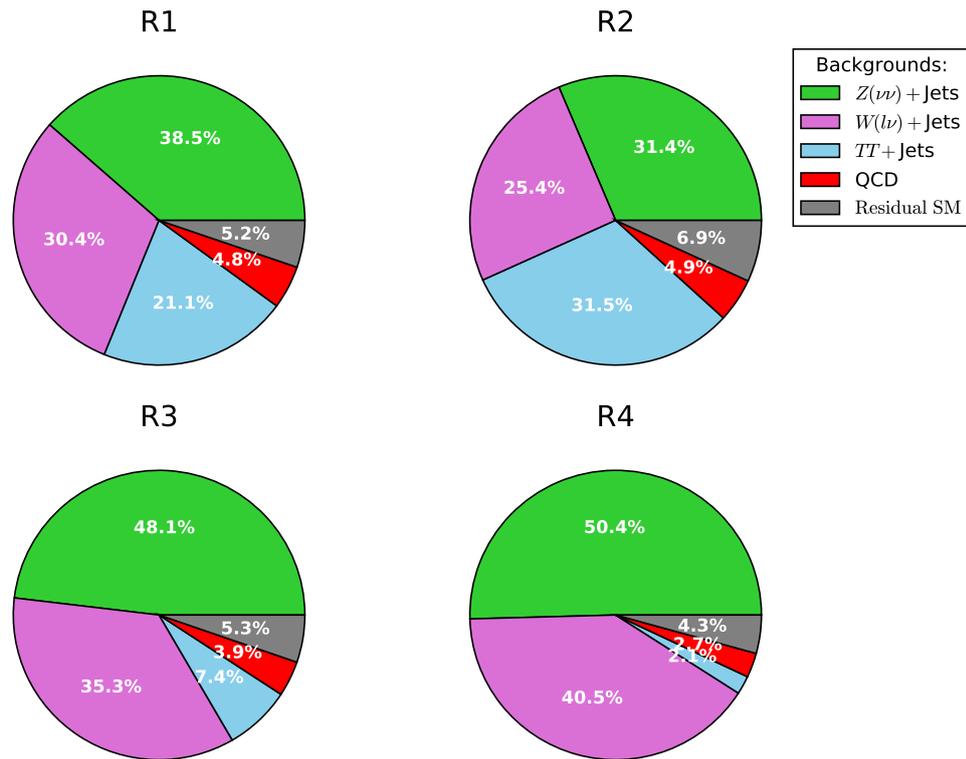


FIGURA 4.21 – Estimativa, baseada na simulação, dos percentuais das contribuições de cada *background* nas quatro sub-regiões de busca.

que também faz uso de uma análise multivariada. A definição exata da região de controle para os eventos com $Z(\nu\nu) + \text{Jatos}$ é realizada na subseção 4.6.4.

4.6.1 $W(l\nu) + \text{Jatos}$

Produção via interação eletrofraca do bóson W , o qual decai em um lépton carregado e um neutro (neutrino) da mesma família. A energia transversa perdida está associada ao neutrino produzido, e os eventos que entram na região de busca são aqueles que não tiveram o lépton carregado reconstruído, o que pode ocorrer por estarem fora da região de aceitação do detector, deixarem um sinal muito fraco no detector e não serem reconstruídos ou serem produzidos sobrepostos a alguma atividade hadrônica no evento e, desse modo, não passarem no corte de isolamento. Além do decaimento leptônico, o bóson W tem em torno de 67% de probabilidade de decair em hádrons [8], mas estes eventos não passam pela seleção da região de busca, principalmente por possuírem pouca energia transversa perdida.

O conjunto de correções aos eventos com $W(l\nu) + \text{Jatos}$ é dividido em três etapas, aplicadas em sequência como descrito a seguir.

1. **Fator de escala global:** é o fator de escala aplicado a todos os eventos simulados

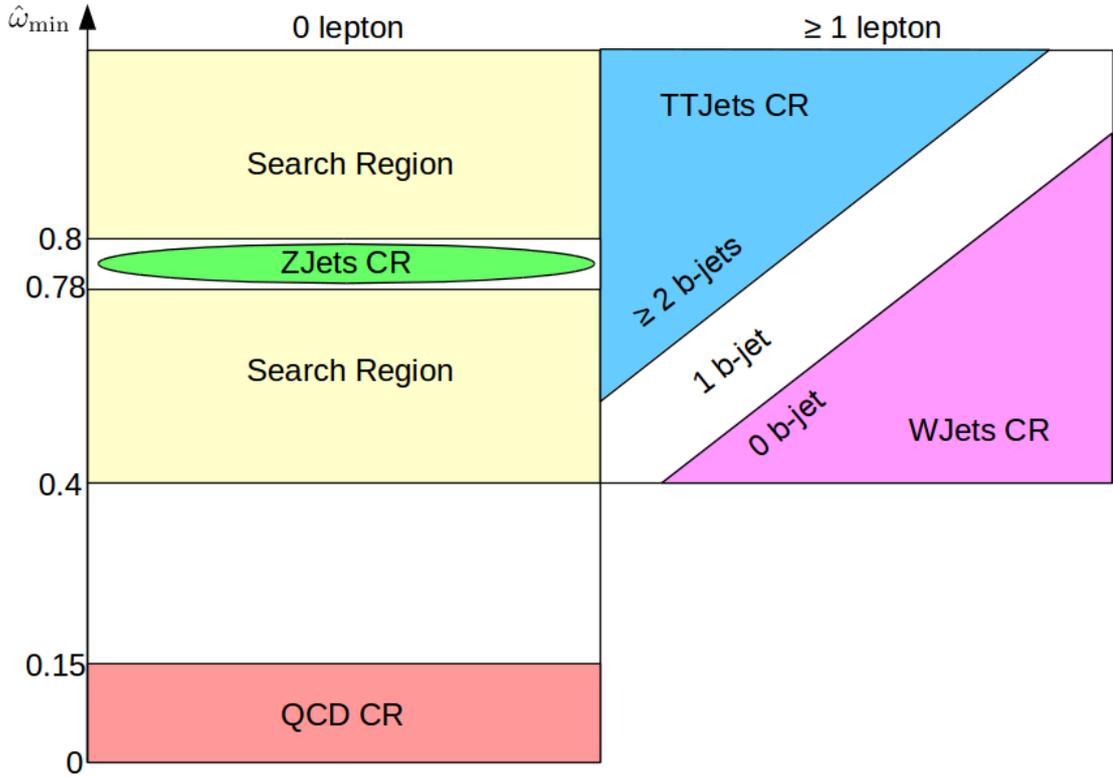


FIGURA 4.22 – Esquema das regiões definidas na análise, todas elas consideram que os eventos tenham passado pela seleção inicial. A região amarela é a região de busca. As regiões de controle dos eventos de $t\bar{t}$ +Jatos, $W(l\nu_l)$ +Jatos, QCD e $Z(\nu\nu)$ +Jatos estão, respectivamente, em azul, roxo, vermelho e verde.

com $W(l\nu_l)$ +Jatos e será denotado por F_W^G . O objetivo desse fator de escala é igualar o número de eventos simulados ao número de eventos reais na região de controle. Visto que a principal contaminação na região de controle dos $W(l\nu_l)$ +Jatos, com pureza igual a 85,7%, consiste de eventos de $t\bar{t}$ +Jatos e vice-versa, F_W^G será calculado em conjunto com o fator de escala global dos eventos de $t\bar{t}$ +Jatos, F_{TT}^G . O conjunto composto por essas duas classes de eventos apresenta uma pureza maior que 92,2%, conseqüentemente produz resultados mais apurados. Para realização do cálculo é adotada a seguinte nomenclatura:

- N_{TT}^{CRTT} - número de eventos simulados de $t\bar{t}$ +Jatos, na região de controle de $t\bar{t}$ +Jatos.
- N_W^{CRTT} - número de eventos simulados de $W(l\nu_l)$ +Jatos, na região de controle de $t\bar{t}$ +Jatos.
- N_{Data}^{CRTT} - número de eventos dos dados reais subtraídos os eventos simulados dos outros backgrounds diferentes de $t\bar{t}$ +Jatos e $W(l\nu_l)$ +Jatos, na região de controle de $t\bar{t}$ +Jatos.
- N_{TT}^{CRW} - número de eventos simulados de $t\bar{t}$ +Jatos, na região de controle de

$W(l\nu_l)+\text{Jatos}$.

- N_W^{CRW} - número de eventos simulados de $W(l\nu_l)+\text{Jatos}$, na região de controle de $W(l\nu_l)+\text{Jatos}$.
- N_{Data}^{CRW} - número de eventos dos dados reais subtraídos os eventos simulados dos outros backgrounds diferentes de $t\bar{t}+\text{Jatos}$ e $W(l\nu_l)+\text{Jatos}$, na região de controle de $W(l\nu_l)+\text{Jatos}$.

Uma vez definida a nomenclatura das quantidades, é possível formar um sistema de duas equações que expressam a contribuição de $t\bar{t}+\text{Jatos}$ e $W(l\nu_l)+\text{Jatos}$ em cada região de controle, de maneira que as duas incógnitas serão os fatores de escala F_{TT}^G e F_W^G . Escrevendo o sistema em forma de uma relação matricial, e usando de relações algébricas, temos que os fatores de escala globais podem ser obtidos através da expressão,

$$\begin{bmatrix} F_{TT}^G \\ F_W^G \end{bmatrix} = \begin{bmatrix} N_{TT}^{CRTT} & N_W^{CRTT} \\ N_{TT}^{CRW} & N_W^{CRW} \end{bmatrix}^{-1} \begin{bmatrix} N_{Data}^{CRTT} \\ N_{Data}^{CRW} \end{bmatrix}. \quad (4.14)$$

O expoente -1 denota a matriz inversa. Os valores obtidos por esse método estão reunidos na tabela 4.5 junto com as incertezas das medições.

	Fator de escala
F_W^G	$0,977 \pm 0,023$
F_{TT}^G	$0,816 \pm 0,017$

TABELA 4.5 – Fatores de escala global para eventos de $t\bar{t}+\text{Jatos}$ e $W(l\nu_l)+\text{Jatos}$.

2. **Correções de forma:** As correções de forma (*shape corrections*) são correções feitas nos eventos simulados de modo a garantir, na região de controle, um bom acordo entre as simulações e os dados reais nas distribuições das variáveis de interesse da análise. O procedimento adotado na correção de forma das amostras simuladas de $W(l\nu_l)+\text{Jatos}$, e que também será usado nos outros processos de *background*, calcula as correções para os eventos em sub-regiões delimitadas pelos valores de duas variáveis escolhidas apropriadamente. As sub-regiões usadas para aplicar as correções são escolhidas por apresentarem um desacordo constante entre os dados reais e a simulação e possuírem uma quantidade de eventos estatisticamente relevante. As correções de forma dos eventos de $W(l\nu_l)+\text{Jatos}$ são mostradas na figura 4.23. As variáveis escolhidas foram $N_{\text{jets}}^{\text{PV}}$ e H_{T}^{PV} , variáveis relacionadas à atividade hadrônica na radiação de estado inicial, que em geral apresentam discrepâncias entre os dados reais e a simulação.
3. **Correções de escala regionais:** Após as correções de forma, são aplicadas correções de escala em cada uma das quatro sub-regiões da região de controle, definidas

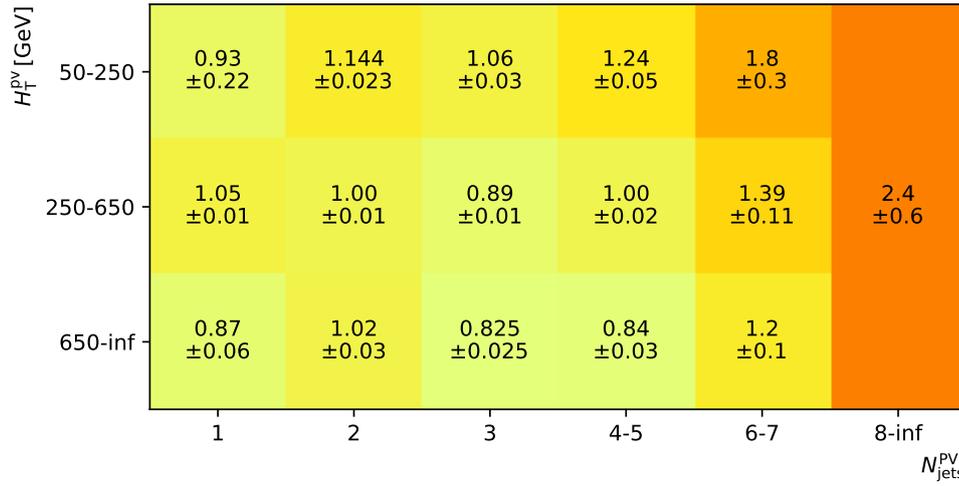


FIGURA 4.23 – Correções de forma dos eventos de $W(l\nu_l)+\text{Jatos}$ e suas incertezas, aplicadas em sub-regiões definidas pelas variáveis $N_{\text{jets}}^{\text{PV}}$ e H_T^{PV} .

da mesma maneira apresentada na figura 4.19. Os eventos de $W(l\nu_l)+\text{Jatos}$ são multiplicados por um fator de escala regional para garantir, que em cada sub-região, o número de eventos simulados seja igual ao de dados reais. A tabela 4.6 mostra os fatores de escala regionais obtidos e suas incertezas.

Região	Fator de escala
F_W^{R1}	$1,10 \pm 0,03$
F_W^{R2}	$1,06 \pm 0,02$
F_W^{R3}	$1,00 \pm 0,01$
F_W^{R4}	$0,989 \pm 0,006$

TABELA 4.6 – Fatores de escala regionais para eventos de $W(l\nu_l)+\text{Jatos}$.

As distribuições das variáveis globais na região de controle dos $W(l\nu_l)+\text{Jatos}$, após serem aplicadas todas as correções, podem ser vistas na figura 4.24. É observado um bom acordo entre os dados reais e a simulação. Deve-se ressaltar que as incertezas sistemáticas ainda não estão incluídas nesses gráficos.

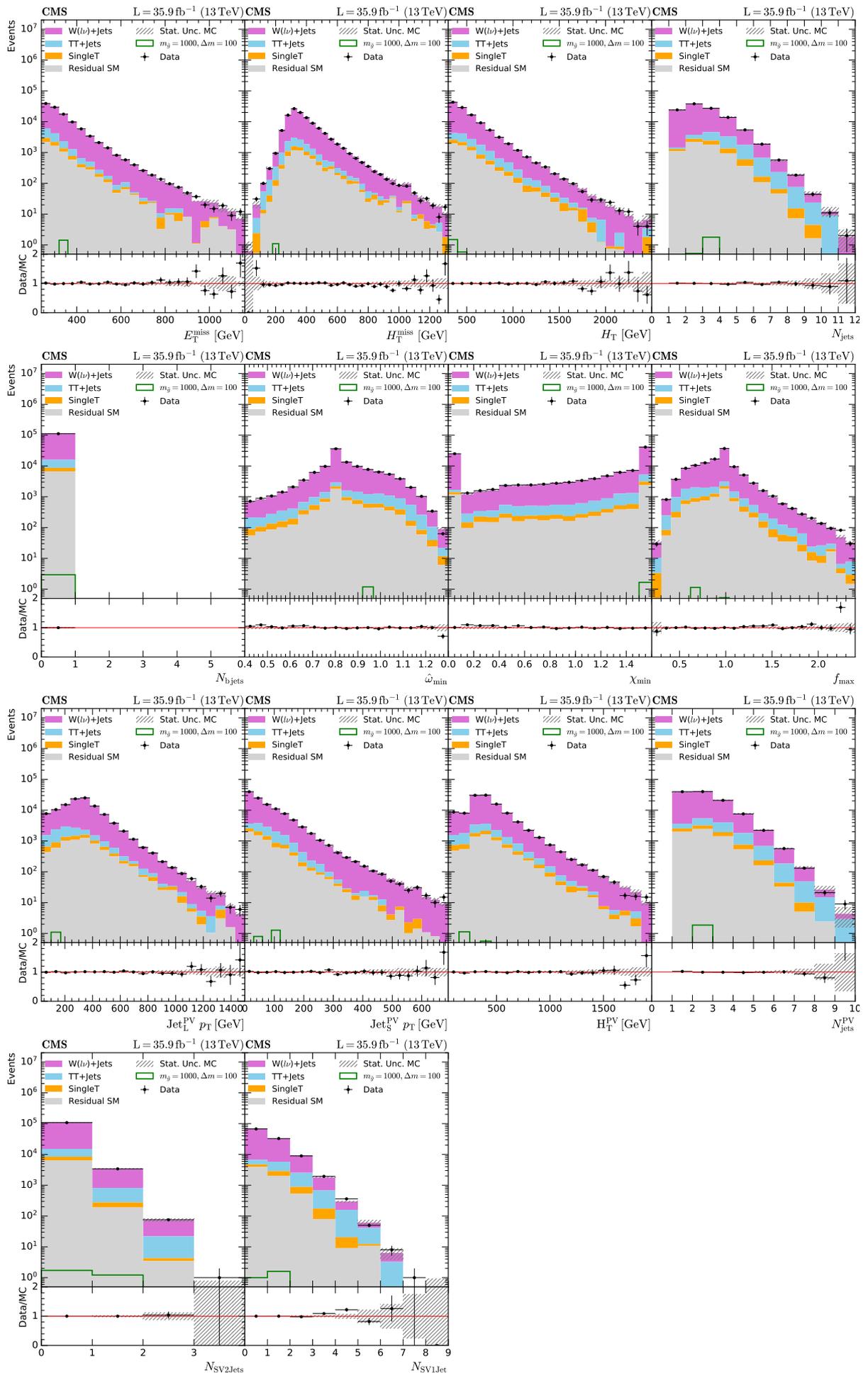


FIGURA 4.24 – Distribuições das variáveis globais na região de controle de $W(l\nu)+\text{Jatos}$, após as correções da simulação.

4.6.2 $t\bar{t}$ +Jatos

Em colisores hadrônicos, os quarks top são produzidos principalmente em pares, sendo o canal $gg \rightarrow t\bar{t}$ responsável pela maior parte de sua produção. O quark top tem um tempo de vida muito curto e decai antes da hadronização ocorrer, quase que 100% das vezes em um bóson W e um quark b [8]. A típica existência de dois quarks b nos produtos do seu decaimento, permite usar essa propriedade na definição da região de controle dos $t\bar{t}$ +Jatos. A energia transversa perdida é originária do neutrino produzido no decaimento leptônico do bóson W. Os eventos que entram na região de busca não possuem léptons carregados pelos mesmos motivos expostos para os eventos de $W(l\nu)$ +Jatos.

O conjunto de correções aos eventos com $t\bar{t}$ +Jatos é dividido em 4 etapas, aplicadas em sequência como descrito a seguir.

1. **Correção da radiação de estado inicial:** os eventos de $t\bar{t}$ +Jatos precisam de uma correção adicional àquelas aplicadas aos outros *backgrounds* dominantes. Essa correção está relacionada à dificuldade de simular a radiação de estado inicial (*Initial State Radiation* - ISR) da produção de eventos com $t\bar{t}$ +Jatos que ocorre via interação forte. O grupo de SUSY do CMS recomenda em [135] a aplicação de correções dependentes do número de jatos originários da radiação de estado inicial, $N_{\text{ISR,jets}}$. Os fatores de correção e suas incertezas são expostos na tabela 4.7.

$N_{\text{ISR,jets}}$	Correção $\pm \sigma_{\text{sist}}$
0	—
1	$0,92 \pm 0,04$
2	$0,82 \pm 0,09$
3	$0,72 \pm 0,14$
4	$0,66 \pm 0,17$
5	$0,56 \pm 0,22$
≥ 6	$0,51 \pm 0,24$

TABELA 4.7 – Fatores de correção da radiação de estado inicial dos eventos de $t\bar{t}$ +Jatos.

2. **Fator de escala global:** Como foi exposto na subseção 4.6.1, o fator de escala global dos eventos de $t\bar{t}$ +Jatos foi computado junto com o de $W(l\nu)$ +Jatos, e o resultado é mostrado na tabela 4.5. A pureza da região de controle de $t\bar{t}$ +Jatos é de 81,8%, mas se considerarmos o conjunto dos eventos de $t\bar{t}$ +Jatos mais $W(l\nu)$ +Jatos, a pureza se torna de 88,5%.
3. **Correções de forma:** As correções de forma para os eventos de $t\bar{t}$ +Jatos são definidas nas sub-regiões delimitadas pelos valores das variáveis $H_{\text{T}}^{\text{miss}}$ e H_{T}^{PV} . Embora tenha menos eventos, a amostra de $t\bar{t}$ +Jatos mostrou um melhor acordo com os

dados que a amostra de $W(l\nu)+\text{Jatos}$ e por isso necessitou de um número menor de sub-regiões para a realização das correções.

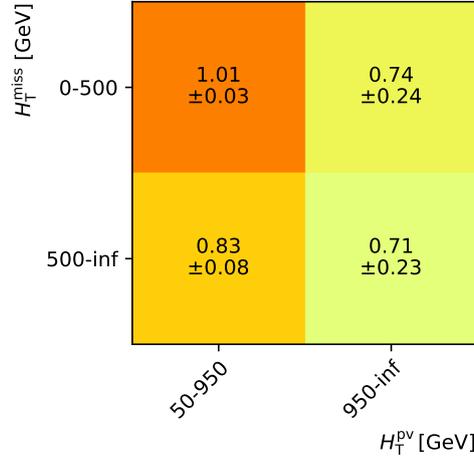


FIGURA 4.25 – Correções de forma dos eventos de $t\bar{t}+\text{Jatos}$ e suas incertezas, aplicadas em sub-regiões definidas pelas variáveis H_T^{miss} e H_T^{PV} .

4. **Correções de escala regionais:** A última correção a ser aplicada é a correção de escala regional. A tabela 4.8 mostra os fatores de escala regionais obtidos para os eventos de $t\bar{t}+\text{Jatos}$ e suas incertezas.

Região	Fator de escala
F_{TT}^{R1}	$1,21 \pm 0,09$
F_{TT}^{R2}	$1,009 \pm 0,023$
F_{TT}^{R3}	$0,85 \pm 0,05$
F_{TT}^{R4}	$0,99 \pm 0,18$

TABELA 4.8 – Fatores de escala regionais para eventos de $t\bar{t}+\text{Jatos}$.

As distribuições das variáveis globais na região de controle dos $t\bar{t}+\text{Jatos}$, após todas as correções, são mostradas na figura 4.26. Um bom acordo entre os dados reais e a simulação é obtido. Os erros estatísticos são tipicamente maiores que os observados nas distribuições de $W(l\nu)+\text{Jatos}$ por causa do menor número de eventos presentes na amostra de $t\bar{t}+\text{Jatos}$.

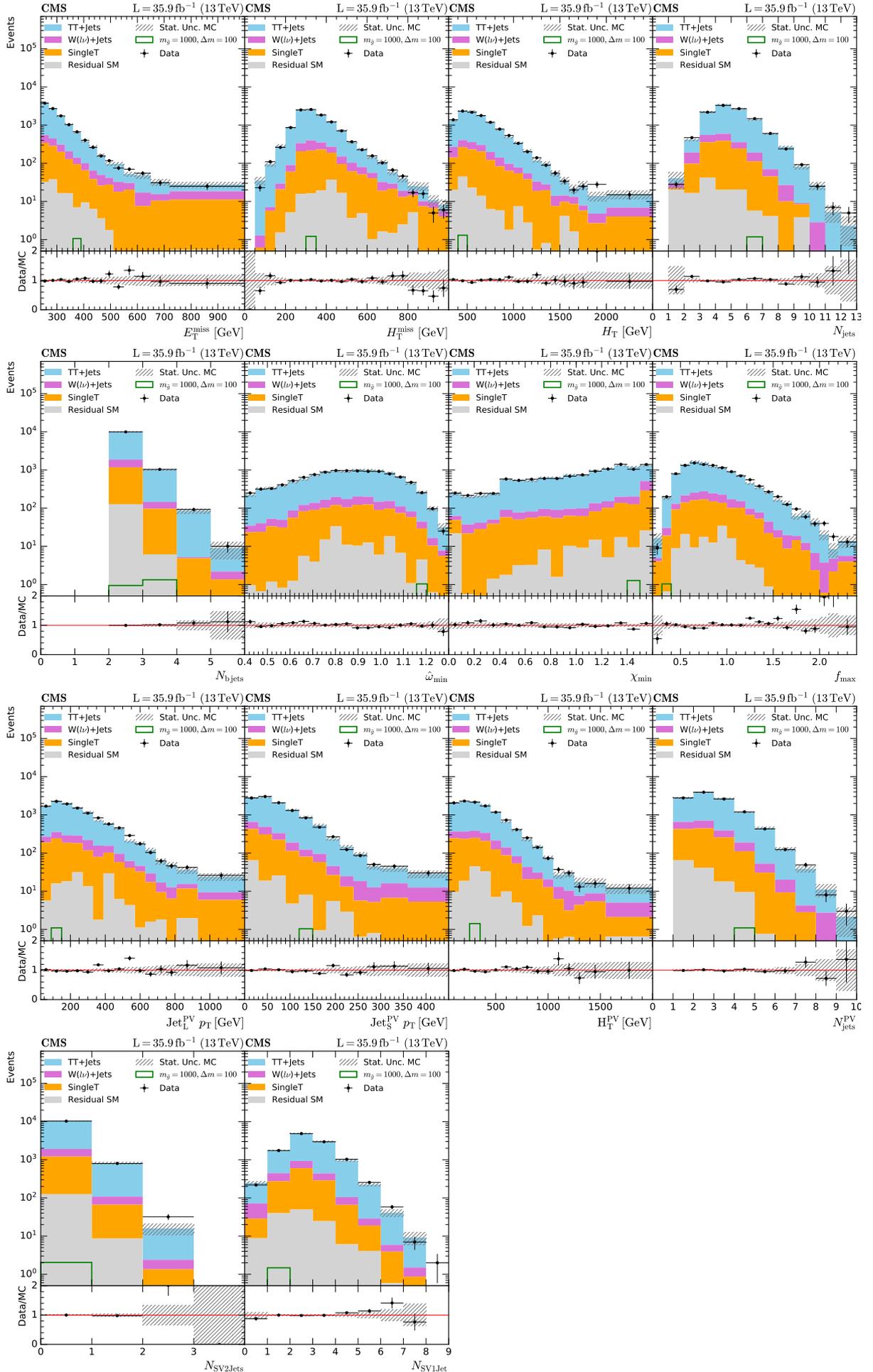


FIGURA 4.26 – Distribuições das variáveis globais na região de controle de $t\bar{t}$ +Jatos, após as correções da simulação.

4.6.3 QCD

Eventos com multijatos, em que os jatos estão associados aos quarks leves produzidos via interação forte no espalhamento duro. A energia transversa perdida nos eventos de QCD é resultado de jatos do evento com p_T mal medido, causando um desbalanço nos momentos transversos. Os eventos de QCD foram suprimidos com o corte em $\hat{\omega}_{\min} \geq 0,4$, em consequência disso, sua contribuição esperada na região de busca é muito menor que a dos outros três *backgrounds* dominantes. Além disso, por serem de difícil simulação e possuírem a maior seção de choque de produção, suas amostras de simulação não são produzidas com uma quantidade adequada de eventos para uma descrição precisa do *background*. A estimativa de eventos de QCD, nas análises em geral, é feita pelo uso de métodos baseados no estudo dos dados reais (*data-driven methods*), o mesmo será feito nesta análise, porém é necessário que a simulação de QCD tenha um acordo razoável com relação aos dados reais para a realização da análise multivariada. Em virtude disso, é realizado nos eventos de QCD o mesmo tratamento de correções aplicados aos outros *backgrounds* dominantes. A região de controle difere da região de busca somente pelo intervalo considerado da variável $\hat{\omega}_{\min}$, sendo definida com o corte superior em $\hat{\omega}_{\min} < 0,15$ visando manter sua pureza maior que 80% (o valor exato é igual a 81,1%). A contaminação de sinal relativa ao modelo com $m_{\tilde{g}} = 1000$ GeV e $\Delta m = 100$ GeV é de 0,06%, figura 4.27.

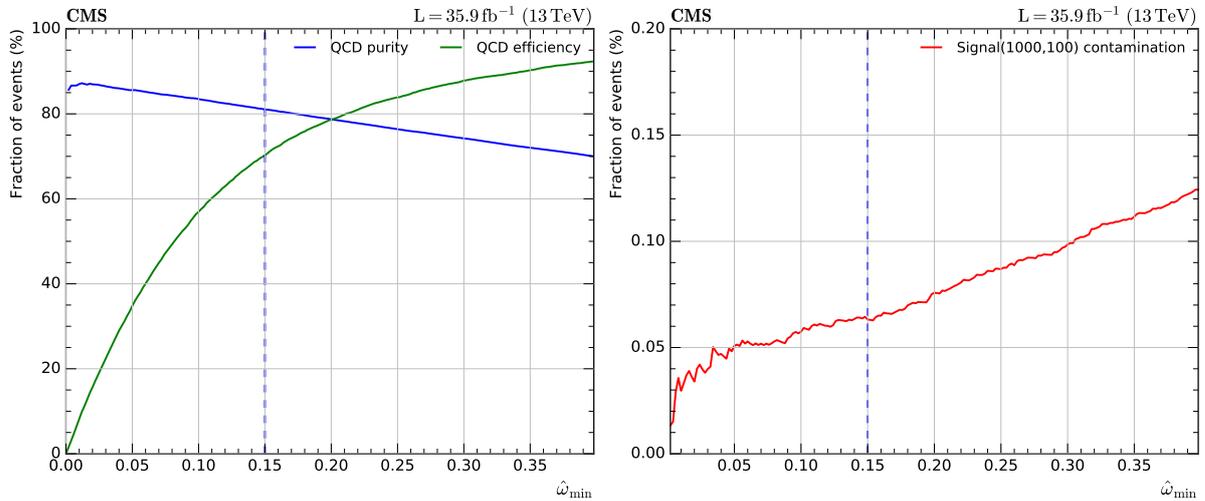


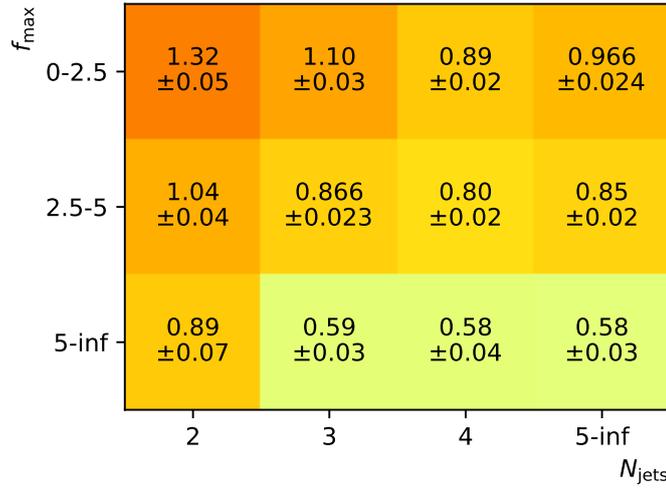
FIGURA 4.27 – Gráfico de controle da pureza e eficiência dos eventos de QCD com relação aos pontos de corte na variável $\hat{\omega}_{\min}$ (esquerda). Gráfico de controle da contaminação por eventos da amostra de sinal, com $m_{\tilde{g}} = 1000$ GeV e $\Delta m = 100$ GeV, com relação aos pontos de corte na variável $\hat{\omega}_{\min}$ (direita). A linha tracejada vertical denota o valor de $\hat{\omega}_{\min}$ em que foi aplicado o corte que define a região de controle de QCD.

1. **Fator de escala global:** o fator de escala global dos eventos de QCD é mostrado na tabela 4.9, e é estimado como o fator multiplicativo nos pesos dos eventos simulados que faz com que o número de eventos simulados seja igual ao número de eventos reais na região de controle.

	Fator de escala
F_{QCD}^G	$1,152 \pm 0,014$

TABELA 4.9 – Fator de escala global para eventos de QCD.

2. **Correções de forma:** As correções de forma para os eventos de QCD são definidas nas sub-regiões delimitadas pelos valores das variáveis N_{jets} e f_{max} , como especificado na figura 4.28.

FIGURA 4.28 – Correções de forma dos eventos de QCD e suas incertezas, aplicadas em sub-regiões definidas pelas variáveis N_{jets} e f_{max} .

3. **Correções de escala regionais:** Como nos casos anteriores, a última etapa é destinada às correções de escala regionais. A tabela 4.10 mostra os resultados obtidos para os eventos de QCD e suas incertezas.

Região	Fator de escala
F_{QCD}^{R1}	$1,03 \pm 0,05$
F_{QCD}^{R2}	$0,98 \pm 0,02$
F_{QCD}^{R3}	$0,98 \pm 0,02$
F_{QCD}^{R4}	$1,04 \pm 0,03$

TABELA 4.10 – Fatores de escala regionais para eventos de QCD.

As distribuições das variáveis globais na região de controle de QCD após todas as correções são mostradas na figura 4.29. Foi alcançado um acordo razoável entre dados reais e a simulação.

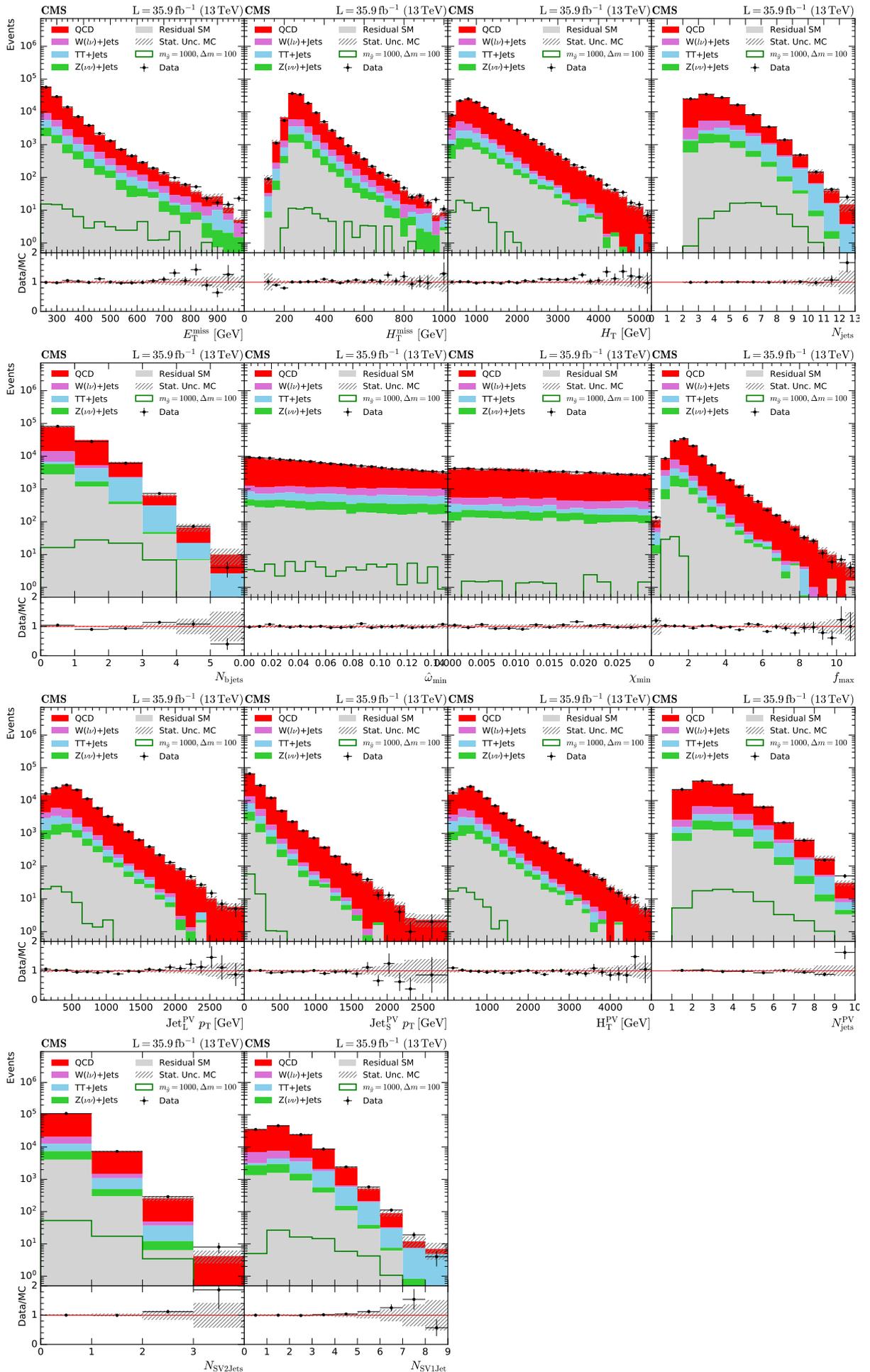


FIGURA 4.29 – Distribuições das variáveis globais na região de controle de QCD após as correções da simulação.

4.6.4 $Z(\nu\nu)+\text{Jatos}$

Produção via interação eletrofraca do bóson Z , o qual decai posteriormente em dois neutrinos. A energia transversa perdida corresponde aos neutrinos produzidos. Os eventos com $Z(\nu\nu)+\text{Jatos}$ são difíceis de serem suprimidos em buscas com estados finais totalmente hadrônicos, e também é difícil a definição de uma região de controle para esse processo. Nesta análise, usa-se o pico característico da distribuição de eventos com $Z(\nu\nu)+\text{Jatos}$, com respeito à variável angular $\hat{\omega}_{\min}$, para definir sua região de controle. O pico está localizado na região com $0,78 \leq \hat{\omega}_{\min} < 0,8$, como mostrado na figura 4.18. Nesse intervalo, estão localizados os eventos com $f_{\max} = 1$ e $\Delta\varphi = \pi^{13}$, valores típicos de eventos com apenas um jato, que são os causadores do pico nessa região. A variável $\hat{\omega}_{\min}$ mostrou-se ser bem simulada nas amostras de MC em geral, sendo uma ótima candidata para ser usada na definição de regiões de controle. O intervalo selecionado ainda contém uma boa quantidade de eventos com mais de um jato, e que varrem completamente os valores possíveis das outras variáveis hadrônicas, que geralmente apresentam discordâncias com os dados reais e precisam ser corrigidas. As quatro sub-regiões de busca também apresentam uma razoável quantidade de eventos, já que os SV1Jet's e SV2Jets's usam jatos com um limiar de corte inferior ao considerado no cálculo de $\hat{\omega}_{\min}$.

Dentro do intervalo $0,78 \leq \hat{\omega}_{\min} < 0,8$, é realizada uma análise multivariada, utilizando as variáveis globais, descrita em detalhes no capítulo 5. Dessa análise gera-se uma variável discriminante (*MVA discriminator*) entre eventos de $Z(\nu\nu)+\text{Jatos}$ e eventos de sinal, cuja distribuição é mostrada na figura 4.30. Os eventos de $Z(\nu\nu)+\text{Jatos}$ se concentram em valores baixos dessa variável, em torno de 0,1, já os diferentes modelos de sinal apresentam comportamentos distintos dependendo do valor de Δm , modelos mais comprimidos, como o modelo com $\Delta m = 30 \text{ GeV}$, possuem uma quantidade menor de jatos reconstruídos e se assemelham mais com eventos de $Z(\nu\nu)+\text{Jatos}$, já os modelos menos comprimidos, como o modelo com $\Delta m = 100 \text{ GeV}$, são bem discriminados e se concentram em valores da variável discriminante próximos de 1. A variável discriminante também possui um grande poder de discriminação entre os eventos de $Z(\nu\nu)+\text{Jatos}$ e eventos com quarks top. Eventos de $Z(\nu\nu)+\text{Jatos}$ e $W(l\nu)+\text{Jatos}$ são extremamente similares nessa região, sendo muito difícil separar suas contribuições, de modo que não é possível definir um região de controle com uma pureza tão alta como nos outros casos. A região de controle de $Z(\nu\nu)+\text{Jatos}$ é definida com um corte superior em 0,6 na variável discriminante, escolhido por ser o valor mínimo que garantia que os eventos selecionados varressem todos os valores possíveis das distribuições das variáveis hadrônicas. A pureza dos $Z(\nu\nu)+\text{Jatos}$ na região de controle é de 66,5%, e a segunda contribuição dominante vem de eventos de $W(l\nu)+\text{Jatos}$ que contribuem com 30,4% dos eventos e apresentaram

¹³No gráfico central da figura 4.3, o intervalo $0,78 \leq \hat{\omega}_{\min} < 0,8$ contém a região em que f_i é igual a 1 para valores de $\Delta\varphi$ superiores a $\pi/2$.

um bom acordo com os dados reais, após as correções da subseção 4.6.1. A contaminação por sinal relativa ao modelo com $m_{\tilde{g}} = 1000$ GeV e $\Delta m = 30$ GeV é de 0,24% como mostrado na figura 4.31.

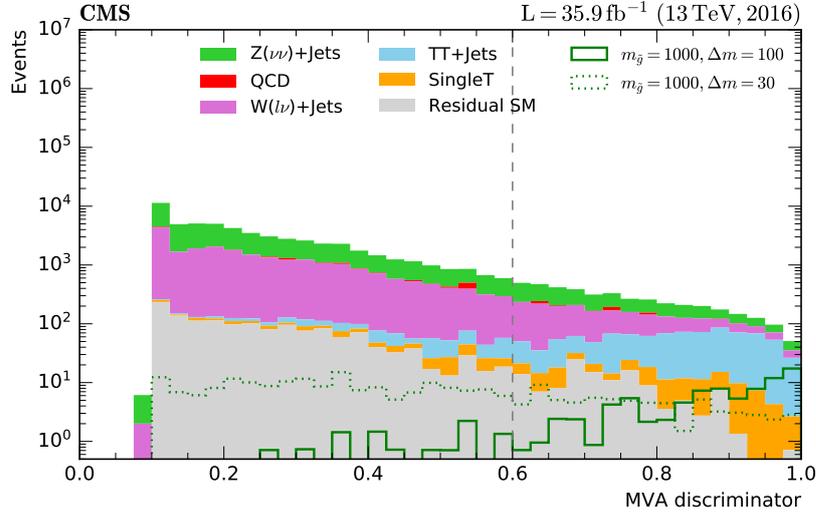


FIGURA 4.30 – Distribuição dos eventos de sinal e *background* no intervalo $0,78 \leq \hat{\omega}_{\min} < 0,8$, descrito na figura 4.22, com respeito à variável discriminante obtida por uma análise multivariada usando redes neurais, a qual será descrita em detalhes no capítulo 5.

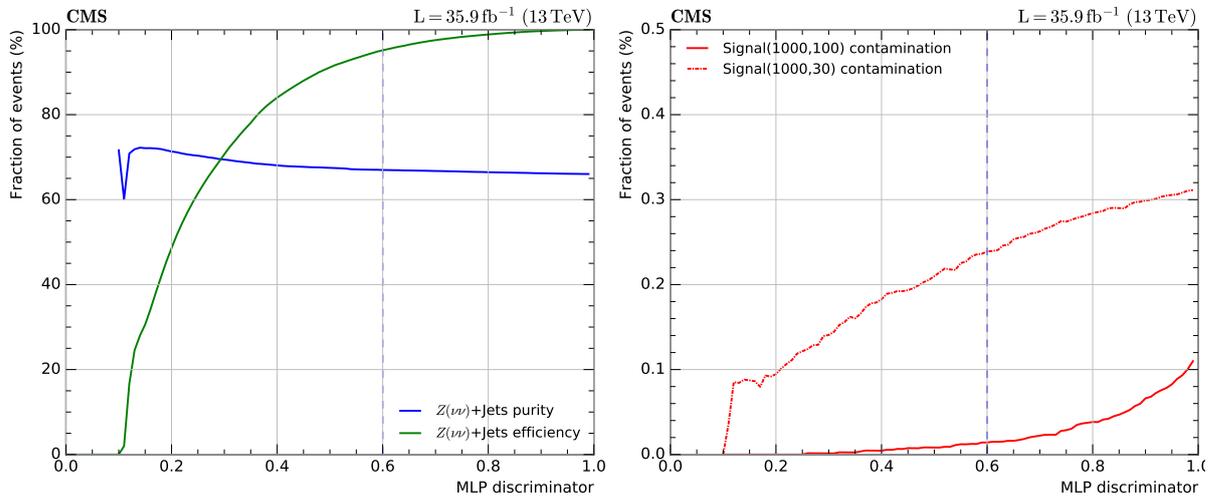


FIGURA 4.31 – Gráfico de controle da pureza e eficiência dos eventos de $Z(\nu\nu)+\text{Jatos}$ com relação aos pontos de corte na variável discriminante (esquerda). Gráfico de controle da contaminação por eventos da amostra de sinal, com $m_{\tilde{g}} = 1000$ GeV e $\Delta m = 100$ GeV, e com $m_{\tilde{g}} = 1000$ GeV e $\Delta m = 30$ GeV, com relação aos pontos de corte na variável discriminante (direita). A linha tracejada vertical denota o valor da variável discriminante em que foi aplicado o corte que define a região de controle de $Z(\nu\nu)+\text{Jatos}$.

1. **Fator de escala global:** o fator de escala global dos eventos de $Z(\nu\nu)+\text{Jatos}$ é mostrado na tabela 4.11, e é estimado como o fator multiplicativo nos pesos dos

eventos simulados que faz com que o número de eventos simulados seja igual ao número de eventos reais na região de controle.

Fator de escala	
F_Z^G	$1,227 \pm 0,011$

TABELA 4.11 – Fator de escala global para eventos de $Z(\nu\nu)+\text{Jatos}$.

2. **Correções de forma:** As correções de forma para os eventos de $Z(\nu\nu)+\text{Jatos}$ são definidas nas sub-regiões delimitadas pelos valores das variáveis E_T^{miss} e $\text{Jet}_S^{\text{PV}} p_T$, como especificado na figura 4.32.

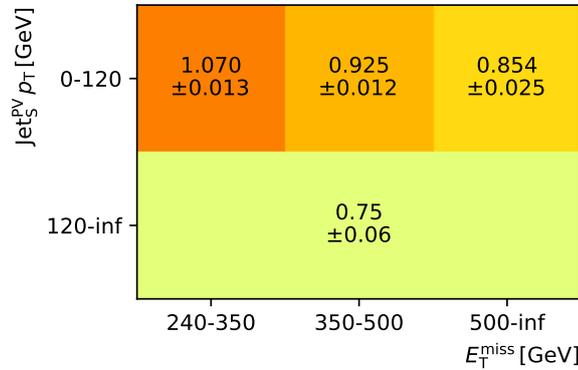


FIGURA 4.32 – Correções de forma dos eventos de $Z(\nu\nu)+\text{Jatos}$ e suas incertezas, aplicadas em sub-regiões definidas pelas variáveis E_T^{miss} e $\text{Jet}_S^{\text{PV}} p_T$.

3. **Correções de escala regionais:** A última etapa destinada às correções das escalas regionais. A tabela 4.12 mostra os resultados obtidos para os eventos de $Z(\nu\nu)+\text{Jatos}$ e suas incertezas.

Região	Fator de escala
F_Z^{R1}	$1,17 \pm 0,12$
F_Z^{R2}	$0,98 \pm 0,01$
F_Z^{R3}	$1,04 \pm 0,04$
F_Z^{R4}	$0,99 \pm 0,09$

TABELA 4.12 – Fatores de escala regionais para eventos de $Z(\nu\nu)+\text{Jatos}$.

As distribuições das variáveis globais na região de controle de $Z(\nu\nu)+\text{Jatos}$, após todas as correções, são mostradas na figura 4.33, apresentando um bom acordo entre os dados reais e a simulação.

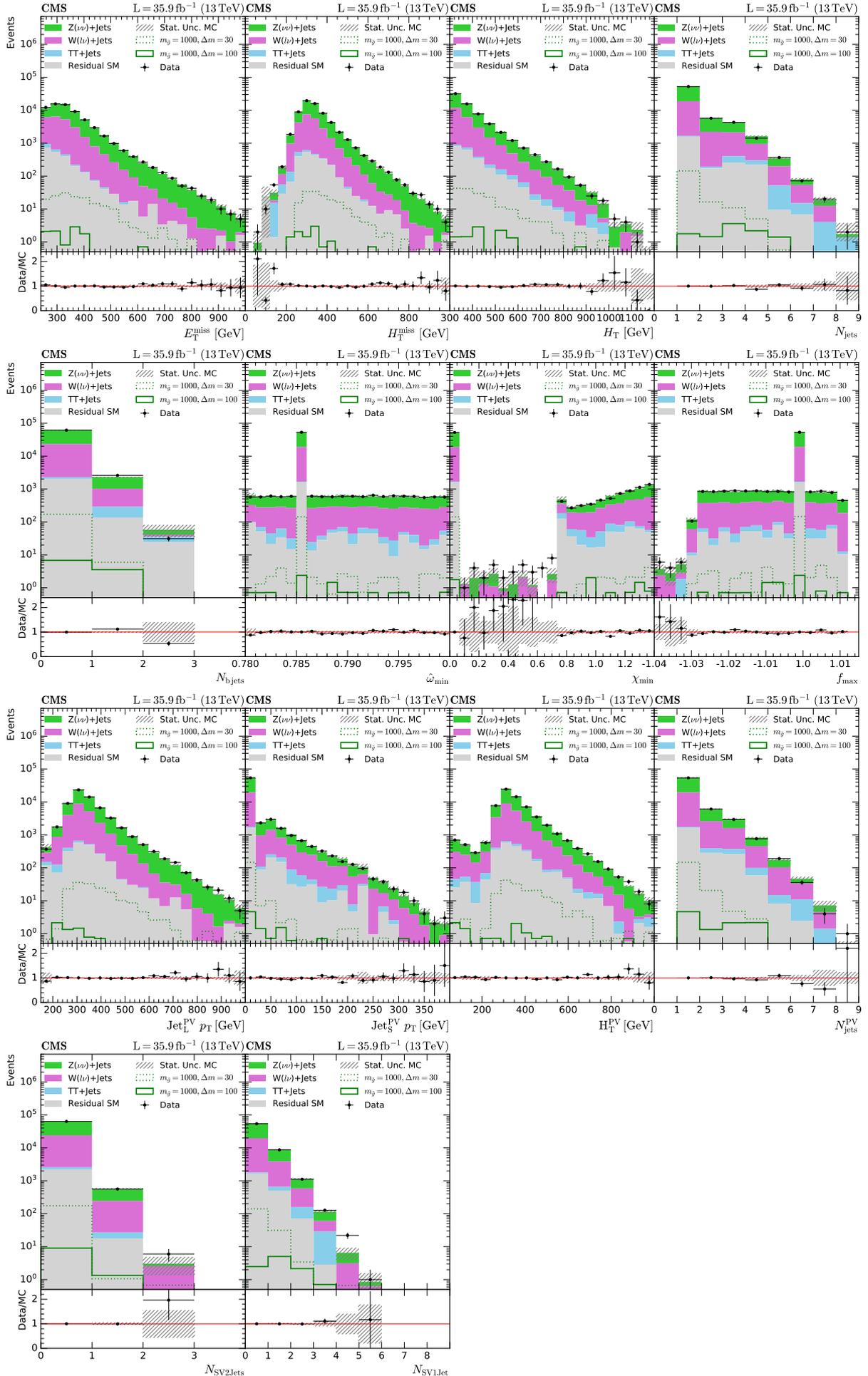


FIGURA 4.33 – Distribuições das variáveis globais na região de controle de $Z(\nu\nu)+\text{Jatos}$ após as correções da simulação.

4.7 Variáveis nas sub-regiões de busca

A região de busca, como visto na seção 4.5, é dividida em quatro sub-regiões (R1, R2, R3 e R4) que são definidas em termos do número de SV1Jet's e SV2Jets's reconstruídos no evento. A análise multivariada, realizada para definir as regiões de sinal, é feita dentro de cada uma dessas quatro sub-regiões de busca, utilizando a informação proveniente das variáveis globais e regionais. As variáveis regionais contêm a informação oriunda dos SV1Jet's e SV2Jets's reconstruídos. Portanto um conjunto distinto de variáveis regionais é definido nas regiões R1, R2 e R3, de modo que R4 não possui variáveis regionais por não possuir SV1Jet's e SV2Jets's reconstruídos. Nas subseções a seguir, as variáveis regionais de cada sub-região são definidas e são mostradas as distribuições do *background* e do sinal com respeito a todas as variáveis definidas na sub-região (globais e regionais).

4.7.1 Variáveis de R1

A região R1 possui pelo menos um SV2Jets reconstruído, no caso de haver mais de um SV2Jets, será utilizado na análise o SV2Jets com a maior significância do SV, S_{SV} , definida na equação (4.13). As variáveis a seguir são construídas com informações da reconstrução do vértice secundário e dos jatos associados a ele.

- Momento transversal do jato líder (jato com maior p_T) associado ao vértice secundário, $\text{Jet}_L^{SV} p_T$.
- Momento transversal do jato sub-líder (jato com menor p_T) associado ao vértice secundário, $\text{Jet}_S^{SV} p_T$.
- Massa invariante, M_{JJ} , dos dois jatos associados ao vértice secundário. Nos eventos de sinal, esses jatos correspondem ao quark e antiquark produzidos no decaimento do gluino, de modo que a massa invariante é limitada teoricamente pela diferença de massa entre o gluino e o neutralino LSP. A expressão da massa invariante entre os jatos líder e sub-líder denotados, respectivamente, por L e S, é dada por

$$M_{JJ} = \sqrt{m_L^2 + m_S^2 + 2(E_L E_S - \vec{p}_L \cdot \vec{p}_S)}, \quad (4.15)$$

onde m , E e \vec{p} são, respectivamente, a massa, energia e o momento do jato.

- Diferença entre a pseudorapidez dos dois jatos associados ao vértice secundário, $\Delta\eta_{JJ} = |\eta_L - \eta_S|$.
- Distância entre os dois jatos associados ao vértice secundário, no plano $(\eta - \phi)$, $\Delta R_{JJ} = \sqrt{(\eta_L - \eta_S)^2 + (\phi_L - \phi_S)^2}$.

- Valor de χ^2 do ajuste dos traços que formaram o vértice secundário, **SV χ^2** .
- Número de graus de liberdade (*number of degrees of freedom* - NDF) do ajuste dos traços que formaram o vértice secundário, **SV NDF**.

A seguir são exibidas as distribuições de sinal e *background* de todas as variáveis da sub-região de busca R1.

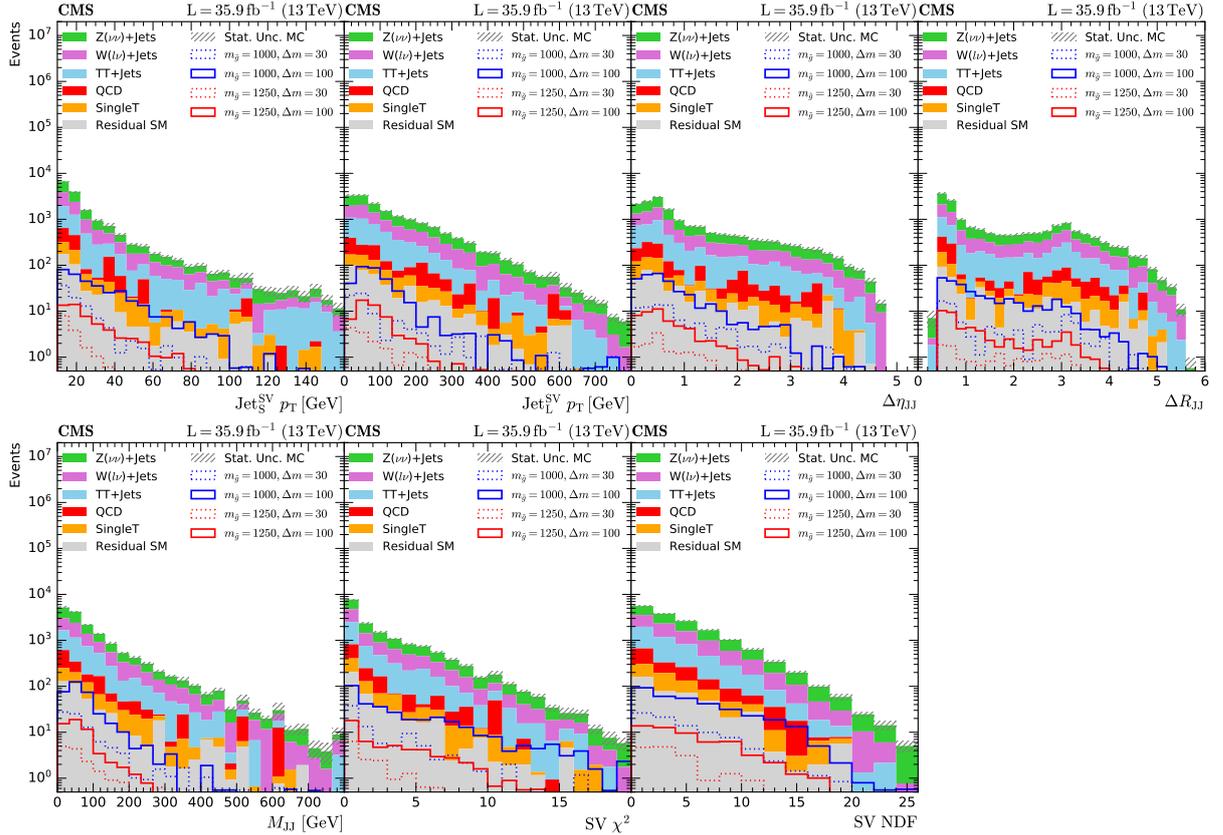


FIGURA 4.34 – Distribuições dos *backgrounds* e sinais, com relação às variáveis regionais, na sub-região de busca R1.

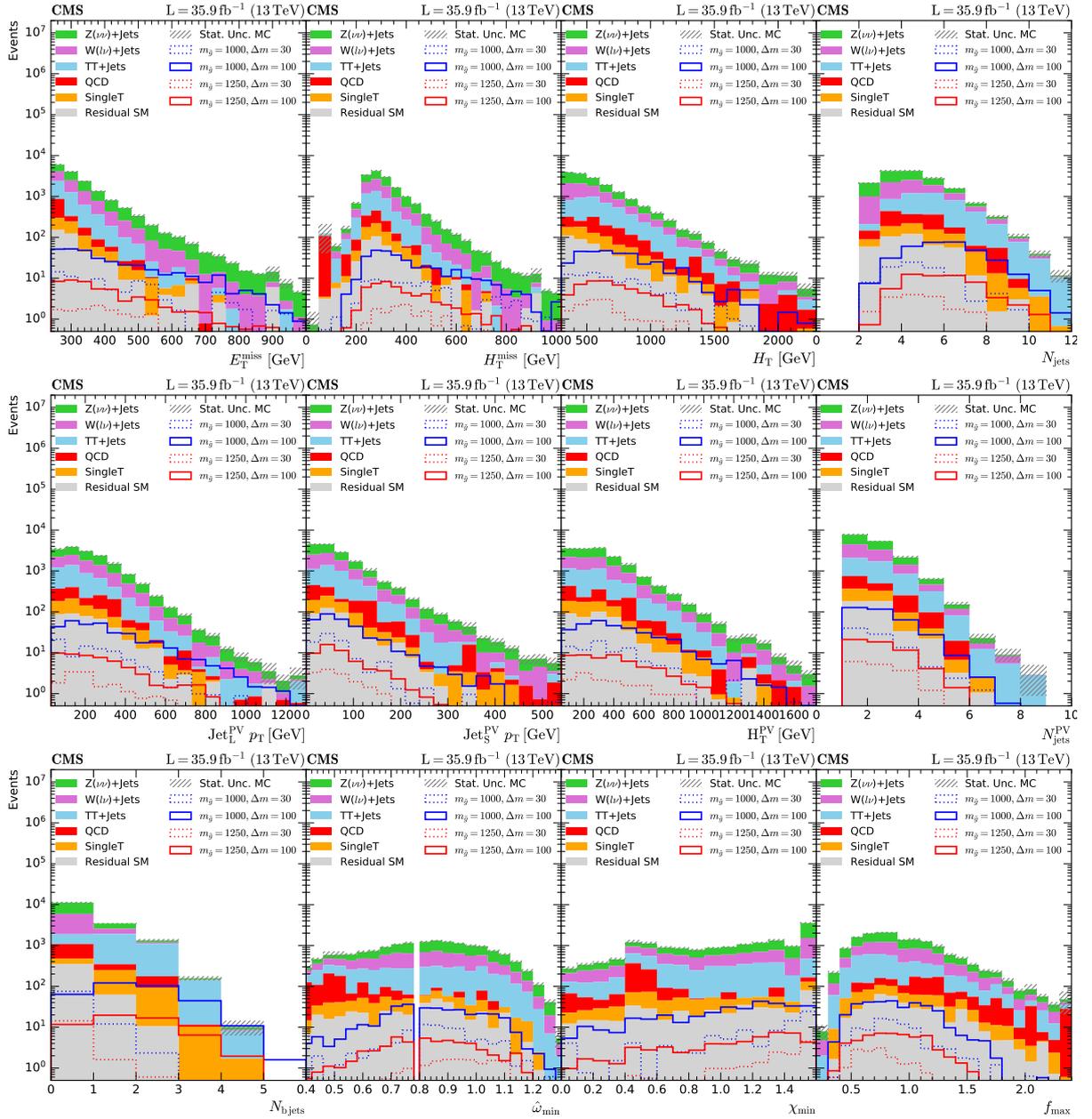


FIGURA 4.35 – Distribuições dos *backgrounds* e sinais, com relação às variáveis globais, na sub-região de busca R1.

4.7.2 Variáveis de R2

A região R2 possui pelo menos dois SV1Jet's reconstruídos, caso haja mais de dois SV1Jet's, serão utilizados na análise os dois com maior S_{SV} . Logo, são utilizados dois vértices secundários cada um associado a um jato. As variáveis a seguir são construídas com informações da reconstrução dos vértices secundários e dos jatos associados a eles.

- Momento transverso do jato correspondente ao SV1Jet com maior S_{SV} , $\text{Jet}_1^{SV} p_T$.
- Momento transverso do jato correspondente ao SV1Jet com o segundo maior S_{SV} , $\text{Jet}_2^{SV} p_T$.
- Valor de χ^2 do ajuste dos traços que formaram o vértice secundário correspondente ao SV1Jet com maior S_{SV} , $SV_1\chi^2$.
- Número de graus de liberdade do ajuste dos traços que formaram o vértice secundário correspondente ao SV1Jet com maior S_{SV} , $SV_1 \text{ NDF}$.
- Valor de χ^2 do ajuste dos traços que formaram o vértice secundário correspondente ao SV1Jet com o segundo maior S_{SV} , $SV_2\chi^2$.
- Número de graus de liberdade do ajuste dos traços que formaram o vértice secundário correspondente ao SV1Jet com o segundo maior S_{SV} , $SV_2 \text{ NDF}$.

4.7.2.1 Variáveis *razor*

As variáveis *razor* são projetadas para serem sensíveis a processos envolvendo a produção de um par de partículas pesadas, com cada uma delas decaindo em uma partícula invisível mais jatos [136,137]. Os jatos e outros produtos visíveis de cada decaimento são combinados para formar dois “megajatos” que são usados na definição das variáveis. Os megajatos dos eventos de sinal em R2 são naturalmente compostos pelos jatos que compõem os SV1Jet's reconstruídos no evento. Os jatos não reconstruídos nos decaimentos dos gluinos são oriundos de quarks pouco energéticos que não produziram grandes discrepâncias nos resultados. Portanto os dois megajatos são definidos simplesmente como Jet_1^{SV} e Jet_2^{SV} .

A variável *razor* M_R é a estimativa de uma escala de massa global, M_Δ , que no limite de megajatos sem massa é dada por $M_\Delta = (m_{\tilde{g}}^2 - m_{\tilde{\chi}_1}^2)/m_{\tilde{g}}$. A variável M_R é definida em termos dos momentos dos megajatos segundo a expressão a seguir,

$$M_R \equiv \sqrt{(|\vec{p}^{j1}| + |\vec{p}^{j2}|)^2 - (p_z^{j1} + p_z^{j2})^2}, \quad (4.16)$$

com j_1 e j_2 denotando, respectivamente, os jatos Jet_1^{SV} e Jet_2^{SV} , \vec{p} sendo o momento dos jatos e p_z sua componente no direção do eixo do feixe. A variável *razor* M_T^R é construída inteiramente de informação do plano transverso (p_T dos megajatos e E_T^{miss}) através da expressão abaixo,

$$M_T^R \equiv \sqrt{\frac{E_T^{\text{miss}}(p_T^{j_1} + p_T^{j_2}) - \vec{E}_T^{\text{miss}} \cdot (\vec{p}_T^{j_1} + \vec{p}_T^{j_2})}{2}}. \quad (4.17)$$

Com ela, pode-se definir a variável *razor*, R^2 , cuja definição é

$$R \equiv \frac{M_T^R}{M_R}. \quad (4.18)$$

A razão R quantifica o fluxo de energia no plano perpendicular ao eixo do feixe. Essas variáveis possuem maior poder de discriminação com relação aos eventos do MP em modelos não comprimidos, mas ainda podem contribuir na análise multivariada em conjunto com as outras variáveis.

A seguir são exibidas as distribuições de sinal e *background* de todas as variáveis da sub-região de busca R2.

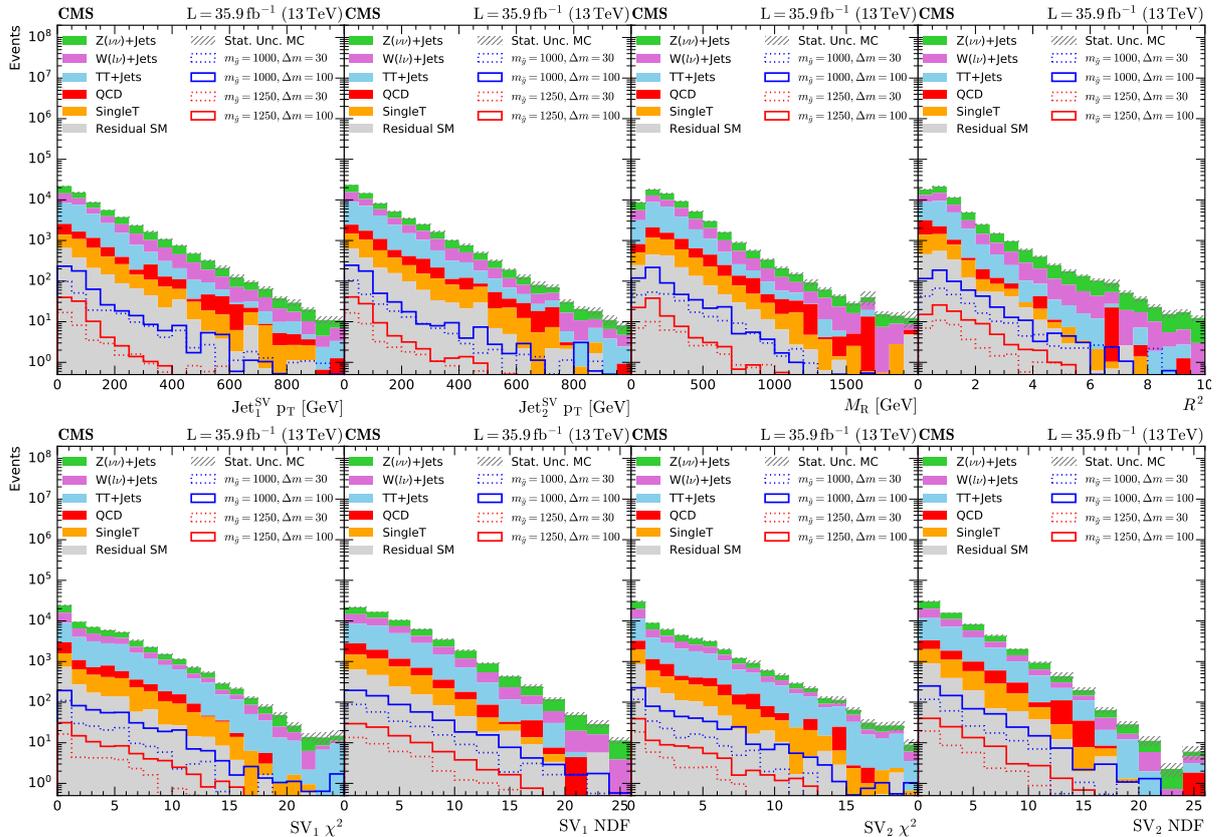


FIGURA 4.36 – Distribuições dos *backgrounds* e sinais, com relação às variáveis regionais, na sub-região de busca R2.

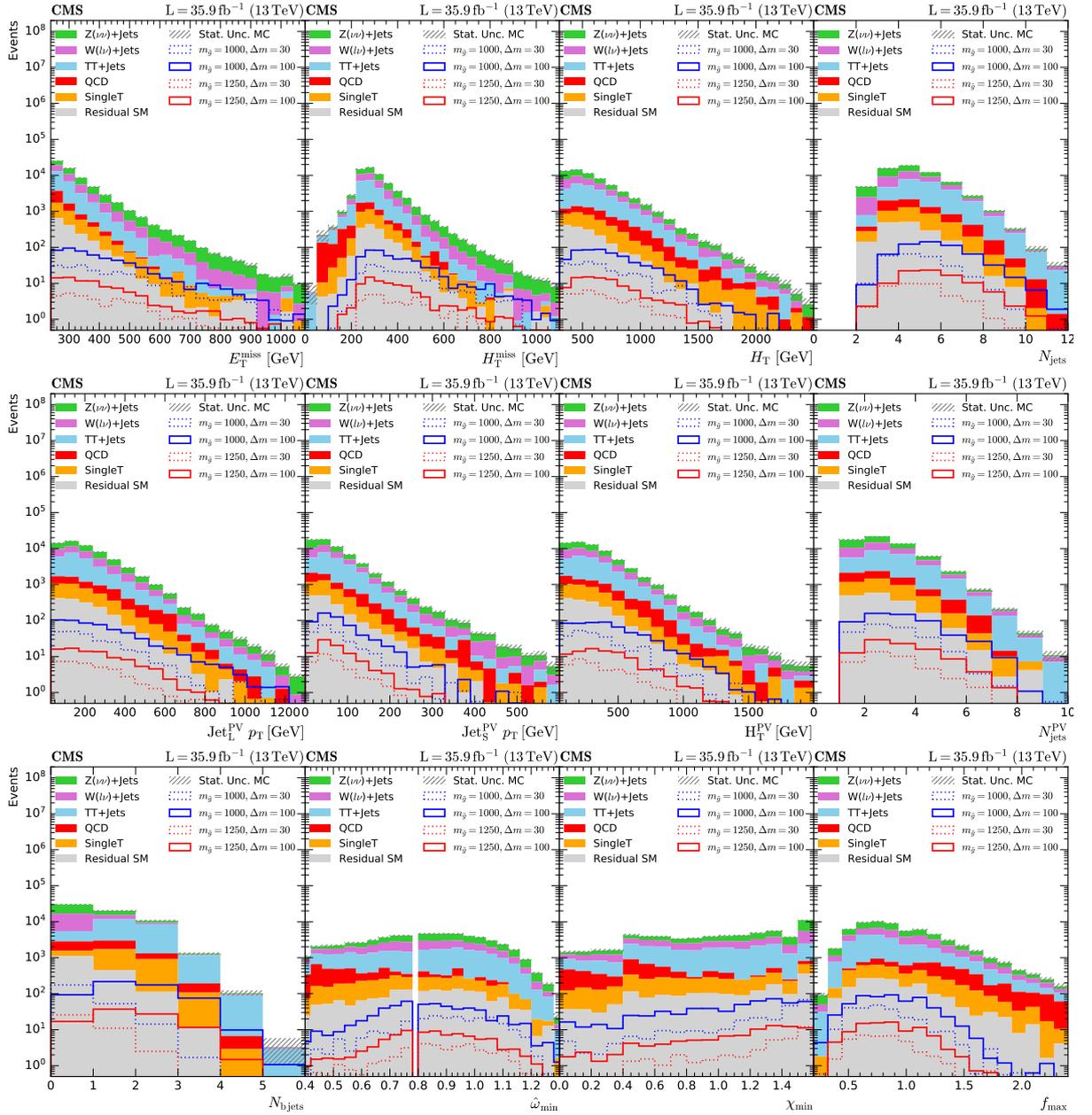


FIGURA 4.37 – Distribuições dos *backgrounds* e sinais, com relação às variáveis globais, na sub-região de busca R2.

4.7.3 Variáveis de R3

A região R3 possui apenas um SV1Jet reconstruído, o que limita a quantidade de informação que pode ser obtida. As três variáveis a seguir são construídas com informações da reconstrução do vértice secundário e do jato associado a ele.

- Momento transverso do jato correspondente ao SV1Jet, $Jet^{SV} p_T$.
- Valor de χ^2 do ajuste dos traços que formaram o vértice secundário, $SV\chi^2$.
- Número de graus de liberdade do ajuste dos traços que formaram o vértice secundário, $SV NDF$.

A seguir são exibidas as distribuições de sinal e *background* de todas as variáveis da sub-região de busca R3.

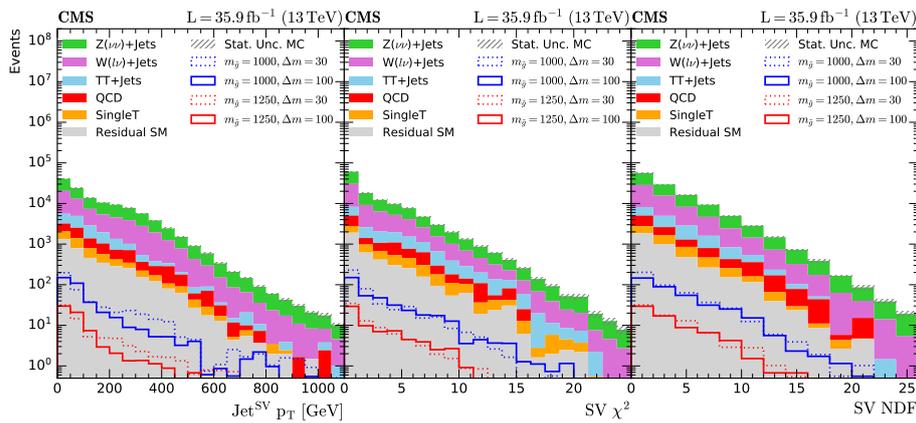


FIGURA 4.38 – Distribuições dos *backgrounds* e sinais, com relação às variáveis regionais, na sub-região de busca R3.

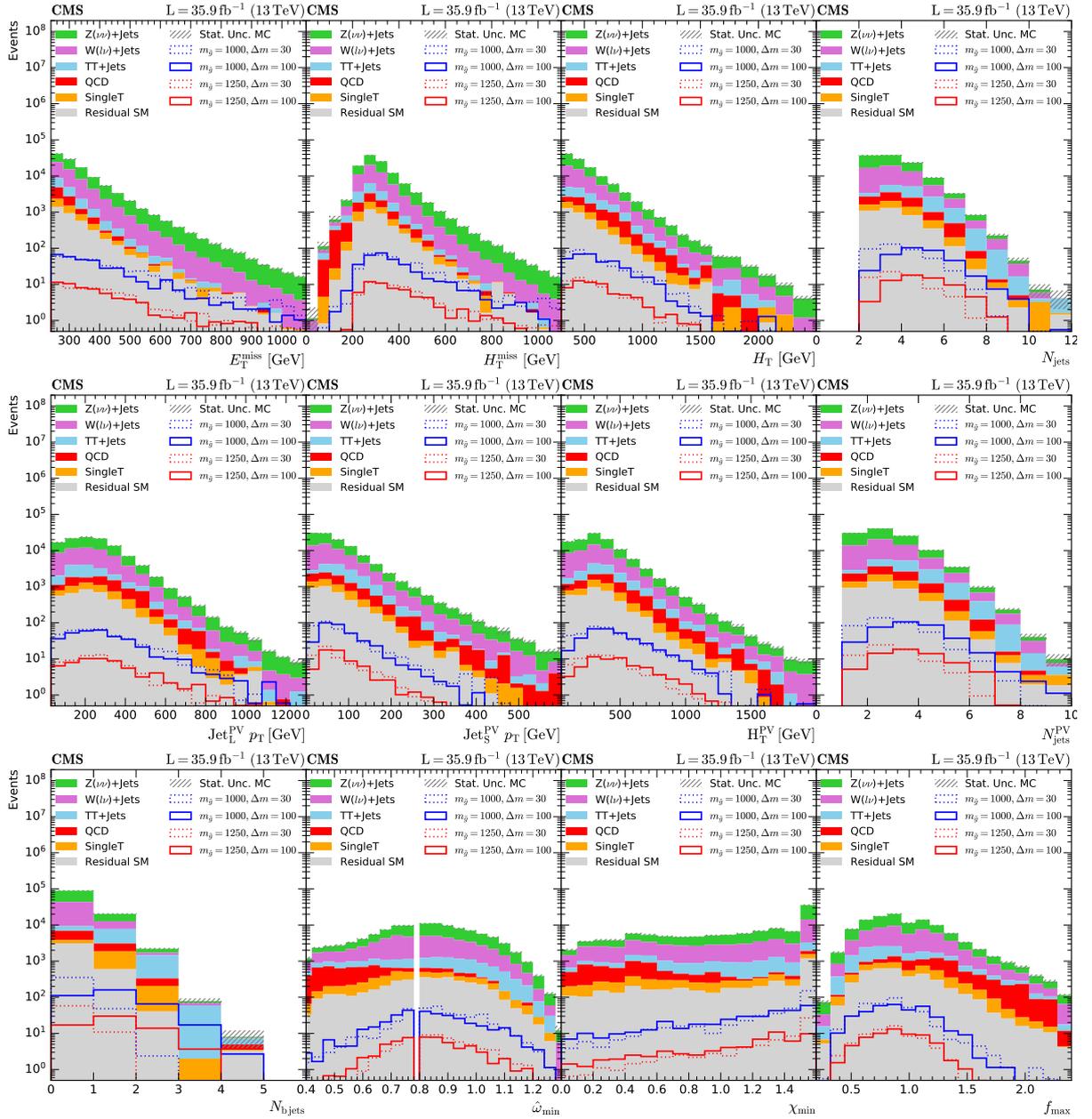


FIGURA 4.39 – Distribuições dos *backgrounds* e sinais, com relação às variáveis globais, na sub-região de busca R3.

4.7.4 Variáveis de R4

Como mencionado anteriormente, a sub-região R4 não possui variáveis regionais. As distribuições de sinal e *background* de todas as variáveis de R4 são exibidas a seguir.

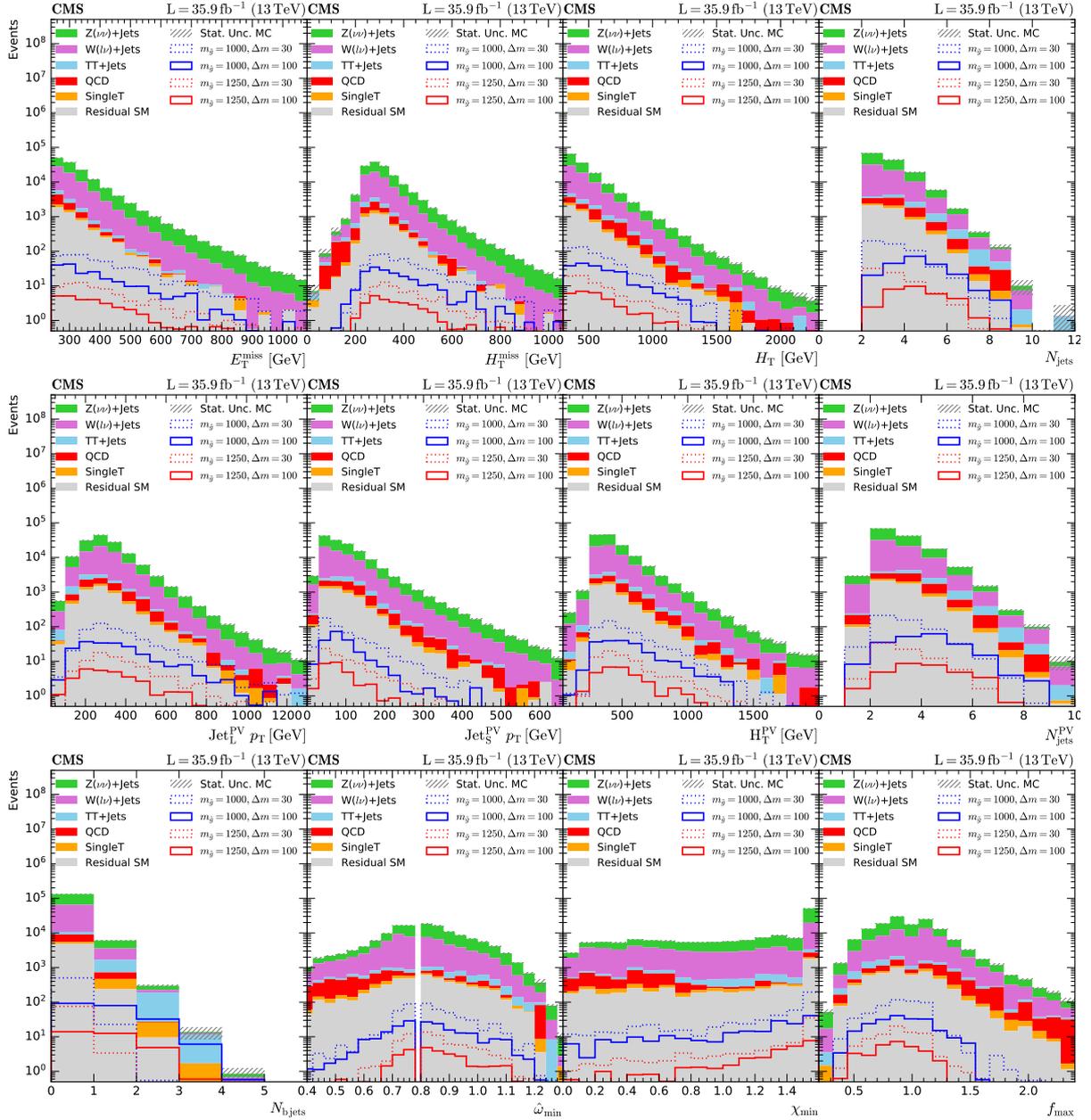


FIGURA 4.40 – Distribuições dos *backgrounds* e sinais, com relação às variáveis globais, na sub-região de busca R4.

4.7.5 Validação das Variáveis

As variáveis utilizadas na busca são validadas em três regiões distintas, subdivididas seguindo o critério exposto na figura 4.19, onde é observado se as distribuições das simulações de MC com respeito às variáveis de busca estão em acordo com os dados. As três regiões de validação são descritas nas figuras 4.41, 4.42 e 4.43, considerando que os eventos tenham passado pela seleção inicial. Os gráficos de todas as variáveis de busca utilizadas em cada sub-região são mostrados no Apêndice C.

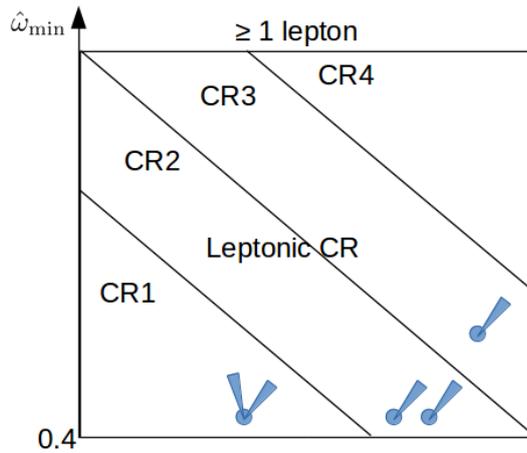


FIGURA 4.41 – Região Leptônica: possui pelo menos um múon ou elétron reconstruído com $p_T \geq 10 \text{ GeV}$ e o valor de $\hat{\omega}_{\min}$ é maior que 0,4.

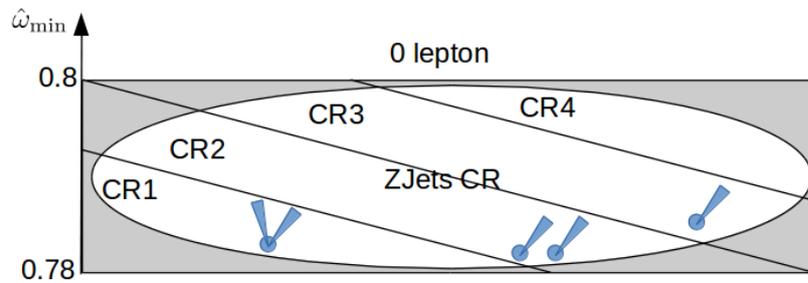


FIGURA 4.42 – Região de $Z(\nu\nu)+\text{Jatos}$: não possui múons ou elétrons reconstruídos, o valor de $\hat{\omega}_{\min}$ se encontra no intervalo $0,78 \leq \hat{\omega}_{\min} < 0,8$ e o valor do discriminante da análise multivariada é menor que 0,6.

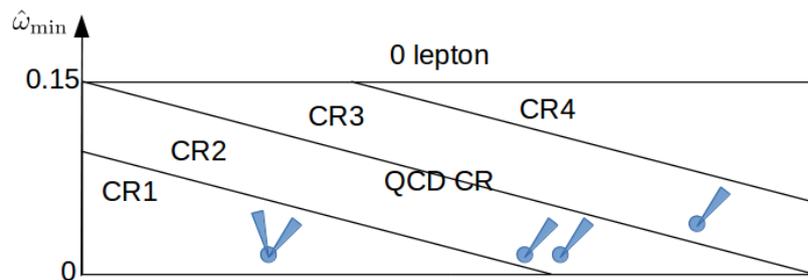


FIGURA 4.43 – Região de QCD: não possui múons ou elétrons reconstruídos e o valor de $\hat{\omega}_{\min}$ é menor que 0,15.

5 Análise Multivariada

Cada vez mais, estudos realizados tanto em áreas científicas quanto no mundo dos negócios envolvem uma enorme quantidade de dados e têm gerado demandas por métodos e ferramentas para interpretar e analisar dados de alta dimensionalidade, em que cada conjunto de dados contém um grande número de medições e variáveis. Existem várias abordagens amplamente utilizadas para a análise multivariada (*multivariate analysis* - MVA) de dados, incluindo ajuste de modelos lineares e não lineares para associar uma variável à outra, ou a um conjunto de outras variáveis, e análise de correlação para capturar como as variáveis interagem ou influenciam umas as outras [138,139]. A prática usual para analisar dados multivariados inclui várias etapas como pré-processamento, a garantia de qualidade, redução de dimensionalidade, teste de hipóteses, modelos preditivos, análise de correlação, modelagem gráfica, visualização e interpretação dos resultados.

A implementação de uma análise multivariada é o ponto central desse projeto, pois possibilita uma melhora significativa na sensibilidade da análise em regiões críticas do *grid* de simulação dos sinais, correspondentes a modelos mais comprimidos e com gluinos possuindo grandes massas. Particularmente, são utilizadas técnicas de aprendizado de máquina (*machine learning*) baseadas em estudos científicos de sistemas computacionais em que algoritmos e modelos estatísticos são treinados para executar uma tarefa específica de maneira eficaz, sem usar instruções explícitas, confiando em padrões e inferência. O termo *machine learning* foi cunhado em 1959 por Arthur Samuel [140] e é visto como um subconjunto da inteligência artificial (IA), que é a parte da ciência da computação preocupada em projetar sistemas computacionais inteligentes, isto é, sistemas que exibem características que associamos à inteligência, no comportamento humano.

As técnicas implementadas são as Redes Neurais Artificiais (*Artificial Neural Networks* - ANN) [141] e as Árvores de Decisão Reforçadas (*Boosted Decision Trees* - BDT) [142], escolhidas devido ao seu reconhecido mérito e recorrente uso em análises de física de partículas de altas energias. Este capítulo começa fazendo uma descrição das técnicas utilizadas na análise. Posteriormente, descreve-se a análise multivariada realizada na definição da região de controle dos $Z(\nu\nu)$ +Jatos e na definição das regiões de sinal.

5.1 Aprendizado de Máquina

A análise multivariada baseada no aprendizado de máquina é um método de análise de dados que automatiza a construção de modelos analíticos, fazendo uso simultâneo de toda informação possível sobre o sistema em questão para produzir os resultados desejados. O caso de interesse neste projeto é a classificação de padrões, ou seja, usar informações dos diferentes tipos de processos físicos, armazenados nas amostras de Monte Carlo, para encontrar regiões no espaço das variáveis, associadas a cada um deles.

A figura 5.1 mostra o exemplo de um sistema que possui duas classes diferentes de eventos (A e B), em que as informações conhecidas sobre cada classe estão armazenadas em duas variáveis (x_1 e x_2)¹. Na parte de cima, são observados os dados simulados, em que se sabe a qual classe pertence cada evento. O objetivo da análise multivariada, em problemas de classificação, é encontrar a curva de separação que melhor discrimina as duas classes de eventos, a partir das informações contidas nas variáveis do sistema. Na parte de baixo, são expostos os dados reais do sistema, em que não se conhece, a priori, a classe de cada evento. Contudo, utilizando a curva de separação obtida na análise multivariada (em verde), é possível classificar os eventos reais baseando-se na região do espaço de variáveis em que ele se encontra. No exemplo da figura 5.1, os seis eventos dos dados reais que possuem valores positivos nas variáveis x_1 e x_2 , são classificados como eventos da classe A, enquanto os outros são classificados como eventos da classe B.

A análise multivariada, usando técnicas de aprendizagem de máquina, é realizada através do processo descrito na figura 5.2. Os dados com informações sobre as classes dos eventos² são divididos entre a amostra de treinamento, amostra de validação e amostra de teste. O algoritmo de aprendizagem é executado sobre a amostra de treinamento. O algoritmo possui um conjunto de parâmetros que são atualizados à medida que o algoritmo processa os dados repetidas vezes, cada conjunto de valores atribuídos aos parâmetros do algoritmo define um modelo de classificação. Portanto o modelo modifica-se à medida que esse processo de repetição (conhecido como processo de treinamento) ocorre. Após cada etapa de atualização do modelo, durante o processo de treinamento, uma cópia dele processa os dados presentes na amostra de validação. Com isso, os resultados processados pelo modelo, relativos às amostras de treinamento e validação, são avaliados e comparados, etapa conhecida como validação cruzada (*Cross Validation*). O objetivo final do processo de treinamento é gerar um modelo que classifique da forma mais eficiente possível os dados simulados e que a performance do modelo na amostra de treinamento seja

¹Este é tipicamente o tipo de problema abordado nesta análise, porém as informações dos eventos estarão armazenadas em um número maior de variáveis.

²Dados com informações sobre as classes dos eventos, em física de partículas, são os dados simulados, porém em outras áreas pode-se fazer uso de dados históricos, que consiste de dados reais obtidos anteriormente, em que é conhecida a classe de cada evento.

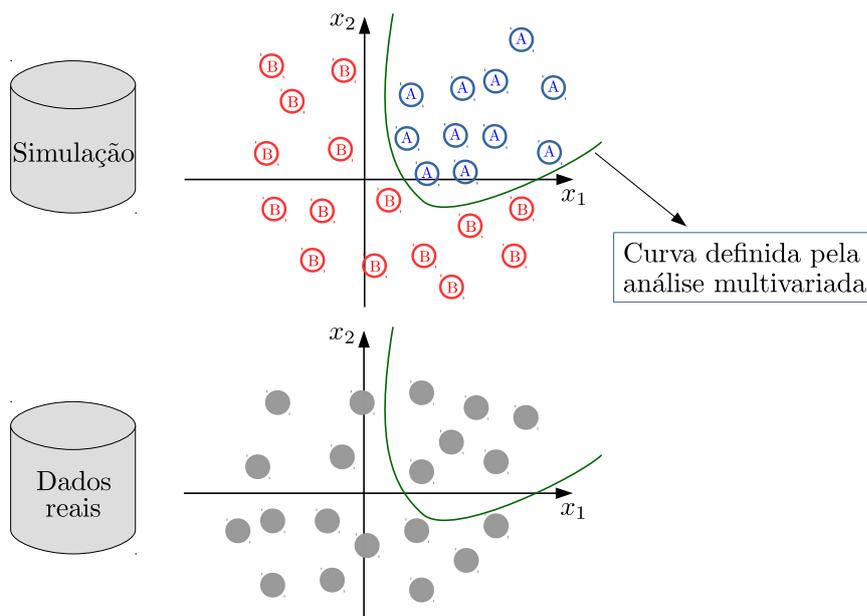


FIGURA 5.1 – Exemplo de um sistema que possui duas classes de eventos (A e B), em que as informações conhecidas sobre cada classe estão armazenadas nas variáveis x_1 e x_2 . A curva de separação (em verde) é obtida da análise multivariada realizada nos dados simulados e posteriormente é aplicada nos dados reais para realizar a classificação dos eventos.

estatisticamente similar à performance sobre a amostra de validação. Esse último requerimento é incluído para evitar que o modelo descreva as flutuações estatísticas particulares da amostra de treinamento (situação conhecida pelo termo em inglês, *overtraining*), que não são características gerais do sistema, como exemplificado na figura 5.3, produzindo uma diminuição de sua performance na amostra de validação e também nos dados reais. O modelo otimizado, caso não apresente *overtraining*³, é aplicado sobre a amostra de teste, e o resultado de sua performance é analisado. Se considerado bom o suficiente ou sendo considerado o melhor entre diversos modelos, o modelo vencedor é aplicado sobre os dados reais para realizar a classificação.

Os modelos construídos no processo de aprendizagem de máquina, são funções das variáveis de entrada, que resultam em uma ou mais saídas de números reais que são processadas para realizar a classificação dos eventos. No exemplo da figura 5.1, só existem duas classes de eventos (A e B), logo, a saída do modelo é um número real. A variável real que carrega os valores das saídas do modelo com respeito aos eventos de entrada é chamada de “discriminante” do modelo. A distribuição do discriminante entre duas classes de eventos, com respeito aos eventos da amostra de treinamento, teria a forma apresentada na figura 5.4. A partir dessa distribuição, define-se o melhor valor do discriminante para realizar-se o corte que define a curva de separação entre as duas classes de eventos no

³Na prática, alguns pacotes computacionais incluem opções para controle do *overtraining*.

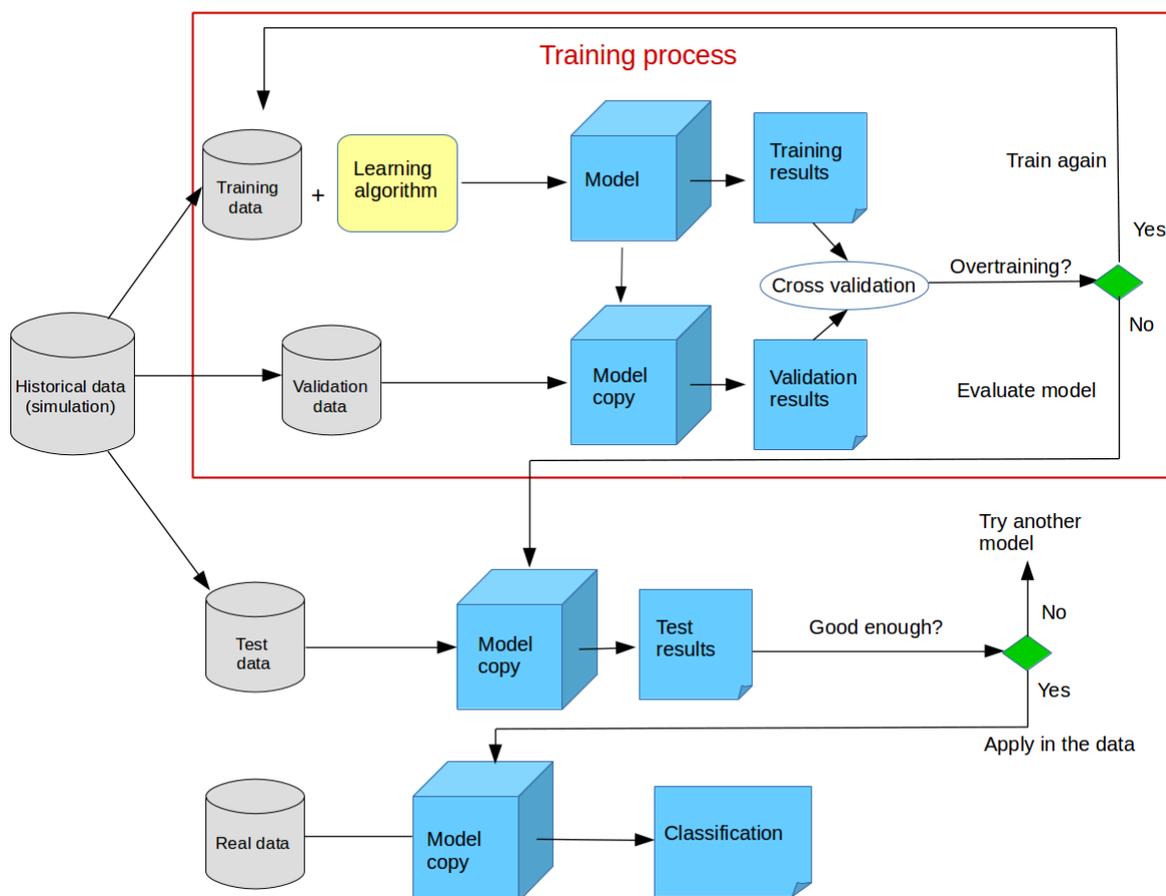


FIGURA 5.2 – Processo completo do treinamento e seleção de modelos em análises multivariadas usando técnicas de aprendizagem de máquina.

espaço das variáveis.

5.2 Árvores de Decisão

Árvore de decisão é um método de classificação que implica em uma sequência de cortes nas variáveis dos eventos, seguindo uma regra preestabelecida. Esses cortes são realizados recursivamente até que seja satisfeito certo critério de encerramento, gerando regiões no espaço das variáveis que serão designadas como regiões da classe de sinal ou da classe de *background*, dependendo da predominância do tipo de eventos dentro dela.

O algoritmo de construção de uma árvore de decisão começa encontrando o corte a ser aplicado na totalidade de eventos do sistema que produza a melhor separação entre sinal e background, o conjunto total de eventos é chamado de nóculo raiz. Entre os critérios usados para a seleção do melhor corte, um dos mais utilizados é o do índice Gini, definido

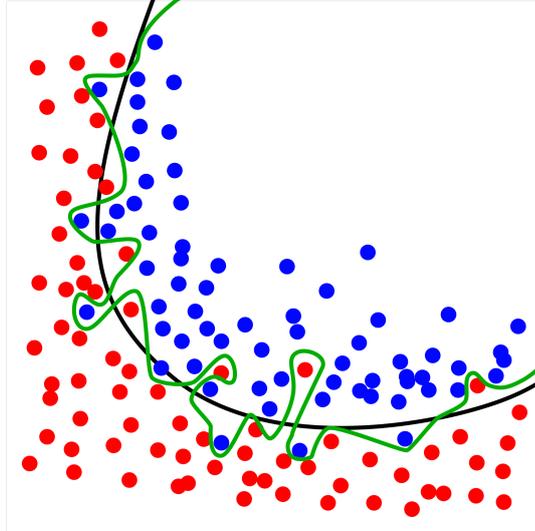


FIGURA 5.3 – Exemplo de uma situação de *overtraining* em que os pontos azuis e vermelhos representam duas classes diferentes de eventos na amostra de treinamento. A curva preta representa a curva de separação ideal, que considera somente as propriedades gerais do sistema, já a curva verde é resultado de um treinamento excessivo em que o modelo considera na classificação as flutuações estatísticas da amostra de treinamento, que não serão observadas nos dados reais.

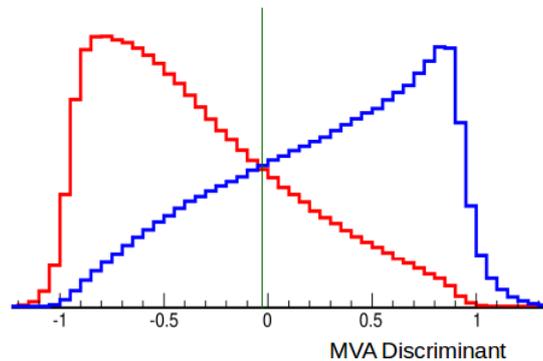


FIGURA 5.4 – Exemplo da distribuição do discriminante de um modelo de classificação entre duas classes de eventos, com respeito aos eventos da amostra de treinamento. As diferentes classes estão representadas pelas cores vermelha e azul. A reta vertical em um valor ligeiramente menor que zero, representa o ponto de corte escolhido para separar as duas classes de eventos do problema.

por

$$\text{Gini} \equiv \left(\sum_{i=1}^N w_i \right) P(1 - P), \quad \text{tal que} \quad P \equiv \frac{\sum_s w_s}{\sum_s w_s + \sum_b w_b}, \quad (5.1)$$

sendo N o número de eventos dentro do nóculo, w_i denota o peso de cada evento, P é denominada a pureza da amostra no nóculo, $\sum_s w_s$ é a soma dos pesos dos eventos de sinal e $\sum_b w_b$ é a soma dos de *background*, presentes no nóculo. Assim, a regra escolhida seleciona o corte que maximiza a diferença entre o índice de Gini do nóculo atual e a soma

do índice de Gini dos nós resultantes. A árvore de decisão é construída realizando um corte no nó raiz e nos nós resultantes sucessivamente até que sejam satisfeitas condições de encerramento que evitam a produção de nós resultantes com pureza inferior ao nó de origem ou até ser atingida profundidade máxima estipulada para a árvore, o nó raiz possui profundidade igual a 1.

Um exemplo de árvore de decisão pode ser visto na figura 5.5, onde temos um sistema formado por 500 eventos de sinal e 800 eventos de background, classificados utilizando-se três variáveis (x_1 , x_2 e x_3). Focando no nó 2, observa-se que esse nó possui 400 eventos de sinal e 200 de *background* e o corte escolhido foi aquele que maximizou a expressão

$$\text{Gini}_2 - (\text{Gini}_4 + \text{Gini}_5). \quad (5.2)$$

A árvore de decisão do exemplo formou 5 regiões no espaço das variáveis, 3 de sinal e 2 de background. Dessa forma, quando aplicada aos dados reais, os eventos são classificados dependendo de qual dessas regiões eles se encontram.

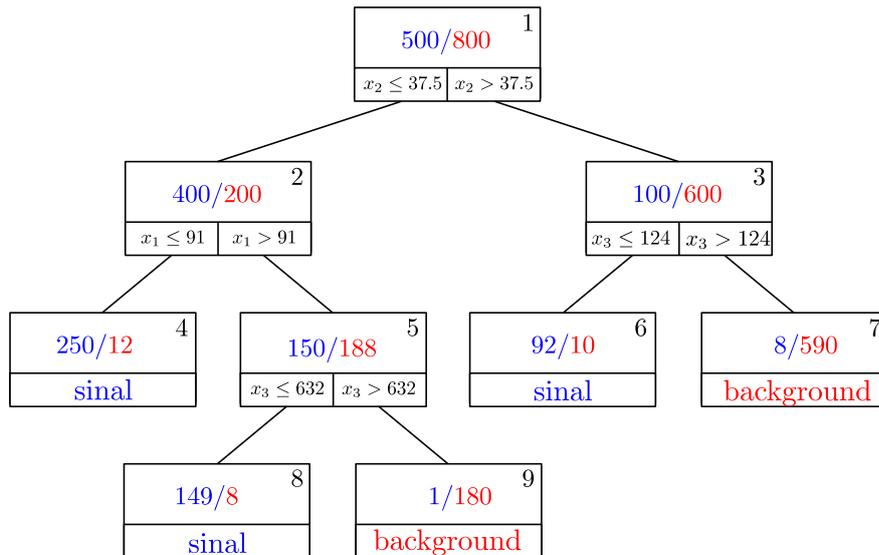


FIGURA 5.5 – Exemplo de uma árvore de decisão aplicada sobre um sistema formado por duas classes de eventos, com 500 eventos na classe de sinal (azul) e 800 eventos na classe de *background* (vermelho). A árvore possui 9 nós, formados com cortes em três variáveis diferentes (x_1 , x_2 e x_3), e gera 3 regiões de sinal e duas regiões de *background* no espaço das variáveis.

O uso de uma árvore de decisão na classificação de eventos fica muito dependente da amostra utilizada na sua construção, tendo limitações na generalização de seu resultado. Por esse motivo, o método de árvores de decisão é geralmente utilizado conjuntamente com algoritmos que visam melhorar sua performance, os algoritmos de *boosting*. Eles funcionam produzindo várias árvores de decisão recursivamente, utilizando os resultados da árvore anterior para alterar o peso dos eventos utilizados na árvore seguinte. O resultado final

é expresso através de uma função, $T(\mathbf{x})$, chamada de classificador, que é uma espécie de média ponderada do resultado de cada árvore de decisão criada.

Considera-se o caso em que são criadas M árvores de decisão, onde identifica-se cada árvore através do índice m . Para descrever o funcionamento de um algoritmo de *boosting*, faz-se as seguintes definições:

- \mathbf{x}_i - conjunto de variáveis correspondentes ao evento i ;
- y_i - indica o tipo de evento na simulação, é igual a 1 se for um evento de sinal e -1 se for de background;
- w_i^m - peso do evento i na árvore m ;
- $T^m(\mathbf{x}_i)$ - indica como o evento i foi classificado na árvore m , é igual a 1 se foi classificado como sinal e -1 se foi classificado como background;
- $I^m(\mathbf{x}_i)$ - é igual a 1 caso o evento i tenha sido classificado incorretamente ($y_i \neq T^m(\mathbf{x}_i)$) na árvore m e 0 se foi classificado corretamente ($y_i = T^m(\mathbf{x}_i)$);
- β - taxa de aprendizagem do algoritmo de *boosting*.

Um algoritmo de boosting bastante utilizado é o AdaBoost (*Adaptive Boosting*). Nele, define-se o parâmetro, a^m , dado por

$$a^m \equiv \beta \cdot \ln \left(\frac{1 - \text{err}^m}{\text{err}^m} \right), \quad \text{com} \quad \text{err}^m \equiv \frac{\sum_{i=1}^N w_i^m I^m(\mathbf{x}_i)}{\sum_{i=1}^N w_i^m}. \quad (5.3)$$

O parâmetro a^m é utilizado para modificar os valores dos pesos que serão usados pela árvore seguinte através da expressão

$$w_i^{m+1} = w_i^m \cdot \exp(a^m I^m(\mathbf{x}_i)). \quad (5.4)$$

A função $T(\mathbf{x}_i)$, que é a variável discriminante do modelo, é definida por,

$$T(\mathbf{x}_i) = \frac{1}{M} \sum_{m=1}^M a^m T^m(\mathbf{x}_i). \quad (5.5)$$

A técnica de aprendizado de máquina baseada em árvores de decisão, é mais fácil de ser implementada, e mais rápida de ser treinada que a técnica baseada em redes neurais, introduzida na seção seguinte. Contudo, os resultados obtidos com redes neurais são geralmente melhores, como no caso dessa análise. O poder de discriminação das melhores árvores de decisão será comparado com os resultados obtidos para as melhores redes neurais, através da comparação das curvas ROC dos modelos. As árvores de decisão são computadas utilizando o pacote computacional TMVA [142].

5.3 Redes Neurais Artificiais

A neurociência é a área que estuda o sistema nervoso biológico, visando desvendar seu funcionamento, estrutura, desenvolvimento e eventuais alterações. Atualmente trata-se de uma área multidisciplinar, sendo a neurocomputação uma de suas vertentes que lida com o desenvolvimento de modelos computacionais baseados em algoritmos estruturados, inspirados no sistema nervoso biológico, chamados de Redes Neurais Artificiais (RNA). Entre algumas definições de redes neurais temos:

“Sistemas de processamento de informação projetados com inspiração tomada do sistema nervoso, ou mais especificamente, do cérebro, e com particular ênfase na solução de problemas.” (Castro [145])

“Uma rede neural (artificial) é um processador maciçamente paralelamente distribuído, constituído de unidades de processamento simples, que têm uma propensão natural para armazenar conhecimento experimental e torná-lo disponível para o uso.” (Haykin [146])

A primeira publicação relacionada à neurocomputação data de 1943, por meio do artigo de McCulloch e Pitts [147], no qual os autores realizam o primeiro modelamento matemático inspirado no neurônio biológico, resultando na primeira concepção de neurônio artificial. Em 1949, foi proposto o primeiro método de treinamento de redes neurais artificiais, baseado em hipóteses e observações de caráter neurofisiológico, que se denominou de regra de aprendizado de Hebb [148]. Entre outros trabalhos realizados no desenvolvimento de modelos matemáticos fundamentados no neurônio biológico, destaca-se o trabalho de Frank Rosenblatt que, entre 1957 e 1958, desenvolveu o primeiro neurocomputador, denominado Mark I - Perceptron, idealizando o modelo básico do Perceptron [149] que despertou interesse de outros pesquisadores pela sua capacidade em reconhecer padrões simples. Em 1960, foi desenvolvida a rede Adaline (*ADaptive LINear Element*), e em seguida a Madalaine, Adaline múltipla, cujo o aprendizado é baseado na regra Delta, também conhecida como algoritmo de aprendizado LMS (*Least Mean Square*).

Devido à impossibilidade do Perceptron e Adaline em realizar a correta classificação de padrões para classes não linearmente separáveis, não houve muitas pesquisas na área nos anos subsequentes. A retomada de um maior interesse em pesquisas na área ocorre no final dos anos 80, impulsionada por diversos fatores, tais como o desenvolvimento de computadores com maior capacidade de processamento, a criação de algoritmos de otimização mais eficientes e novas descobertas sobre o sistema nervoso biológico. Um dos principais trabalhos desse período foi a publicação do livro *Parallel distributed processing* em 1986 [150], onde foi desenvolvido um algoritmo que permitia ajustar os pesos em uma

rede com mais de uma camada. O algoritmo, denominado *backpropagation*, era capaz de classificar corretamente padrões não linearmente separáveis, e motivou novas pesquisas em redes neurais artificiais.

O termo “aprendizagem profunda” (*deep learning*) refere-se a uma classe de algoritmos que usa múltiplas camadas, redes neurais profundas (*deep neural networks* - DNNs), para extrair progressivamente características de alto nível de dados brutos [151]. O impacto da aprendizagem profunda na indústria começou no início dos anos 2000, quando as redes neurais convolucionais⁴ (*convolutional neural networks* - CNNs) já processavam cerca de 10% a 20% de todos os cheques escritos nos EUA. Aplicações industriais de aprendizagem profunda para reconhecimento de voz em larga escala começaram por volta de 2010. Houve um grande avanço no 2012, que foi publicado por um trabalho da Imagenet [152] desencadeando o renascimento atual em redes neurais. Foram reunidas maneiras diferentes de acelerar o processo de aprendizagem, incluindo o uso inteligente de GPUs e alguns truques matemáticos novos como o uso de funções de ativação ReLU (*REctified Linear Unit*), e mostrou-se que em poucas semanas eles poderiam treinar uma rede muito complexa a um nível que poderia superar as abordagens convencionais para visão computacional⁵.

As redes neurais são empregadas em diferentes áreas como engenharia, ciências e negócios. Entre as inúmeras potenciais áreas de aplicação, podemos citar algumas onde as redes neurais já são bastante usadas [141]:

- **Aproximador universal de funções:** tem como objetivo mapear o relacionamento funcional entre variáveis (tipicamente reais) de um sistema a partir de uma amostra de seus valores representativos;
- **Controle de processos:** tem como objetivo identificar ações de controle para que atinjam requisitos de qualidade, eficiência e segurança. É geralmente usado em robótica, aeronaves, elevadores, satélites, etc;
- **Reconhecimento/classificação de padrões:** visa a associação de cada padrão de entrada a uma das classes previamente definidas. É geralmente usado em reconhecimento de imagens, voz, escrita, etc.
- **Agrupamento de dados (clusterização):** visa identificar e detectar similaridades e particularidades entre diversos padrões de entrada, a fim de agrupá-los em diferentes classes;

⁴Rede neural convolucional é uma classe de DNNs mais comumente aplicada na análise de imagens virtuais.

⁵Visão computacional é um campo científico dedicado a estudar como computadores podem ser desenvolvidos para obter um alto grau de conhecimento de imagens digitais e vídeos.

- **Sistemas de previsão:** têm como objetivo estimar valores futuros de um processo levando-se em consideração diversas medidas prévias observadas em seu domínio. É usado na previsão de séries temporais, previsões de mercados financeiros, previsões climáticas, etc;
- **Memórias associativas:** O objetivo consiste em recuperar padrões corretos mesmo se os seus elementos constituintes forem apresentados de forma incerta ou imprecisa. É utilizado em processamento de imagens, transmissão de sinais, identificação de caracteres manuscritos, etc.

5.3.1 Neurônio Biológico e Artificial

A célula elementar do sistema nervoso cerebral é o neurônio, cujo papel é conduzir impulsos elétricos, oriundos de reações físico-químicas, sob determinadas condições de operação. O neurônio biológico pode ser dividido em três partes principais: nos dendritos, no corpo celular e no axônio, como mostrado na figura 5.6.

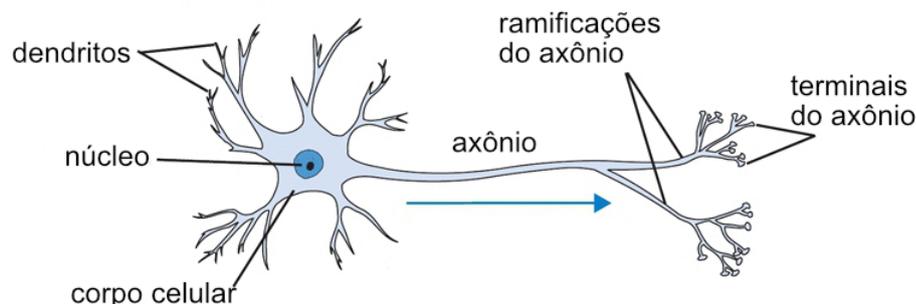


FIGURA 5.6 – Representação de um neurônio biológico [153].

Os dendritos são formados por vários prolongamentos finos ligados ao corpo celular, cuja a principal função consiste de captar os estímulos vindos de outros neurônios ou do próprio meio externo, onde os mesmos podem estar em contato. O corpo celular é incumbido de processar todas as informações oriundas dos dendritos, a fim de produzir um potencial de ativação que indicará se o neurônio poderá disparar um impulso elétrico ao longo de seu axônio. O axônio, que é constituído por um único prolongamento, conduz os impulsos elétricos para outros neurônios conectores ou para aqueles que se conectam diretamente com o tecido muscular. A sua terminação é também constituída de ramificações denominadas terminações do axônio ou terminações sinápticas. Não existe contato físico entre os neurônios, desse modo, a transmissão de impulsos elétricos de um neurônio para o outro é feita e ponderada por elementos neurotransmissores dependentes da química cerebral [154]. O potencial elétrico de uma membrana neural, quando em repouso (polarizada), assume valor negativo, ou seja, existe uma concentração maior de íons negativos

internamente à membrana em relação ao seu exterior. No momento em que o neurônio é estimulado (despolarizado) além de um limiar de ativação (em torno de -55 mV), ocorre o disparo de um impulso elétrico prolongado ao longo de todo o seu axônio [155].

Em analogia ao que acontece nos neurônios biológicos, o neurônio artificial recebe um sinal de entrada representado pelo conjunto de variáveis (que caracterizam um objeto, uma imagem, um evento físico, etc.), $\{x_1, \dots, x_n\}$, e esse sinal é ponderado pelos pesos sinápticos, $\{w_1, \dots, w_n\}$, formando o potencial de ativação $I = \sum_{i=0}^n w_i \cdot x_i = \mathbf{w} \cdot \mathbf{x}$, onde $x_0 \equiv -1$ e $w_0 \equiv \theta$, sendo θ o limiar de ativação do neurônio artificial. Por fim, o neurônio produz um sinal de saída, $Y(\mathbf{x})$, que é o resultado do processamento da informação de entrada pela chamada função de ativação, $g(I)$. O neurônio artificial em termos matemáticos é definido por um conjunto de pesos sinápticos e pela função de ativação que produzirá o resultado da saída do neurônio. Na figura 5.7 pode-se ver uma representação esquemática de um neurônio artificial. Os elementos dentro do contorno vermelho constituem o neurônio artificial, eles processam os n sinais de entrada e fornecem o sinal de saída denotado por Y .

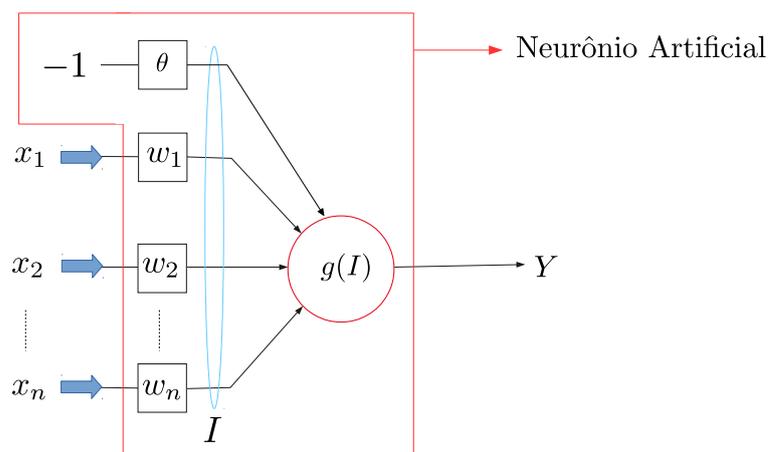


FIGURA 5.7 – Representação esquemática de um Neurônio Artificial.

5.3.2 Rede Perceptron

O Perceptron foi idealizado em 1958 por Rosenblatt [149], ele é a forma mais simples de uma rede neural artificial, sendo formado por um único neurônio. Ele foi concebido como um modelo computacional inspirado na retina para agir como um elemento de percepção eletrônica de sinais. A figura 5.7 é uma representação esquemática de uma rede Perceptron. A função de ativação da rede Perceptron é geralmente a função degrau ou a função sinal, logo, somente duas possibilidades de valores são produzidos pela saída.

Considerando uma rede Perceptron com apenas dois sinais de entrada e a função de ativação como sendo a função sinal, o seu funcionamento é descrito na figura 5.8. Ao

receber os sinais de entrada, $\{x_1, x_2\}$, o potencial ativador é computado, resultando na expressão,

$$I = w_1 \cdot x_1 + w_2 \cdot x_2 - \theta. \quad (5.6)$$

Caso I seja um valor positivo, a saída da rede será 1, e caso I seja zero ou um valor negativo, a saída será -1 . Os valores 1 e -1 representam, respectivamente, as saídas da rede associadas as classes predefinidas, A e B.

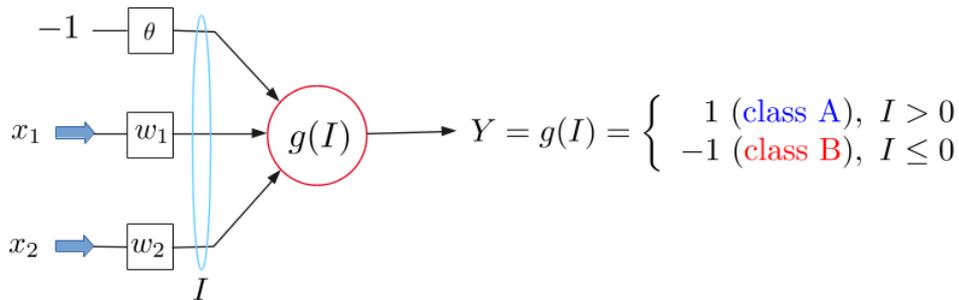
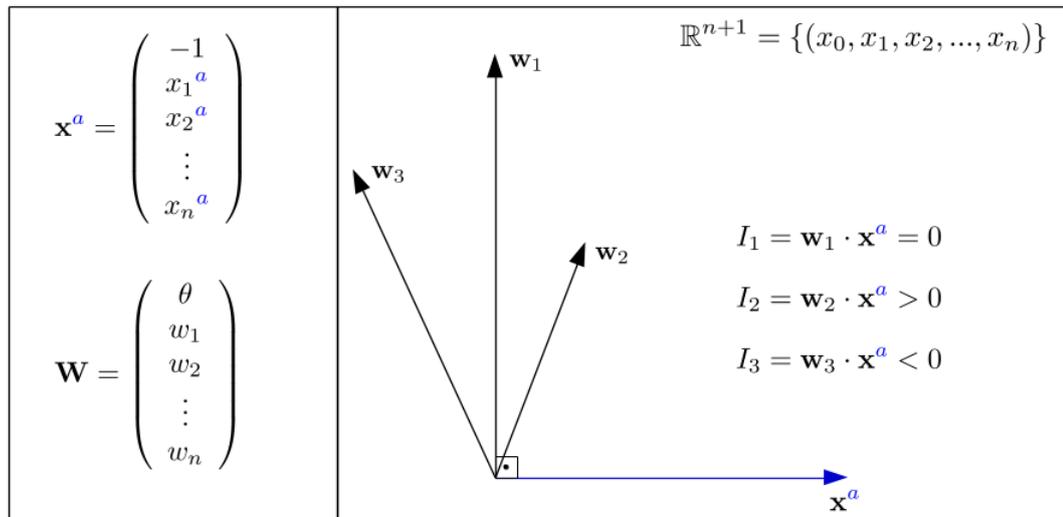


FIGURA 5.8 – Exemplo do funcionamento de uma rede Perceptron que recebe entradas compostas por duas variáveis $\{x_1, x_2\}$.

O funcionamento da rede Perceptron depende de uma prévia definição dos valores dos parâmetros w_1 , w_2 e θ . De um modo geral, os pesos sinápticos e limiares de ativação de redes neurais são definidos previamente no processo de treinamento da rede. Nesse processo, é submetida à rede uma amostra de treinamento possuindo m entradas com n variáveis. A classificação correta das entradas na amostra de treinamento já é previamente conhecida e denotada por d . O processo de treinamento visa comparar a saída da rede, Y , como o resultado esperado, d , de modo a alterar os valores dos parâmetros da rede até que a rede gere os valores corretos de Y . Isso é alcançado fazendo-se uso de um algoritmo de aprendizagem projetado para realizar o correto ajuste dos parâmetros da rede.

Para entender os algoritmos de aprendizagem descritos a seguir, é muito útil observar o comportamento dos vetores \mathbf{x}^a e \mathbf{w} no espaço \mathbb{R}^{n+1} constituído pelo conjunto de todas as $(n + 1)$ -tuplas ordenadas, ou seja, $\mathbb{R}^{n+1} = \{(x_0, x_1, x_2, \dots, x_n); x_i \in \mathbb{R}\}$, em que o índice $a = 1, \dots, m$ representa cada entrada da amostra de treinamento. Na figura 5.9 são mostrados os três resultados possíveis do produto escalar entre uma entrada genérica da amostra de treinamento, \mathbf{x}^a , e o vetor \mathbf{w} que representa os pesos da rede. Os resultados dependem da configuração vigente dos pesos da rede, que na figura 5.9 é denotada pelos índices 1, 2 e 3 em \mathbf{w} e I . O caso 1 representa a situação em que o produto escalar é igual a zero, pode-se observar que isso ocorre porque os vetores \mathbf{x}^a e \mathbf{w}_1 são ortogonais, já no caso 2 o produto escalar é positivo devido ao ângulo entre os vetores \mathbf{x}^a e \mathbf{w}_2 ser agudo, e por fim no caso 3 o produto escalar é negativo porque o ângulo entre \mathbf{x}^a e \mathbf{w}_3 é obtuso.

O algoritmo de aprendizagem da rede Perceptron baseia na regra de aprendizado de

FIGURA 5.9 – Três casos possíveis do produto escalar $\mathbf{w} \cdot \mathbf{x}^a$.

Hebb [148], representada pela expressão

$$\mathbf{w}^{\text{nov}} = \mathbf{w}^{\text{atual}} + \eta(d^a - Y^a)\mathbf{x}^a. \quad (5.7)$$

O parâmetro η é denominado taxa de aprendizagem e exprime o quão rápido o processo de treinamento da rede é conduzido até a sua convergência. O valor de η está no intervalo compreendido em $0 < \eta < 1$, sendo escolhido um valor suficientemente pequeno para evitar instabilidades no processo de treinamento. Segundo a equação (5.7), ao comparar-se a saída esperada para a entrada a , d^a , com a saída obtida pela rede Perceptron, Y^a , caso a rede tenha classificado a entrada corretamente, a correção realizada no vetor de pesos atual, $\mathbf{w}^{\text{atual}}$, será igual a zero, logo, o novo vetor de pesos, \mathbf{w}^{nov} , será igual ao seu antecessor. Entretanto, caso a classificação esteja errada, existem duas possibilidades matemáticas distintas.

- A primeira possibilidade acontece quando $d^a = 1$ e $Y^a = -1$, sabe-se que $Y^a = -1$ ocorre quando o produto escalar $\mathbf{w}^{\text{atual}} \cdot \mathbf{x}^a$ é negativo, o que implica em um ângulo obtuso entre os vetores $\mathbf{w}^{\text{atual}}$ e \mathbf{x}^a . A correção aplicada ao vetor de pesos atual, segundo a regra de Hebb, é dada por $+2\eta\mathbf{x}^a$ que somando-se ao vetor $\mathbf{w}^{\text{atual}}$ resulta no vetor \mathbf{w}^{nov} , figura 5.10 à esquerda. O vetor \mathbf{w}^{nov} possui um ângulo menor, com relação a \mathbf{x}^a , do que o vetor $\mathbf{w}^{\text{atual}}$, quando esse ângulo for menor que 90° , a rede classificará a entrada \mathbf{x}^a corretamente.
- A segunda possibilidade acontece quando $d^a = -1$ e $Y^a = 1$, assim o produto escalar $\mathbf{w}^{\text{atual}} \cdot \mathbf{x}^a$ é positivo, o que implica em um ângulo agudo entre os vetores $\mathbf{w}^{\text{atual}}$ e \mathbf{x}^a . A correção ao vetor de pesos atual é dada por $-2\eta\mathbf{x}^a$ que somando-se ao vetor $\mathbf{w}^{\text{atual}}$ resulta no vetor \mathbf{w}^{nov} , figura 5.10 à direita. O vetor \mathbf{w}^{nov} possui um ângulo

maior, com relação a \mathbf{x}^a do que o vetor $\mathbf{w}^{\text{atual}}$, quando esse ângulo for maior que 90° , a rede classificará a entrada \mathbf{x}^a corretamente.

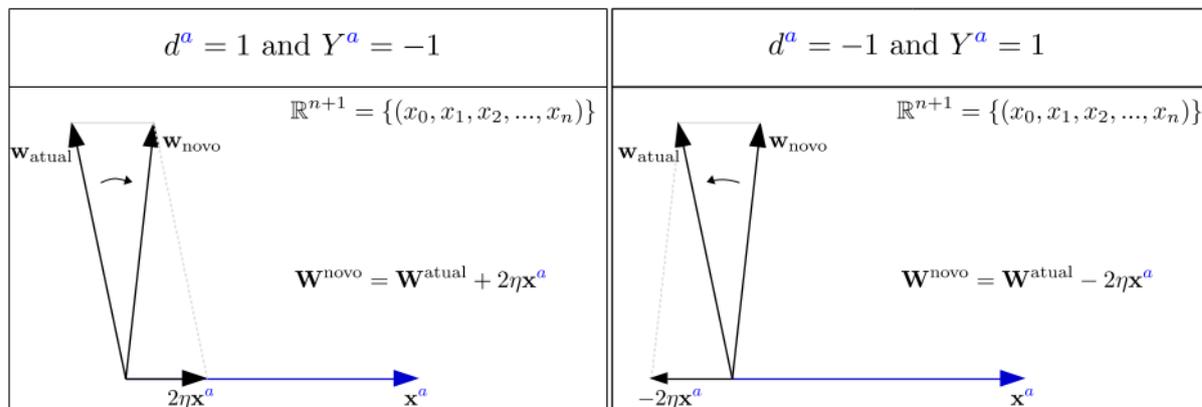


FIGURA 5.10 – Os dois casos possíveis em que a saída da rede Perceptron não é igual à saída esperada.

O processo de treinamento inicia a rede neural com um vetor de pesos, \mathbf{w} , com valores aleatórios no intervalo $0 < w_i < 1$. Então, são submetidas à rede, uma por uma, todas as m entradas da amostra de treinamento. Para cada entrada, é aplicada a regra de Hebb e corrige-se o vetor de pesos se a classificação estiver incorreta. Após serem submetidas todas as m entradas, caso alguma das entradas da amostra de treinamento tiver sido classificada incorretamente, submete-se todas as entradas da amostra novamente. O número de vezes em que é realizado esse *loop* sobre as entradas da amostra de treinamento é chamado de “número de épocas”. O treinamento da rede Perceptron termina quando, em uma época, todas as entradas tiverem sido classificadas corretamente. Logo, a rede Perceptron só é aplicável aos casos em que as variáveis de entrada, $\{x_1, \dots, x_n\}$, são linearmente separáveis. No caso de duas variáveis apenas, isso significa que as classes podem ser separadas por uma linha reta no espaço das variáveis.

5.3.3 Rede Adaline

A rede Adaline é formada por um único neurônio, igualmente à rede Perceptron, ela também produz uma separação linear entre as classe mas possui a vantagem de poder ser aplicada a sistemas que não são perfeitamente separáveis por um hiperplano. Idealizada por Widrow e Hoff em 1960 [156], sua principal aplicação se destinava a sistemas de chaveamento de circuitos telefônicos, sendo uma das primeiras aplicações industriais de RNAs. Com respeito à estrutura e ao funcionamento da rede Adaline, ela é igual a rede Perceptron, ou seja, as figuras 5.7 e 5.8 também são válidas para a rede Adaline. A diferença do rede Adaline com respeito a rede Perceptron é o algoritmo de aprendizagem.

A rede Adaline usa o algoritmo de aprendizagem denominado regra Delta, considerado a grande contribuição da rede Adaline por ser o precursor do algoritmo de treinamento usado para treinar redes Perceptron Multicamadas.

O idéia básica da regra Delta consiste em ajustar os pesos do neurônio, \mathbf{w} , minimizando a diferença entre a saída esperada, d , e o potencial de ativação, I , das entradas da amostra de treinamento. Para isso, define-se a função erro quadrático, $E(\mathbf{w})$, como sendo,

$$E(\mathbf{w}) \equiv \frac{1}{2} \sum_{a=1}^m (d^a - I^a)^2. \quad (5.8)$$

Considera-se que a amostra de treinamento possua m entradas. Então, propõe-se que a correção, $\Delta\mathbf{w}$, aplicada ao vetor, \mathbf{w} , seja proporcional ao gradiente da função erro quadrático com relação a \mathbf{w} ,

$$\Delta\mathbf{w} \equiv -\eta \cdot \nabla E(\mathbf{w}) = \eta \sum_{a=1}^m (d^a - I^a) \mathbf{x}^a, \quad (5.9)$$

sendo η a taxa de aprendizagem. O sinal negativo significa que a adaptação dos pesos deve ser efetuada na direção oposta àquela do gradiente, pois o objetivo é minimizar o erro quadrático. Por simplicidade, a atualização de \mathbf{w} pode ser realizada discretamente após a inserção de cada entrada a da amostra de treinamento,

$$\mathbf{w}^{\text{novo}} = \mathbf{w}^{\text{atual}} + \eta (d^a - I^a) \mathbf{x}^a. \quad (5.10)$$

O processo de treinamento da rede Adaline segue os mesmo passos do treinamento da rede Perceptron, porém a regra de Hebb é substituída pela regra Delta da equação (5.10), encerrando o treinamento quando é atingido o vetor de pesos ótimo, \mathbf{w}^* , no qual a função erro quadrático é mínima. Ele é atingido dentro de uma certa precisão que depende do parâmetro de aprendizagem. Um valor muito grande de η faz com que a rede convirja mais rapidamente, mas ficará oscilando em torno do ponto de mínimo, \mathbf{w}^* , com uma precisão pequena, já um valor muito pequeno de η fará com que a rede demore mais para convergir, mas o ponto de mínimo será alcançado com uma precisão maior, visto que as oscilações em torno dele terão menor amplitude. A escolha do η apropriado depende do problema atacado. Na prática, o treinamento é geralmente encerrado quando a função erro quadrático aparenta convergir para um valor mínimo, como mostrado na figura 5.11 que descreve o comportamento da função erro quadrático ao longo das épocas enquanto está sendo treinada para encontrar o vetor de pesos ótimo.

Como foi mencionado, a rede Adaline pode ser aplicada a problemas nos quais as classes não podem ser separadas perfeitamente por um hiperplano. Desse modo, o processo de treinamento se incube de encontrar o melhor hiperplano para separar as classes, ou

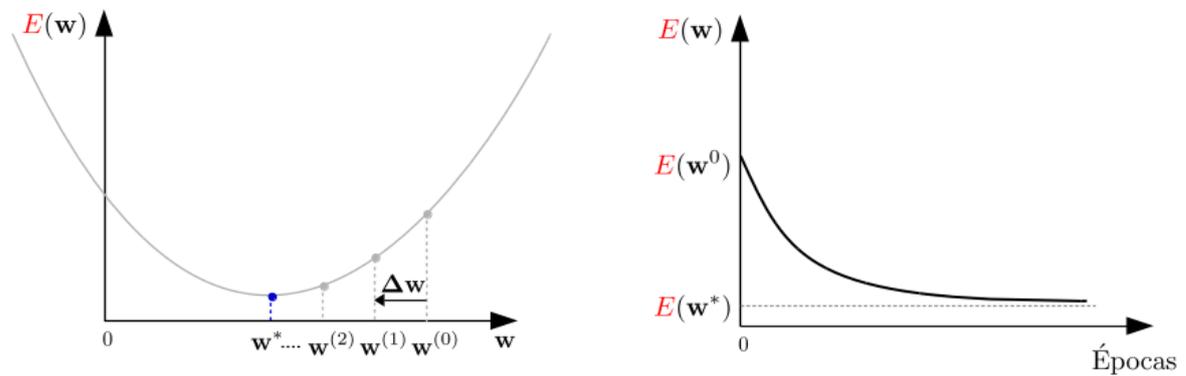


FIGURA 5.11 – Evolução da função erro quadrático com relação ao vetor de pesos (esquerda) e com relação ao número de épocas (direita).

seja, com o menor erro de classificação. A figura 5.12 mostra o hiperplano ótimo para um exemplo com duas variáveis de entrada.

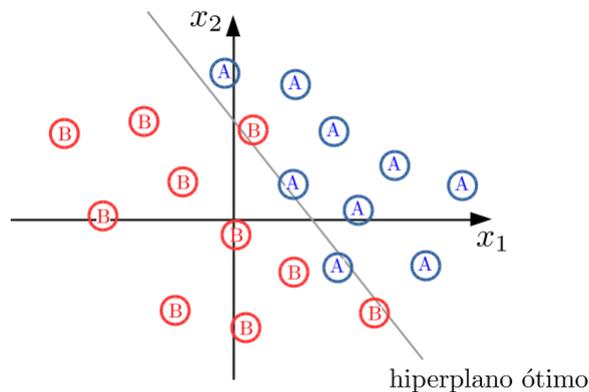


FIGURA 5.12 – Exemplo de um sistema com duas classes de eventos, A e B, cuja a informação está contida em um conjunto de duas variáveis, x_1 e x_2 . A reta diagonal representa o hiperplano de separação ótimo encontrado pela rede Adaline.

5.3.4 Aspectos gerais das RNAs

Nas seções anteriores foram apresentadas a estrutura de um neurônio artificial e duas redes historicamente importantes para o desenvolvimento da área de neurocomputação. Ao longo dos anos foram projetadas redes neurais mais complexas para acompanhar as exigências da indústria e da área acadêmica para a resolução de problemas cada vez mais difíceis. Num modo geral, essas redes são formadas por vários neurônios interconectados e organizados em arquiteturas compostas por mais de uma camada de neurônios. No figura 5.13, é exposto um exemplo de uma rede Perceptron Multicamadas que será discutida na seção 5.3.5. Pode-se observar que a rede em questão é composta por três camadas de neurônios (os neurônios são representados pelos círculos vermelhos), onde as camadas 1 e

2 são denominadas camadas escondidas e a camada 3, que produz o sinal de saída, recebe o nome de camada de saída. Generalizando, pode-se criar redes com um número arbitrário de camadas, caso uma rede tenha N camadas, ela possuirá $N - 1$ camadas escondidas e uma camada de saída.

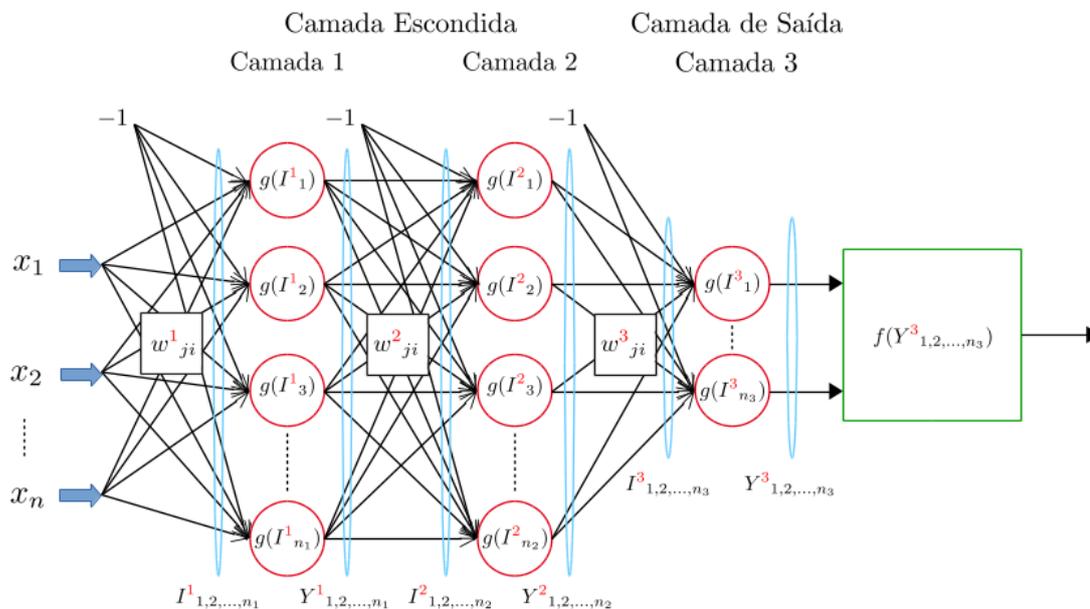


FIGURA 5.13 – Exemplo da forma geral de uma rede Perceptron Multicamadas com 3 camadas. O quadrado verde no final da rede representa o pós-processamento dos sinais de saída para a identificação da classe do evento, quando há mais de um neurônio na camada de saída.

As camadas escondidas são responsáveis por extrair as características associadas ao processo ou sistema a ser inferido. A camada de saída é responsável pela produção dos resultados finais da rede. Muitas vezes é utilizada também a expressão “camada de entrada”, referindo-se as variáveis de entrada da rede e seu pré-processamento, essa terminologia pode gerar confusão para iniciantes. Neste trabalho, o uso do termo “camada” é usado somente para se referir as camadas de neurônios da rede.

5.3.4.1 Treinamento

O processo de treinamento de uma rede neural consiste de passos ordenados, necessários para a sintonização dos pesos sinápticos e limiares de seus neurônios, tendo-se como objetivo final a otimização das respostas produzidas pelas saídas da rede, as quais são representativas do sistema físico que está sendo mapeado. O conjunto desses passos ordenados visando o treinamento da rede é denominado de algoritmo de aprendizagem. Cada tipo de RNA possui algoritmos de aprendizagem com processos de treinamento específicos, os quais podem ser classificados como,

- Treinamento supervisionado: tendo-se disponível na amostra de treinamento os sinais de entrada e suas respectivas saídas desejadas. Os pesos sinápticos e limiares são ajustados supervisionando-se a defasagem entre as respostas produzidas pela rede em relação àsquelas desejadas, sendo essa diferença usada no procedimento de ajuste. A rede será considerada treinada quando tal defasagem tiver alcançado seu valor mínimo;
- Treinamento não-supervisionado: diferente do treinamento supervisionado, as saídas referentes aos sinais de entrada não estão disponíveis na amostra de treinamento. Desse modo, o treinamento é realizado pelo ajuste dos pesos e limiares da rede visando a auto-organização com relação as particularidades do sistema, identificando subconjuntos (*clusters*) que possuam similaridades. A quantidade máxima de possíveis *clusters* pode ser especificada, a priori, baseando-se em um conhecimento prévio do sistema estudado.
- Treinamento com reforço: usado em problemas em que se deve tomar decisões sequencialmente, tendo como objetivo avaliar a melhor decisão que se deve tomar em uma situação específica. O treinamento ajusta os parâmetros da rede baseando-se em informações advindas da interação com o sistema que está sendo mapeado, seguindo um processo de tentativa e erro. Em cada etapa, a entrada é avaliada como satisfatória ou não, caso seja satisfatória, são feitos incrementos nos pesos e limiares visando reforçar esse comportamento.

5.3.4.2 Pré-processamento das entradas

As diferentes variáveis que compõem o sinal de entrada da rede podem possuir domínios bem distintos, os quais podem estar (e geralmente estão) fora do intervalo de variação dinâmica das funções de ativação, o que dificultaria bastante o processo de treinamento. Devido a isso, recomenda-se redimensionar as variáveis de entrada antes de fornecê-las a rede neural. O procedimento adotado nesta análise, redimensiona cada variável de entrada, em todas as amostras, utilizando a expressão:

$$x = \frac{x_{\text{raw}} - \bar{x}}{\sigma_x} . \quad (5.11)$$

Sendo \bar{x} e σ_x , respectivamente, a média e o desvio-padrão da amostra de treinamento. O valor da variável antes do pré-processamento é denotado por x_{raw} , e x é o valor redimensionado da variável que resulta, nas amostras, em distribuições com média próxima a zero e desvio padrão em torno de um.

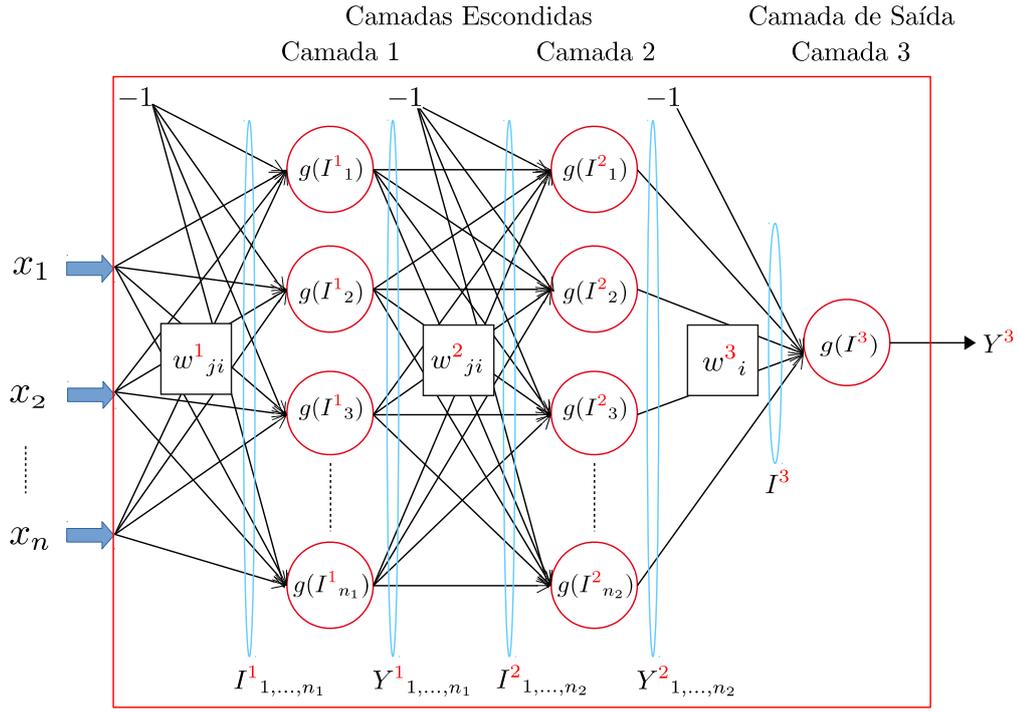
5.3.5 Rede Perceptron Multicamadas

Os diferentes tipos de RNAs existentes podem ser classificados pela forma de sua arquitetura, ou seja, disposição dos neurônios, forma de interligação entre eles e constituição de suas camadas. Um dos tipos principais é a arquitetura *feedforward*, em que o fluxo de informação segue somente uma direção, ou seja, da camada de entrada à camada de saída. Nos casos mais simples em que existe somente uma camada, estão inseridas as redes Perceptron e Adaline. Redes que apresentam múltiplas camadas são usadas para aproximação de funções, classificação de padrões, identificação de sistemas, robótica, controle de processos, etc. Entre as principais redes com arquitetura *feedforward* de múltiplas camadas, estão o Perceptron Multicamadas (*Multilayer Perceptron* - MLP), as redes neurais convolucionais (*convolutional neural network* - CNN) e as redes de base radial (*radial basis function* - RBF).

Um neurônio artificial pode discriminar no máximo duas classes de eventos e apenas de forma linear. A análise de sistemas mais complexos requer o uso de vários neurônios conectados de maneira adequada entre si, formando uma rede capaz de discriminar duas ou mais classes de eventos de forma não linear. A Rede Perceptron Multicamadas, ou rede MLP, é a rede neural com arquitetura *feedforward* mais tradicional em classificação de padrões e já tem seu uso difundido em física de partículas e altas energias. A topologia de uma rede MLP é determinada pelo número de camadas e neurônios presentes em cada camada. Na prática, a aplicação da rede MLP se dá pela tentativa de diferentes topologias, escolhendo-se a que discrimina melhor as classes de eventos. Esta análise faz uso de redes MLP para resolver problemas de classificação envolvendo duas classes eventos, geralmente referidas como classe de sinal e classe de *background*. Portanto são utilizadas redes MLP com apenas um neurônio na camada de saída, como mostrado no exemplo da figura 5.14 que apresenta um rede MLP com 3 camadas.

A informação na rede MLP é processada da seguinte maneira. O sinal de entrada, $\{x_1, \dots, x_n\}$ é fornecido à primeira camada de neurônios da rede, esses dados são ponderados pela matriz de pesos da primeira camada, w^1_{ji} , resultando nos potenciais de ativação I^1_i . Aplica-se a função de ativação sobre I^1_i , resultando no vetor Y^1_i que é a saída da primeira camada e servirá de entrada para a segunda camada de neurônios. Esse processo se repete nas camadas seguintes até chegar à camada de saída que produzirá o classificador $Y^3(\mathbf{x})$, que representa a variável discriminante da rede MLP.

A produção de uma MLP, destinada a classificar as classes de eventos de um determinado sistema, implica em definir uma função de ativação, estipular a topologia da rede e encontrar os valores adequados dos pesos sinápticos por meio de um algoritmo de treinamento. Na versão tradicional da rede MLP, as funções de ativação geralmente escolhidas eram as funções tangente hiperbólica e sigmóide, figura 5.15 e o algoritmo de treinamento



$$I^1_j = w^1_{ji} \cdot x_i \rightarrow Y^1_j = g(I^1_j) \rightarrow I^2_j = w^2_{ji} \cdot Y^1_i \rightarrow Y^2_j = g(I^2_j) \rightarrow I^3 = w^3_i \cdot Y^2_i \rightarrow Y^3 = g(I^3)$$

FIGURA 5.14 – Rede Perceptron Multicamadas com três camadas e apenas um neurônio na última camada. Abaixo da rede é mostrado o fluxo completo de processamento da informação na rede, desde as variáveis de entrada, x_i , até o sinal de saída, Y^3 .

conhecido como *backpropagation* era usado extensamente. Segundo esse algoritmo, as correções dos pesos são realizadas sequencialmente, começando pelos pesos na última camada e depois passando para a camada anterior, até chegar a primeira. Como no caso da rede Adaline, o objetivo é alcançar o mínimo de uma função, chamada de função de perda (*loss function*) ou função objetiva (*objective function*), $f^a(\mathbf{w})$. Considerando o exemplo da figura 5.14, ao definir-se a função de perda como sendo a função erro quadrático,

$$f^a(\mathbf{w}) \equiv \frac{1}{2} \sum_{j=1}^{n_3} (d_j^a - (Y^3_j)^a)^2, \tag{5.12}$$

as correções, Δw^l_{ji} , aos pesos de cada camada, l , são proporcionais ao gradiente da função erro quadrático com relação a w^l_{ji} , apontando no sentido de decréscimo da função de perda, de modo que

$$(w^l_{ji})^{\text{nov}} = (w^l_{ji})^{\text{atual}} - \eta \cdot g^{la}_{ji}, \text{ com } g^{la}_{ji} \equiv \frac{\partial f^a(\mathbf{w})}{\partial w^l_{ji}}. \tag{5.13}$$

A versão tradicional da rede MLP foi ultrapassada em termos de performance quando começaram a surgir inovações para melhorar o desempenho das redes MLP profundas.

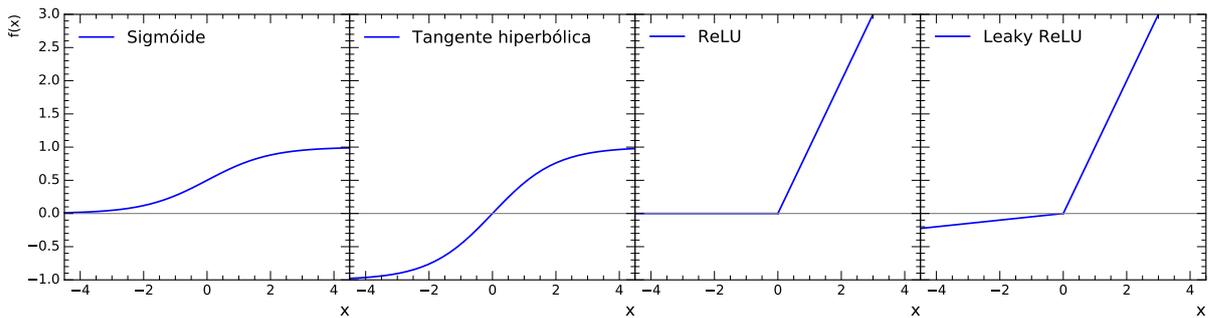


FIGURA 5.15 – Funções de ativação usadas nas redes MLP.

Embora não exista um conceito exato, são tipicamente consideradas redes rasas, aquelas compostas por uma ou duas camadas de neurônios enquanto aquelas compostas por mais de duas camadas são chamadas de redes neurais profundas (*deep neural networks - DNN's*). Nos últimos anos, o uso de redes profundas na solução de problemas tem aumentado, devido ao aumento na rapidez do processo de treinamento e melhora na performance de classificação. O uso de novas funções de ativação como a função ReLU e suas variantes, figura 5.15, e o surgimento de novos algoritmos de aprendizagem, como o Adam (*Adaptive Moment Estimation*), contribuíram bastante para esse novo cenário.

Conforme mostrado na equação (5.13), a atualização dos pesos depende de gradientes da função de perda, que resulta em gradientes das funções de ativação. Entretanto, as funções sigmóide e tangente hiperbólica alcançam uma região de saturação em valores distantes de zero, fazendo com que os gradientes das funções de ativação tendam a zero, dificultando a convergência do processo de aprendizagem. A função de ativação ReLU não sofre do problema do gradiente e é mais simples de ser calculada, aumentando bastante o processo de convergência. Ela é uma das funções de ativação mais usadas no momento (provavelmente a mais usada), sendo a recomendação padrão para redes MLP e CNN's. A função ReLU é igual a zero em valores negativos, o que pode fazer com que neurônios que resultem em valores negativos nunca mais sejam ativados. Variantes da função ReLU, como a função Leaky ReLU, propõem pequenas modificações para evitar que a função seja zero em valores negativos. Os benefícios decorrentes do uso das variantes da função ReLU não são muito claros, em geral, as funções ReLU são utilizadas a não ser que seja observado algum problema decorrente de seu uso. Em problemas de duas classes, a função sigmóide continua sendo usada na camada de saída.

O Adam é um algoritmo de otimização baseado em estimativas adaptativas dos momentos do gradiente de ordem mais baixa [157], sendo m a estimativa associada ao primeiro momento e v a estimativa associada ao segundo momento. Seus parâmetros tipicamente requerem pouco ajuste e sua configuração padrão é dada por $\eta = 0,001$, $\beta_1 = 0,9$,

$\beta_2 = 0,999$ e $\epsilon = 10^{-8}$. Os pesos das camadas são ajustados pela expressão:

$$(w_{ji}^l)^{\text{nov}} = (w_{ji}^l)^{\text{atual}} - \eta \cdot \frac{(m_{ji}^{la})^{\text{nov}}}{\sqrt{(v_{ji}^{la})^{\text{nov}} + \epsilon}}, \quad (5.14)$$

tal que

$$(m_{ji}^{la})^{\text{nov}} = \frac{\beta_1 \cdot (m_{ji}^{la})^{\text{atual}} + (1 - \beta_1) \cdot g_{ji}^{la}}{1 - \beta_1^t}, \quad (5.15)$$

$$(v_{ji}^{la})^{\text{nov}} = \frac{\beta_2 \cdot (v_{ji}^{la})^{\text{atual}} + (1 - \beta_2) \cdot (g_{ji}^{la})^2}{1 - \beta_2^t}. \quad (5.16)$$

Os tensores m e v são iniciados todos os valores iguais a zero e o expoente t é igual ao número da época em que estão sendo computados os novos pesos. O Adam geralmente tem melhores resultados que os outro algoritmos de otimização. A tabela 5.1 compara as características dos diferentes tipos de redes neurais apresentados nesta seção. As características são identificadas na primeira coluna à esquerda, sendo o número de classes que a rede é capaz de separar, o tipo de separação realizada no espaço das variáveis, o algoritmo de treinamento tipicamente utilizado, a função de ativação usada nos neurônios das camadas escondidas e a função de ativação usada na camada de saída, em problemas com duas classes de eventos. As redes neurais computadas na seções seguintes utilizam o pacote computacional Keras [158].

	Perceptron	Adaline	MLP (Tradicional)	MLP (Profunda)
Número de classes	2	2	≥ 2	≥ 2
Separação	Linear	Linear	Não linear	Não linear
Treinamento	Regra de Hebb	Regra Delta	<i>Backpropagation</i>	Adam
Função de ativação	Sinal	Linear	Sigmóide	ReLU
Saída (2 classes)	Linear	Sinal	Sigmóide	Sigmóide

TABELA 5.1 – Exposição das características de diferentes tipos de redes neurais.

5.4 Região de controle de $Z(\nu\nu)+\text{Jatos}$

A região de controle de eventos de $Z(\nu\nu)+\text{Jatos}$ foi definida na subseção 4.6.4 através de uma análise multivariada. Nesta seção, são descritos todos os aspectos da análise realizada e seus resultados, sendo adotada uma abordagem didática para o completo entendimento de cada etapa do procedimento. As análises multivariadas realizadas neste trabalho têm sempre como objetivo a classificação das entradas da amostra fornecida em duas classes distintas, as quais são chamadas de classe de sinal e de *background*. Os eventos da classe de sinal são sempre compostos por uma mistura de eventos provenientes de cada

ponto do *grid* de sinal, e os eventos da classe de *background* são compostos por eventos de *background* cuja composição utilizada na definição da região de controle de $Z(\nu\nu)+\text{Jatos}$ é diferente da utilizada na definição das regiões de sinal. A classificação é realizada por RNAs e o valor de saída da rede é tratado com uma variável discriminante entre as duas classes de eventos, na qual são aplicados cortes para definir as regiões de interesse.

5.4.1 Amostras de entrada

As amostras de entrada utilizadas na definição da região de controle de $Z(\nu\nu)+\text{Jatos}$ são compostas por eventos em torno do pico que existe na distribuição de $\hat{\omega}_{\min}$ da figura 4.18. Como mostrado na seção 4.6.4, esse pico é composto em sua maior parte por eventos que possuem apenas um jato reconstruído com $p_T > 30 \text{ GeV}$. No esquema da figura 4.22, essa região é retratada pelo retângulo branco dentro da região de busca, no intervalo $0,78 \leq \hat{\omega}_{\min} < 0,8$. Nessa região, a contribuição predominante do *background* é de eventos com $Z(\nu\nu)+\text{Jatos}$. A segunda maior contribuição é de eventos com $W(l\nu)+\text{Jatos}$ que não tiveram o lépton carregado reconstruído, resultando em eventos extremamente semelhantes aos eventos de $Z(\nu\nu)+\text{Jatos}$, de tal modo, que a discriminação desses dois tipos de *background* não é muito efetiva. Portanto o foco da análise multivariada é definir uma região com uma contaminação menor de eventos de sinal e que ainda possua uma quantidade suficiente de eventos com $Z(\nu\nu)+\text{Jatos}$ para que sejam feitas as correções na simulação. Para isso, a classe de *background* é definida como eventos $Z(\nu\nu)+\text{Jatos}$, e a classe de sinal como uma mistura de eventos de sinal oriundos de diferentes pontos do *grid*. Os eventos correspondentes às outras contribuições de *background* não são usados na análise multivariada, explicitando nesse caso a diferença entre os termos “eventos de *background*” e “classe de *background*”, o primeiro termo sempre se refere a todos os processos do MP, já o segundo refere-se exclusivamente aos processos utilizados na análise multivariada. As variáveis utilizadas para discriminar as duas classes de eventos são as variáveis globais, exceto N_{SV1Jet} e N_{SV2Jets} que são usadas na definição das regiões de validação.

A mistura de eventos de diferentes pontos do *grid* de sinal é realizada de modo a dar maior peso aos pontos com maior massa do gluino e menor Δm , por serem os pontos com menor sensibilidade devido à baixa seção de choque em pontos de maior massa do gluino e a dificuldade de reconstrução de eventos extremamente comprimidos. Sendo assim, os pesos dos eventos de sinal correspondentes a cada ponto $(m_{\tilde{g}}, \Delta m)$ são normalizados para que sua soma seja igual a 1, e depois eles são reponderados através da equação,

$$w(m_{\tilde{g}}, \Delta m) = (0,0009 m_{\tilde{g}} - 0,125) \cdot (-0,012857 \Delta m + 1,38571). \quad (5.17)$$

A ponderação foi feita para que a contribuição dos casos mais extremos do *grid* de sinal,

$(m_{\tilde{g}} = 1250 \text{ GeV}, \Delta m = 30 \text{ GeV})$ e $(m_{\tilde{g}} = 250 \text{ GeV}, \Delta m = 100 \text{ GeV})$ tenham, respectivamente, peso 1 e 0,01 na mistura de eventos de sinal usada na análise multivariada, como pode ser visto na figura 5.16. Ao final, os pesos da mistura são normalizados novamente para que sua soma seja igual a 1.

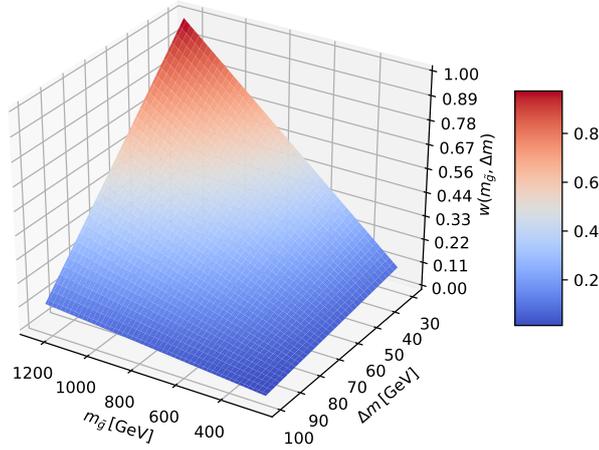


FIGURA 5.16 – Função utilizada na ponderação dos eventos da classe de sinal, descrita pela equação (5.17).

Os eventos da classe de *background* também são normalizados para que sua soma seja igual a 1, porém, para o caso específico da definição da região de controle de $Z(\nu\nu)+\text{Jatos}$, são realizadas outras 4 análises multivariadas, nas quais são consideradas normalizações dos eventos da classe de *background* iguais a 2, 3, 4 e 5. O aumento no valor da normalização dos eventos da classe de *background* faz com que eles tenham maior influência na definição da distribuição da variável de saída da rede neural, quanto maior o valor da normalização, mais concentrados à esquerda estarão os eventos de $Z(\nu\nu)+\text{Jatos}$. Por outro lado, à medida que o valor da normalização da classe de *background* aumenta, os eventos de sinal perdem influência na distribuição da variável de saída, conseqüentemente há uma perda no poder de discriminação entre os eventos das classes de sinal e *background*.

Os eventos das classes de sinal e *background* são separados em 6 amostras distintas mas com o mesmo número de eventos simulados ($16,6\%$) e estatisticamente idênticas. A amostra 1 é usada no treinamento do algoritmo de aprendizagem, a amostra 2 é usada na validação da aprendizagem e otimização da rede neural e as amostras 3, 4, 5 e 6 são as amostras de teste que são usadas para checar se a rede neural tem resultados estatisticamente compatíveis entre todas as amostras. Após a definição da rede neural e aplicação do corte na variável discriminante, as estimativas físicas são feitas utilizando somente os eventos das amostras 2, 3, 4, 5 e 6. A amostra 1 não é utilizada porque a rede neural foi treinada usando os eventos dessa amostra, portanto, mesmo que seja evitado o *overtraining*, é uma prática comum não utilizar essa amostra em estimativas posteriores

para evitar qualquer tipo de viés nos resultados. O esquema da figura 5.17 mostra, em resumo, como é feita a divisão de amostras.

As distribuições das classes de sinal e *background*, com respeito as 12 variáveis de entrada na amostra de treinamento, são mostradas na figura 5.18⁶ e a matriz de correlação linear entre elas é mostrada na figura 5.19. Nessa região, não há léptons reconstruídos e a maior parte dos eventos possui apenas um jato com $p_T > 30$ GeV, isso faz com que as variáveis hadrônicas E_T^{miss} , H_T^{miss} , H_T , H_T^{PV} e $\text{Jet}_L^{\text{PV}} p_T$ apresentem uma grande correlação linear entre elas. Ademais, os eventos de sinal possuem em geral uma maior atividade hadrônica, em comparação aos eventos com $Z(\nu\nu)+\text{Jatos}$, devido aos jatos pouco energéticos oriundos do decaimento dos gluinos.

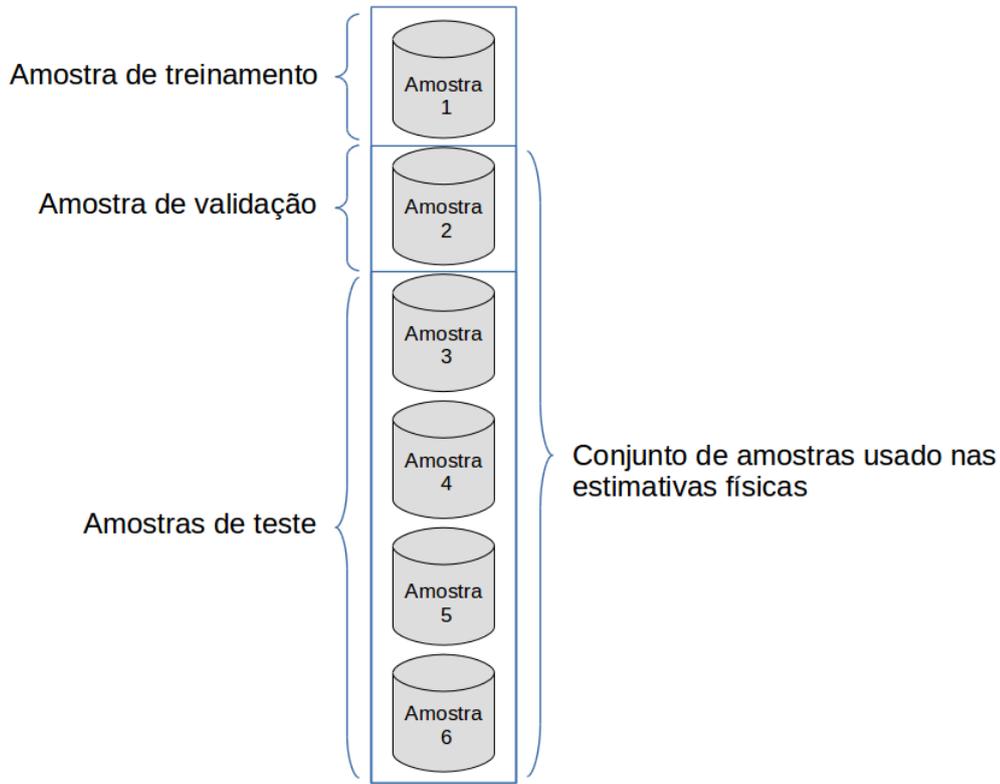


FIGURA 5.17 – Figura esquemática representando a divisão dos dados simulados entre as amostras de treinamento, validação e teste.

⁶Os nomes das variáveis que aparecem nos gráficos são aqueles utilizados no código da análise. A seguir é exibida a correspondência entre eles e os nomes adotados neste trabalho: $E_T^{\text{miss}} \Leftrightarrow \text{MET}$, $H_T^{\text{miss}} \Leftrightarrow \text{MHT}$, $H_T \Leftrightarrow \text{HT}$, $H_T^{\text{PV}} \Leftrightarrow \text{SumPVJetsPt}$, $\text{Jet}_L^{\text{PV}} p_T \Leftrightarrow \text{LeadingPVJetPt}$, $\text{Jet}_S^{\text{PV}} p_T \Leftrightarrow \text{SubLeadingPVJetPt}$, $N_{\text{jets}} \Leftrightarrow \text{NJets}$, $N_{\text{jets}}^{\text{PV}} \Leftrightarrow \text{NPVJets}$, $N_{\text{bjets}} \Leftrightarrow \text{NBJets}$, $\hat{\omega}_{\text{min}} \Leftrightarrow \text{OmegaMin}$, $\chi_{\text{min}} \Leftrightarrow \text{ChiMin}$ e $f_{\text{max}} \Leftrightarrow \text{FMax}$.

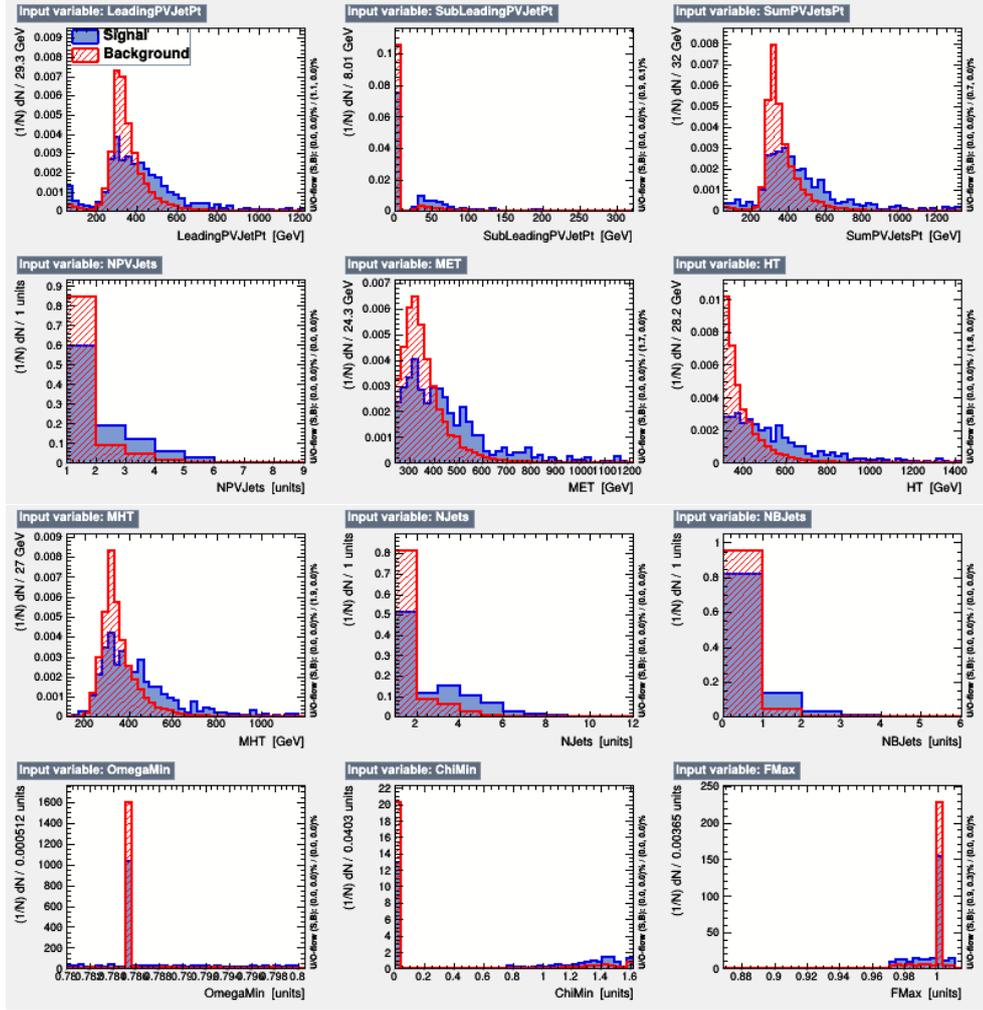


FIGURA 5.18 – Distribuições das classes de sinal e *background* com respeito as 12 variáveis de entrada na amostra de treinamento.

O quadro abaixo resume as informações gerais da análise multivariada realizada na definição da região de controle de $Z(\nu\nu)+\text{Jatos}$.

- ★ **Objetivo:** Definir uma região de controle para eventos com $Z(\nu\nu)+\text{Jatos}$;
- ★ **Região das amostras:** Eventos que passam pela pré-seleção da seção 4.4, estão no *plateau* do *trigger* de busca ($H_T^{\text{miss}} \geq 220 \text{ GeV}$, $E_T^{\text{miss}} \geq 240 \text{ GeV}$) e possuem $0,78 \leq \hat{\omega}_{\min} < 0,8$;
- ★ **Classe de sinal:** Mistura de eventos simulados de diferentes pontos do *grid* de sinal;
- ★ **Classe de *background*:** Eventos simulados de $Z(\nu\nu)+\text{Jatos}$;
- ★ **Variáveis de entrada:** E_T^{miss} , H_T^{miss} , H_T , H_T^{PV} , $\text{Jet}_L^{\text{PV}} p_T$, $\text{Jet}_S^{\text{PV}} p_T$, N_{jets} , $N_{\text{jets}}^{\text{PV}}$, N_{bjets} , $\hat{\omega}_{\min}$, χ_{\min} e f_{\max} .

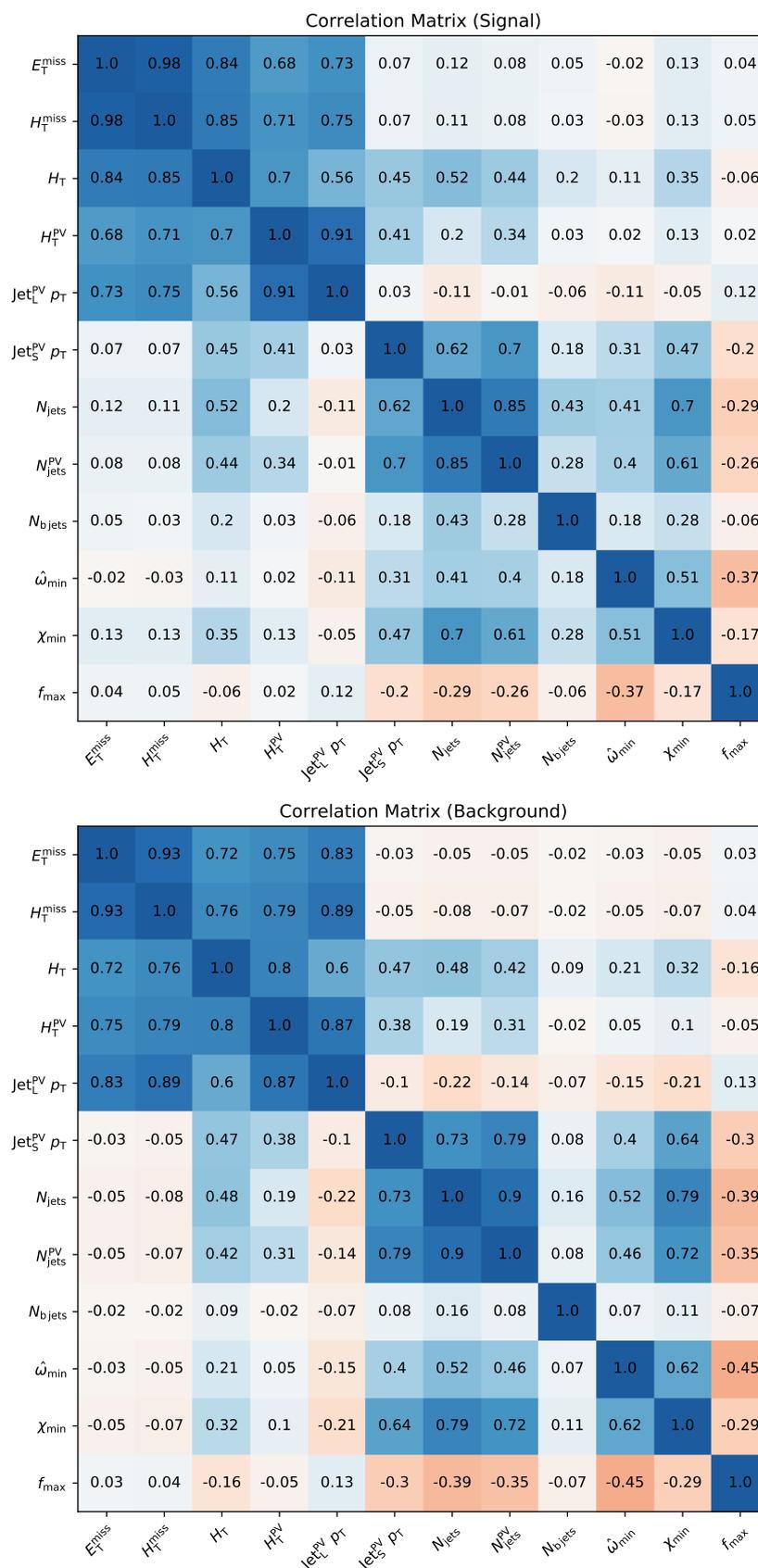


FIGURA 5.19 – Matriz de correlação linear das classes de sinal (acima) e *background* (abaixo) com respeito as 12 variáveis de entrada na amostra de treinamento.

5.4.2 Treinamento

São treinadas 100 redes neurais candidatas varrendo um *grid* dependente do número de camadas e do número de neurônios por camada. São propostas redes com 1 a 10 camadas. Em cada rede, o número de neurônios por camada é definido com relação ao número de variáveis de entrada. Existindo N variáveis de entrada, o número de neurônios por camada é proposto para ser igual a $2N$, $4N$, $6N$, $8N$, $10N$, $12N$, $14N$, $16N$, $18N$ ou $20N$. O treinamento foi configurado para ter um máximo de 300 épocas ou para encerrar quando, após uma determinada época EP , a função de perda na amostra de validação resulta em valores maiores que o obtido em EP nas 30 épocas subsequentes. O resultado do treinamento é uma rede neural com a configuração de pesos e limiares existentes na época EP , evitando o *overtraining* da rede. Dificilmente o treinamento dura 300 épocas, em geral as redes convergem bem antes desse valor. A figura 5.20 mostra a evolução do valor da função de perda no decorrer das épocas de treinamento da rede com 2 camadas e 48 neurônios por camada e que possui normalização dos eventos da classe de *background* igual a 2.

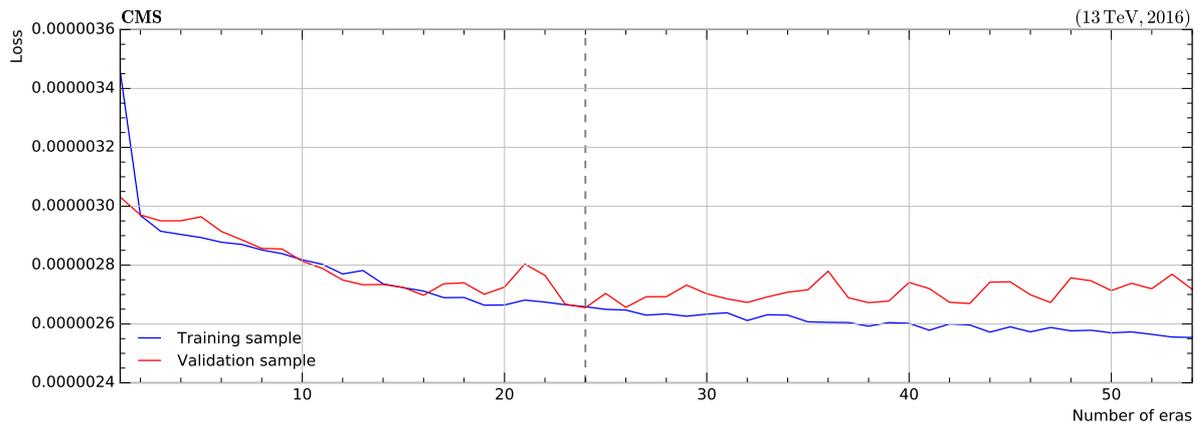


FIGURA 5.20 – Valor da função de perda no decorrer das épocas de treinamento da rede com 2 camadas, 48 neurônios por camada e normalização dos eventos da classe de *background* igual a 2. O treinamento foi encerrado na época 54 e o resultado é a rede neural com configuração de pesos da época 24, denotada no gráfico pela linha vertical tracejada.

A métrica usada para medir a performance das redes neurais após o treinamento é a área sob a curva ROC. A rede neural com melhor performance na definição da região de controle de eventos com $Z(\nu\nu)+\text{Jatos}$ foi justamente aquela citada no parágrafo anterior com 2 camadas, 48 neurônios por camada e normalização dos eventos da classe de *background* igual a 2, possuindo $AUC = 0,819$. Sua curva ROC pode ser observada na figura 5.21, junto com a curva da segunda rede de melhor performance e das duas BDTs com melhor performance de um *grid* de BDTs com 100, 200 ou 300 árvores, profundidades máximas iguais a 3, 4, 5 ou 6 e $\beta = 0,1$; o qual será usado como *grid* de referência

para comparação com os resultados das redes neurais. A melhor BDT, com 200 árvores e profundidade máxima igual a 3, possui $AUC = 0,801$.

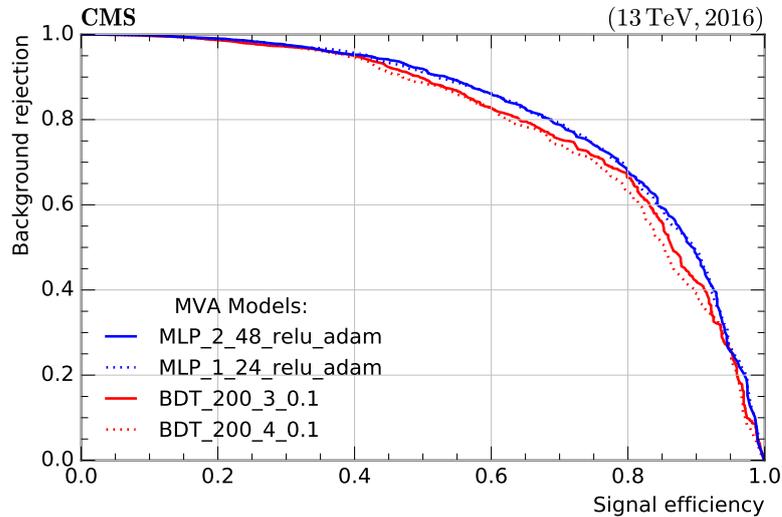


FIGURA 5.21 – Curva ROC das duas redes MLP e das duas BDTs com melhor performance.

5.4.3 Checagem de *overtraining*

Após o treinamento, a saída da rede MLP de melhor performance, com respeito a todos os processos de *background* e sinal, resulta na distribuição observada na figura 4.30, construída com os eventos das amostras 2, 3, 4, 5 e 6. O corte superior em 0,6 na variável discriminante define a região de controle de eventos com $Z(\nu\nu)+J$ atos. Portanto a análise multivariada nesse caso possui duas regiões de interesse separadas pelo corte em 0,6. Na figura 5.22, é verificado se as estimativas das classes de sinal e *background* são estatisticamente semelhantes entre as amostras de treinamento e as outras cinco amostras de dados simulados. Na parte de cima do gráfico são expostas as frações de eventos das classes de sinal e *background* em cada uma das regiões de interesse para as amostras de treinamento e validação. Obviamente não é esperado que as frações relativas à amostra de treinamento sejam exatamente iguais às frações da amostra de validação, mas espera-se que a diferença seja proveniente de incertezas estatísticas. Na parte de baixo do gráfico, as frações relativas às amostras de treinamento e validação são divididas pela média das frações de eventos das amostras 2, 3, 4, 5 e 6 em cada região, e os valores dessas razões são comparados com o desvio-padrão com relação à média. Observa-se que tanto os eventos da classe de sinal quanto os da classe de *background* estão dentro do intervalo de 2σ para a amostra de treinamento, logo, as frações são estatisticamente compatíveis. Essa verificação ressalta o fato de que todas as seis amostras descrevem distribuições estatisticamente semelhantes e que não existe um viés significativo com respeito ao valor

estimado pela amostra de treinamento, que poderia ocorrer no caso de *overtraining*.

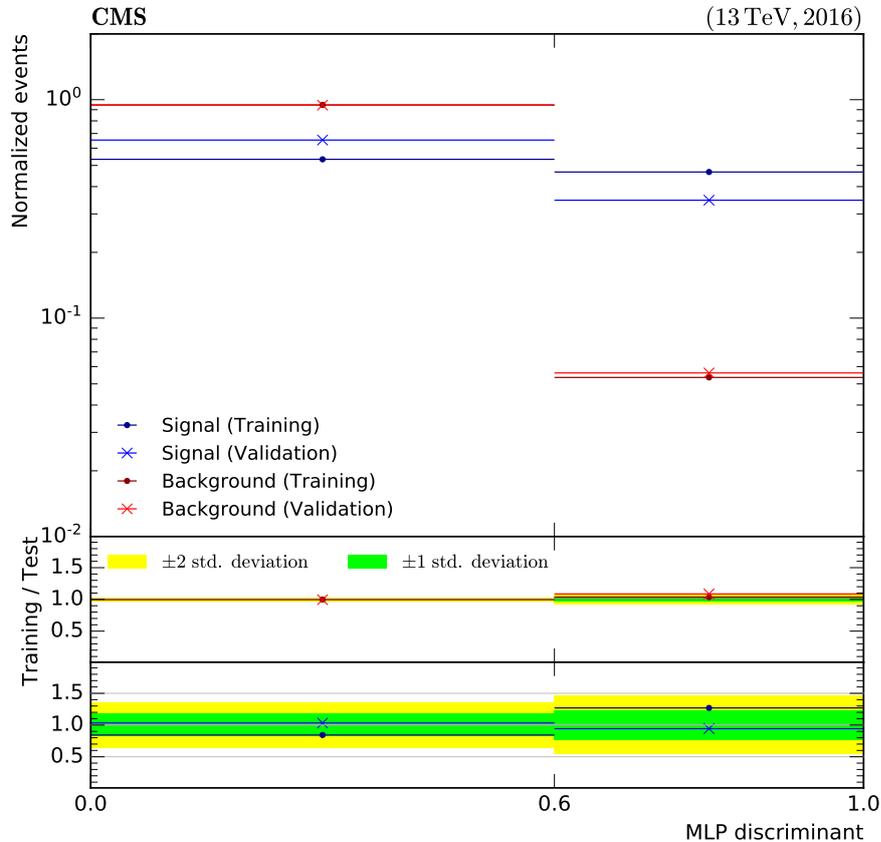


FIGURA 5.22 – Validação da compatibilidade entre as estimativas da amostra de treinamento e as estimativas das outras cinco amostras de dados simulados.

5.5 Regiões de sinal

Esta seção descreve a definição das regiões de sinal da análise usando a variável discriminante produzida pelas RNAs treinadas em cada uma das quatro sub-regiões de busca. A classe de sinal é composta pela mesma mistura de pontos do *grid* de sinal descrita na subseção 5.4.1, já a classe de *background* é composta agora por todos os processos do MP que contribuem como eventos de *background*. Ambas as classes são normalizadas antes do treinamento para que a soma dos pesos dos seus eventos seja igual a 1. Também é utilizado o mesmo procedimento da subseção 5.4.1 na definição das amostras de treinamento, validação e teste; e a mesma metodologia da subseção 5.4.2 para realizar o treinamento das RNAs.

Em R1, a rede neural que obteve a melhor performance possui uma camada escondida com 190 neurônios, sendo treinada em 8 épocas, sem sinais de *overtraining*. Sua curva ROC possui $AUC = 0,804$ enquanto a melhor BDT teve $AUC = 0,800$, portanto, o

método com BDT's não perde muito em performance para as RNAs nessa sub-região. A distribuição da variável discriminante produzida pela rede MLP é exibida na parte inferior da figura 5.23. Nela são identificadas a região de controle MLP CR1 e os *bins* de busca da região de sinal. A região de controle, onde é observado um bom acordo entre a distribuição da simulação e os dados reais com um erro relativo de 2,7% no número total de eventos, é definida com um corte superior na variável discriminante em 0,4, de modo a garantir que a contaminação de sinal em cada *bin* da região de controle permaneça no máximo entre 10 a 20% da incerteza estatística do *background*, gráfico superior à esquerda na figura 5.23, Esse critério também é adotado em R2, R3 e R4. O primeiro *bin* de busca de R1 é obtido pela determinação do corte inferior que maximiza a significância do sinal(1000,30). O modelo $m_{\tilde{g}} = 1250$ GeV prevê uma quantidade muito pequena de eventos, por essa razão a estimativa é realizada utilizando o modelo com $m_{\tilde{g}} = 1000$ GeV. A expressão utilizada para calcular a significância, Sig, é dada por [159],

$$\text{Sig} = \sqrt{2(S + B) \ln \left(1 + \frac{S}{B} \right) - 2S}, \quad (5.18)$$

sendo S o número de eventos de sinal na região e B o número de eventos de *background*. As curvas de significância para modelos com $m_{\tilde{g}} = 1000$ GeV são mostradas no gráfico superior à direita na figura 5.23. Dessa forma, o primeiro *bin* é definido com um corte inferior em 0,97. Os outros *bins* são definidos em intervalos com valores iguais ou próximos a 0,05, totalizando 13 *bins* de busca em R1, numerados de 1 a 13.

Em R2, a melhor performance foi obtida pela rede MLP que possui uma camada escondida com 240 neurônios, treinada em 20 épocas, sem sinais de *overtraining*. Sua curva ROC possui $\text{AUC} = 0,793$ ao passo que a melhor BDT teve $\text{AUC} = 0,792$, logo, os métodos tiveram performances semelhantes nessa sub-região. A região de controle MLP R2 e os *bins* de busca são definidos pelos mesmos critérios adotados em R1. Sendo assim, a região de controle MLP R2 é limitada com um corte superior em 0,32, com um bom acordo com relação aos dados, na forma da curva do discriminante, com um erro relativo de 4,8% no número total de eventos, e o primeiro *bin* de busca é determinado com um corte inferior em 0,95. A numeração dos *bins* de busca começa a partir da numeração do último *bin* da sub-região anterior, desse modo, são adicionados mais 13 *bins* na análise, numerados de 14 a 26.

Em R3, a melhor performance foi da rede MLP com 2 camadas e 30 neurônios em cada camada, treinada em 31 épocas, sem sinais de *overtraining*. Sua curva ROC possui $\text{AUC} = 0,826$ e a melhor BDT teve $\text{AUC} = 0,815$, mostrando uma melhora da performance ao usar RNAs. A região de controle MLP R3 é definida com um corte superior em 0,47, exibindo um ótimo acordo com os dados reais, apresentando um erro relativo de 1,6% no número total de eventos. O ponto de sinal(1000,30) não possui um ponto de máximo em valores

altos da variável discriminante, desse modo, aplicou-se um corte inferior em 0,9, que fica próximo aos pontos de máximo das outras curvas de significância com $m_{\tilde{g}} = 1000 \text{ GeV}$. Um total de 9 *bins* de busca são definidos em R3, numerados de 27 a 35.

Em R4, a rede MLP com 2 camadas e 96 neurônios em cada camada teve a melhor performance, sendo treinada em 39 épocas, sem sinais de *overtraining*. Sua curva ROC possui $AUC = 0,843$ enquanto a melhor BDT teve $AUC = 0,829$, como em R3, há uma melhora da performance ao usar RNAs. A região de controle MLP R4 é definida com um corte superior em 0,50, e possui um ótimo acordo com os dados reais, com um erro relativo de 1,4% no número total de eventos. Como aconteceu em R3, o ponto de sinal(1000,30) não possui uma significância máxima em valores altos da variável discriminante, portanto, aplicou-se um corte inferior em 0,95, que fica próximo ao valor máximo das outras curvas de significância com $m_{\tilde{g}} = 1000 \text{ GeV}$. São definidos 10 *bins* de busca em R4, numerados de 36 a 45.

Nas subseções seguintes, são exibidas as informações gerais de cada análise multivariada e os resultados obtidos em cada sub-região de busca. As distribuições das variáveis de entrada, matrizes de correlação linear e os gráficos de controle e de performance das RNAs são expostos no apêndice D.

5.5.1 Análise multivariada em R1

- ★ **Objetivo:** Definir as regiões de sinal em R1;
- ★ **Região das amostras:** Eventos que estão na região de busca definida na seção 4.5 e possuem pelo menos um SV2Jets;
- ★ **Classe de sinal:** Mistura de eventos simulados de diferentes pontos do *grid* de sinal;
- ★ **Classe de *background*:** Eventos simulados dos processos do MP;
- ★ **Variáveis de entrada:** E_T^{miss} , H_T^{miss} , H_T , H_T^{PV} , $\text{Jet}_L^{\text{PV}} p_T$, $\text{Jet}_S^{\text{PV}} p_T$, N_{jets} , $N_{\text{jets}}^{\text{PV}}$, N_{bjets} , $\hat{\omega}_{\text{min}}$, χ_{min} , f_{max} , M_{JJ} , $\text{Jet}_L^{\text{SV}} p_T$, $\text{Jet}_S^{\text{SV}} p_T$, $\text{SV } \chi^2$, SV NDF , $\Delta\eta_{\text{JJ}}$ e ΔR_{JJ} .

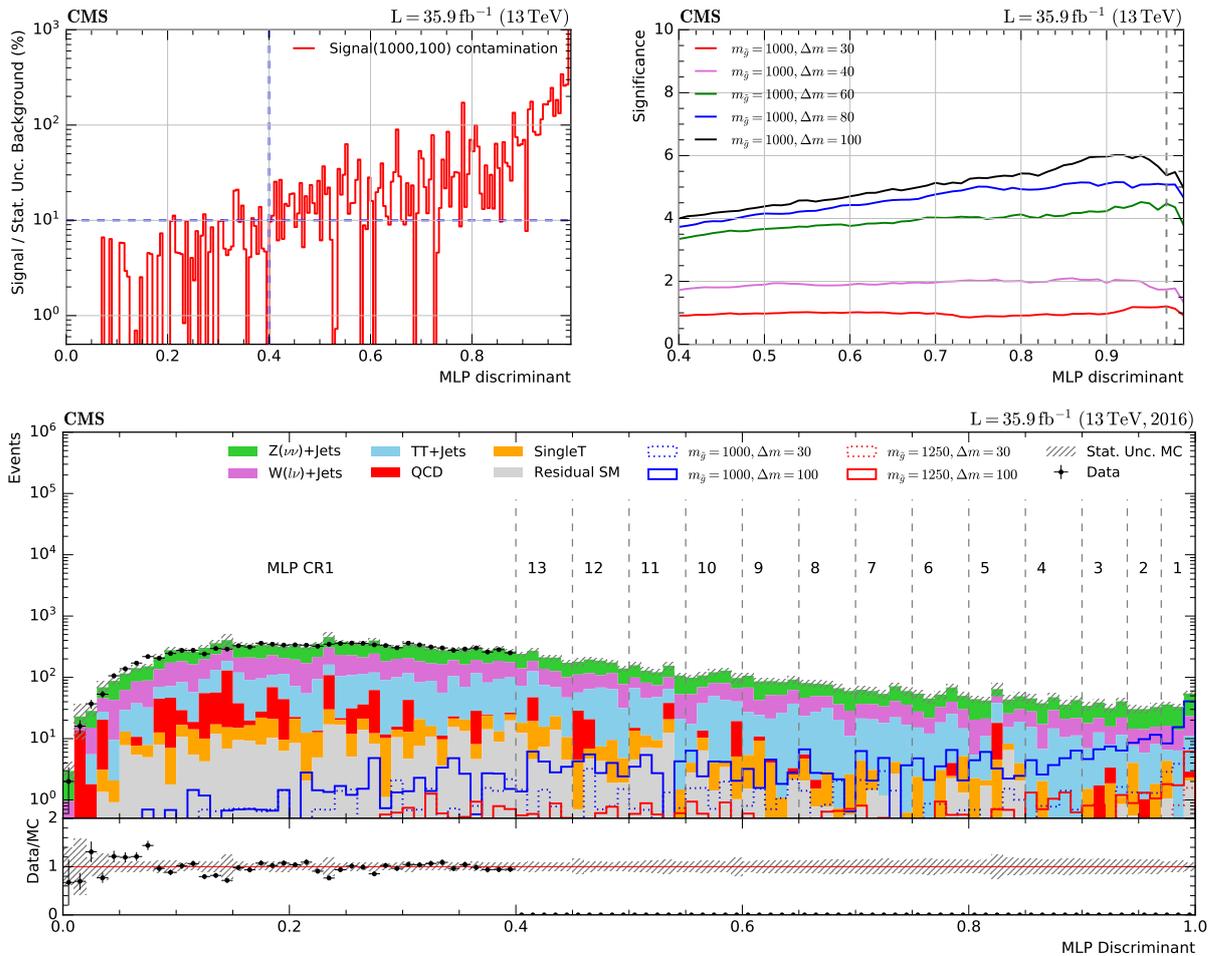


FIGURA 5.23 – Gráficos de R1: Histograma de controle da contaminação de sinal (acima à esquerda). Curva de significância dos modelos com $m_{\tilde{g}} = 1000$ GeV (acima à direita). Distribuição da variável discriminante correspondente à RNA com melhor performance (abaixo). A região de controle e os 13 *bins* de busca são delimitados pelas linhas verticais em 0,4, 0,45, 0,5, 0,55, 0,6, 0,65, 0,7, 0,75, 0,8, 0,85, 0,9, 0,94 e 0,97.

5.5.2 Análise multivariada em R2

- ★ **Objetivo:** Definir as regiões de sinal em R2;
- ★ **Região das amostras:** Eventos que estão na região de busca definida na seção 4.5, não possuem SV2Jets's e possuem dois ou mais SV1Jet's;
- ★ **Classe de sinal:** Mistura de eventos simulados de diferentes pontos do *grid* de sinal;
- ★ **Classe de *background*:** Eventos simulados dos processos do MP;
- ★ **Variáveis de entrada:** E_T^{miss} , H_T^{miss} , H_T , H_T^{PV} , $\text{Jet}_L^{\text{PV}} p_T$, $\text{Jet}_S^{\text{PV}} p_T$, N_{jets} , $N_{\text{jets}}^{\text{PV}}$, N_{bjets} , $\hat{\omega}_{\text{min}}$, χ_{min} , f_{max} , $\text{Jet}_1^{\text{SV}} p_T$, $\text{Jet}_2^{\text{SV}} p_T$, M_R , R^2 , $\text{SV}_1 \chi^2$, $\text{SV}_1 \text{NDF}$, $\text{SV}_2 \chi^2$ e $\text{SV}_2 \text{NDF}$.

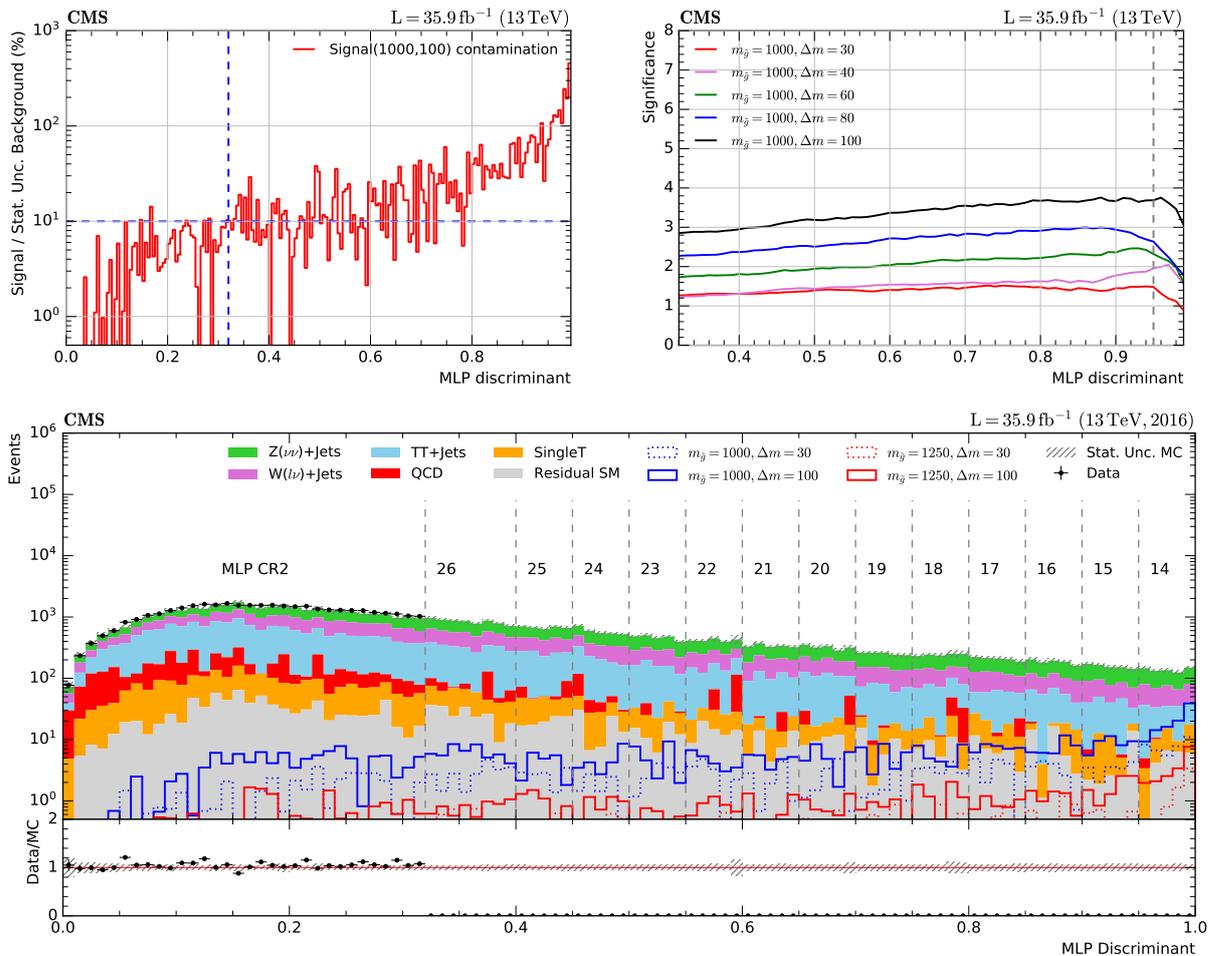


FIGURA 5.24 – Gráficos de R2: Histograma de controle da contaminação de sinal (acima à esquerda). Curva de significância dos modelos com $m_{\tilde{g}} = 1000 \text{ GeV}$ (acima à direita). Distribuição da variável discriminante correspondente à RNA com melhor performance (abaixo). A região de controle e os 13 *bins* de busca são delimitados pelas linhas verticais em 0,32, 0,4, 0,45, 0,5, 0,55, 0,6, 0,65, 0,7, 0,75, 0,8, 0,85, 0,9 e 0,95.

5.5.3 Análise multivariada em R3

- ★ **Objetivo:** Definir as regiões de sinal em R3;
- ★ **Região das amostras:** Eventos que estão na região de busca definida na seção 4.5, não possuem SV2Jets's e possuem um SV1Jet;
- ★ **Classe de sinal:** Mistura de eventos simulados de diferentes pontos do *grid* de sinal;
- ★ **Classe de *background*:** Eventos simulados dos processos do MP;
- ★ **Variáveis de entrada:** E_T^{miss} , H_T^{miss} , H_T , H_T^{PV} , $\text{Jet}_L^{\text{PV}} p_T$, $\text{Jet}_S^{\text{PV}} p_T$, N_{jets} , $N_{\text{jets}}^{\text{PV}}$, N_{bjets} , $\hat{\omega}_{\text{min}}$, χ_{min} , f_{max} , $\text{Jet}^{\text{SV}} p_T$, SV χ^2 e SV NDF.

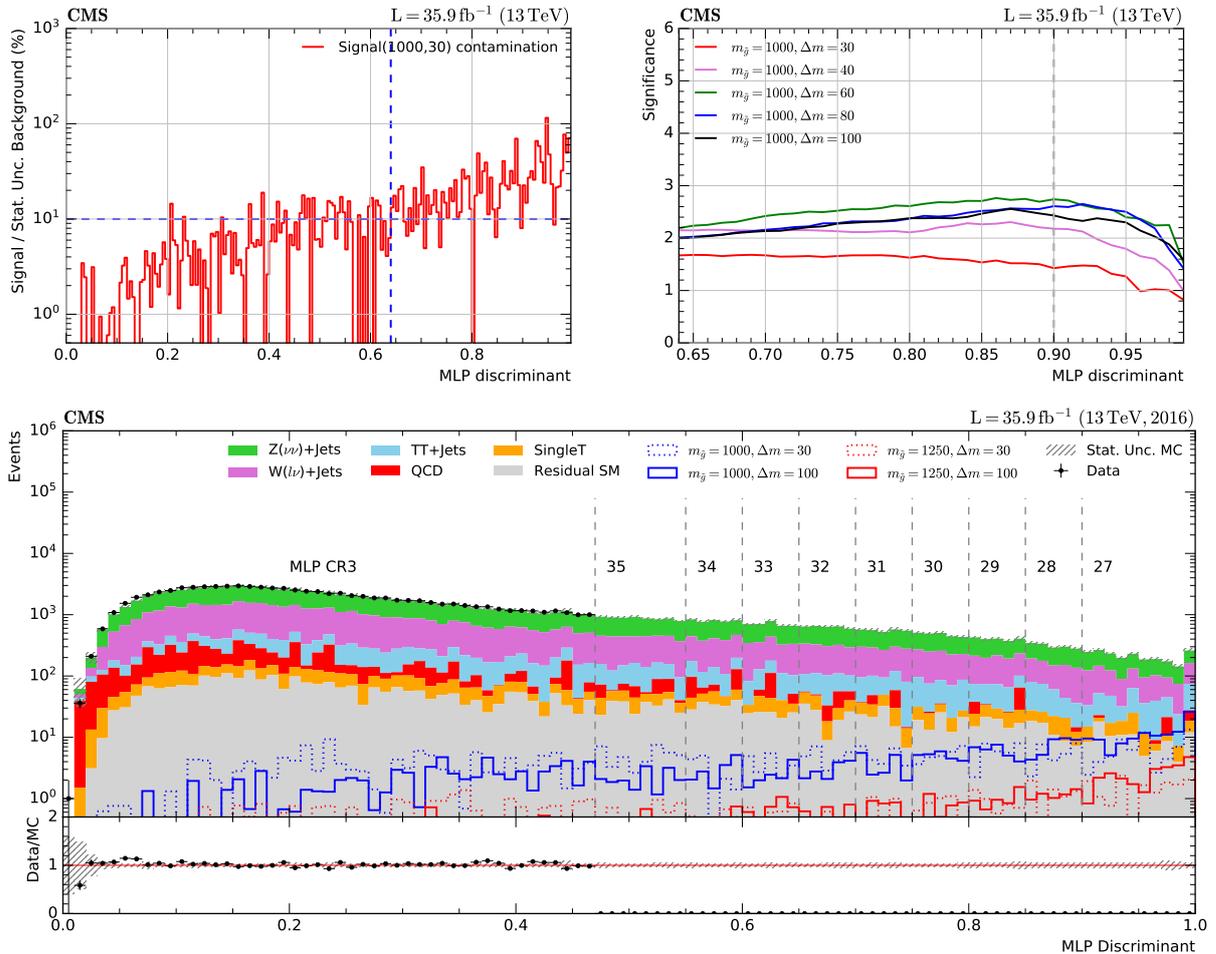


FIGURA 5.25 – Gráficos de R3: Histograma de controle da contaminação de sinal (acima à esquerda). Curva de significância dos modelos com $m_{\tilde{g}} = 1000 \text{ GeV}$ (acima à direita). Distribuição da variável discriminante correspondente à RNA com melhor performance (abaixo). A região de controle e os 9 *bins* de busca são delimitados pelas linhas verticais em 0,47, 0,55, 0,6, 0,65, 0,7, 0,75, 0,8, 0,85 e 0,9.

5.5.4 Análise multivariada em R4

- ★ **Objetivo:** Definir as regiões de sinal em R4;
- ★ **Região das amostras:** Eventos que estão na região de busca definida na seção 4.5, não possuem SV2Jets's nem SV1Jet's;
- ★ **Classe de sinal:** Mistura de eventos simulados de diferentes pontos do *grid* de sinal;
- ★ **Classe de *background*:** Eventos simulados dos processos do MP;
- ★ **Variáveis de entrada:** E_T^{miss} , H_T^{miss} , H_T , H_T^{PV} , Jet_L^{PV} p_T , Jet_S^{PV} p_T , N_{jets} , $N_{\text{jets}}^{\text{PV}}$, N_{bjets} , $\hat{\omega}_{\text{min}}$, χ_{min} e f_{max} .

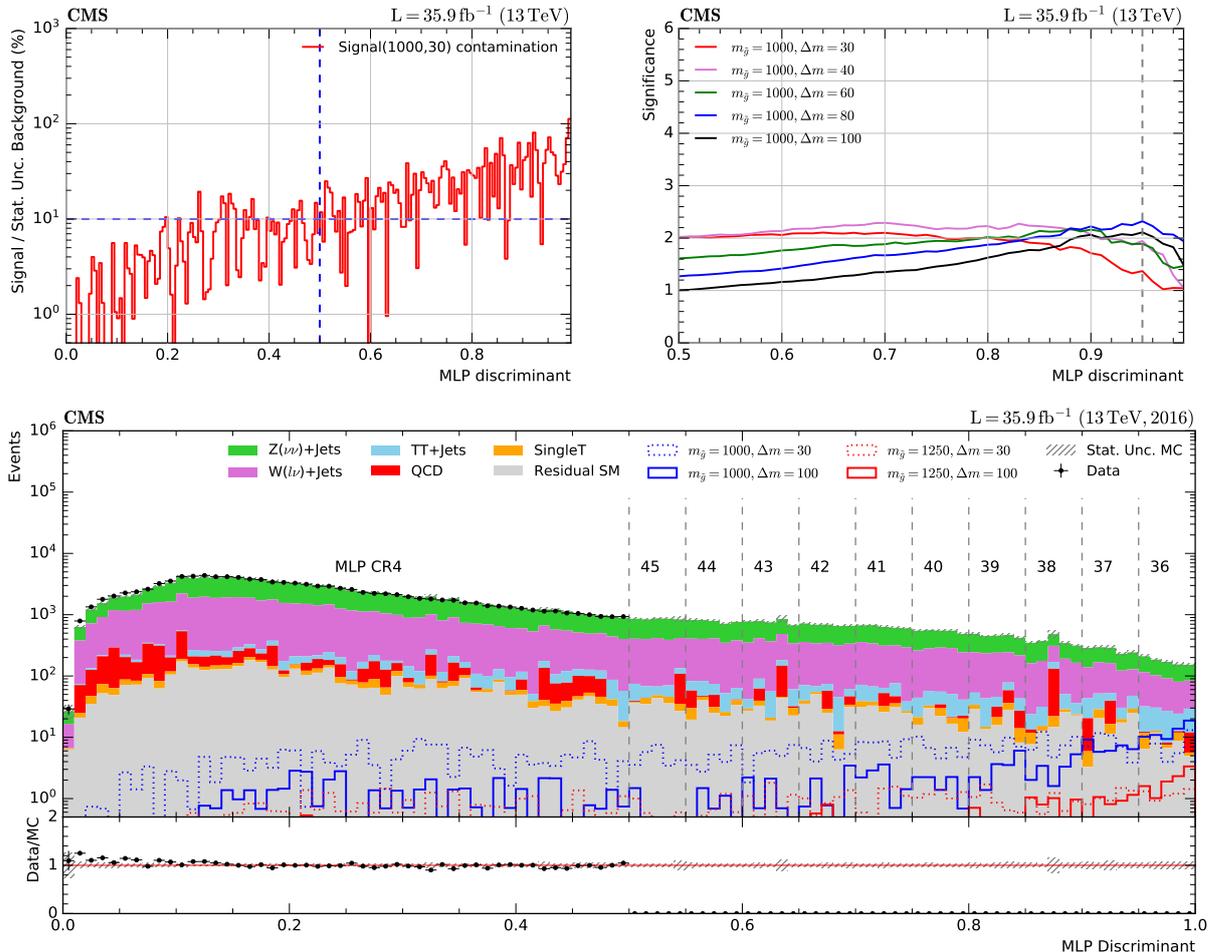


FIGURA 5.26 – Gráficos de R4: Histograma de controle da contaminação de sinal (acima à esquerda). Curva de significância dos modelos com $m_{\tilde{g}} = 1000$ GeV (acima à direita). Distribuição da variável discriminante correspondente à RNA com melhor performance (abaixo). A região de controle e os 10 *bins* de busca são delimitados pelas linhas verticais em 0,50, 0,55, 0,6, 0,65, 0,7, 0,75, 0,8, 0,85, 0,9 e 0,95.

6 Resultados

A análise multivariada do capítulo 5 definiu 45 *bins* de busca. A estimativa dos eventos de sinal e *background* em cada *bin* é mostrada na figura 6.1.

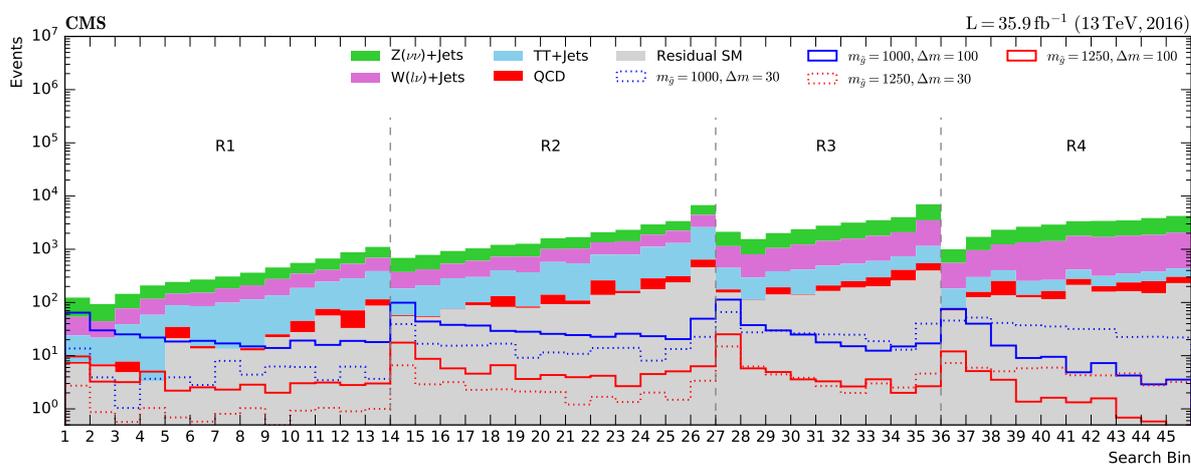


FIGURA 6.1 – Estimativa das contribuições de processos de *background* e sinal em cada *bin* de busca.

Estando a análise em andamento, as incertezas sistemáticas ainda estão sendo avaliadas. Neste capítulo, são descritas as incertezas sistemáticas da análise e exibidos os resultados das incertezas já avaliadas. Ao final, são mostrados os limites de exclusão esperados, considerando as incertezas sistemáticas já avaliadas.

6.1 Incertezas Sistemáticas

As incertezas sistemáticas aplicadas às amostras de sinal estão listadas abaixo. As incertezas já avaliadas têm seus valores exibidos.

- **Luminosidade:** incerteza plana¹ de 2,6% para a luminosidade integrada de 2016, estimada centralmente pelo CMS;

¹Uma incerteza plana possui o mesmo valor em todos os *bins de busca*

- **Trigger:** incerteza plana de 0,9% correspondente à estimativa da eficiência do *trigger* obtida na seção 4.4;
- **Estatística do MC:** incerteza estatística da amostra de MC em cada *bin* de busca. Valores entre 2,7-100%;
- **JER:** incerteza propagada para cada *bin* de busca relativa às correções realizadas na resolução da energia dos jatos;
- **JEC:** incerteza propagada para cada *bin* de busca relativa às correções realizadas na escala de energia dos jatos, variando o p_T dos jatos em ± 1 desvio padrão. Valores entre 1-63%;
- **Correções do MET:** incerteza propagada para cada *bin* de busca relativa às correções realizadas na energia transversa perdida, variando E_T^{miss} em ± 1 desvio padrão. Valores entre 2-70%;
- **Radiação de Estado Inicial:** incerteza dos fatores de correção do estado de radiação inicial, derivados de uma amostra de $t\bar{t}+\text{Jatos}$, conforme exposto na tabela 4.7. Essa correção é aplicada exclusivamente na amostra de $t\bar{t}+\text{Jatos}$ e nas amostras de sinal;
- **Pileup:** incerteza propagada para cada *bin* de busca relativa à modificação na forma da distribuição de interações verdadeiras dos dados ao variar a seção de choque inelástica total em ± 1 desvio padrão;
- **PDF:** a seção de choque depende do PDF conforme exposto na equação 3.4. Para avaliar o efeito nos resultados finais, PDF's alternativas são consideradas, seguindo a recomendação oficial PDF4LHC em [160];
- **Escalas de fatoração e renormalização:** incerteza calculada ao variar as escalas de normalização e fatoração por um fator de dois [161,162];
- **Marcador de b-jatos:** os fatores de escala dos marcadores de b-jatos são funções do p_T e η dos jatos. Sua incerteza sistemática é obtida variando os fatores de escala em ± 1 desvio padrão [163];
- **Algoritmo de associação:** são derivados fatores de escala para os objetos SV1Jet e SV2Jets. A incerteza sistemática é obtida variando os fatores de escala em ± 1 desvio padrão.

Entre as incertezas sistemáticas listadas para as amostras de sinal, quase todas também são aplicadas nas amostras de *background* dominantes, ou seja, $Z(\nu\nu)+\text{Jatos}$, $W(l\nu)+\text{Jatos}$ e $t\bar{t}+\text{Jatos}$. Isso ocorre porque suas estimativas também estão se baseando nas amostras

de MC. As exceções são a incerteza da luminosidade que não é aplicada aos *backgrounds* dominantes, visto que eles já são normalizados para concordar com a escala global dos dados, e a segunda é a incerteza da radiação de estado inicial que só é aplicada nas amostras de sinal e $t\bar{t}$ +Jatos. A incerteza da estatística de MC também é aplicada às contribuições residuais do MP, só não sendo considerada para os eventos de QCD, que serão estimados por um método baseado nos dados reais (*data-driven method*). A seguir, são listadas as incertezas sistemáticas para as amostras de *background* dominantes, baseadas em correções da simulação de MC.

- **Normalização:** incerteza combinada dos fatores de correção globais e regionais aplicados aos pesos dos eventos;
- **Forma:** incerteza dos fatores de correção de forma aplicados aos pesos dos eventos;
- **MLP discriminante:** incerteza sistemática associada ao grau de acordo entre as simulações e os dados reais com relação à variável discriminante nos *bins* de busca.

As contribuições residuais do MP não têm sua simulação corrigida, sendo atribuída uma incerteza plana de 50% para levar em conta possíveis imprecisões no modelamento desses processos. As figuras 6.2, 6.3 e 6.4 mostram as estimativas nominais e as incertezas sistemáticas já avaliadas e dependentes do *bin* de busca, para os eventos de $t\bar{t}$ +Jatos, $W(l\nu)$ +Jatos e $Z(\nu\nu)$ +Jatos, respectivamente.

Para o cálculo dos limites de exclusão esperados, na seção 6.2, é atribuída uma incerteza plana de 10% às amostras de sinal, para suprimir a falta de parte das incertezas sistemáticas; a estimativa do número de eventos de QCD é baseada na amostra de MC com uma incerteza plana de 100%; e a incerteza sistemática do MLP discriminante para os *backgrounds* dominantes é considerada como o erro relativo entre a quantidade de eventos reais e de *background* nas regiões de controle MLP CR1, CR2, CR3 e CR4. Uma abordagem mais precisa para a estimativa da incerteza do MLP discriminante é baseada na avaliação da distribuição da variável discriminante nas regiões de controle de cada *background* dominante. Entretanto, as regiões de controle CR1 e CR2 dos eventos de $Z(\nu\nu)$ +Jatos não possuem quantidade de eventos suficientes para essa avaliação. Uma alternativa é realizar essa avaliação utilizando uma amostra de $Z(\mu\mu)$ +Jatos, incorporando os momentos transversais dos múons ao cálculo da variável E_T^{miss} , numa região de controle com uma definição próxima à adotada na região de busca, mas com o requerimento de dois múons reconstruídos. Essa substituição é válida porque o comportamento cinemático com relação às variáveis hadrônicas é igual entre os processos $Z(\nu\nu)$ +Jatos e $Z(\mu\mu)$ +Jatos. Contudo, essa alternativa ainda está sendo avaliada.

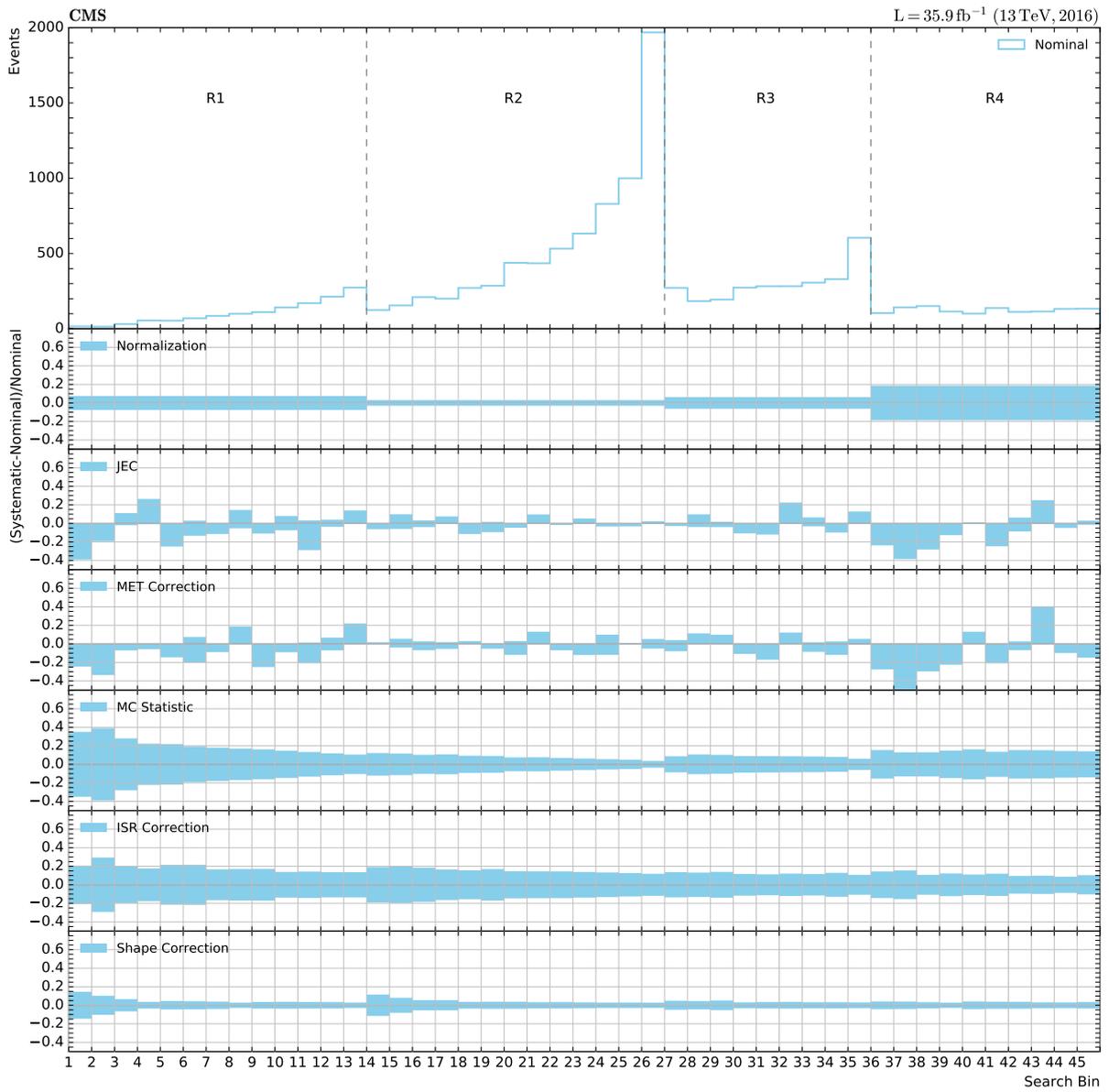


FIGURA 6.2 – Estimativa nominal do número de eventos de $t\bar{t} + \text{Jatos}$ nos *bins* de busca junto com algumas estimativas de suas incertezas sistemáticas.

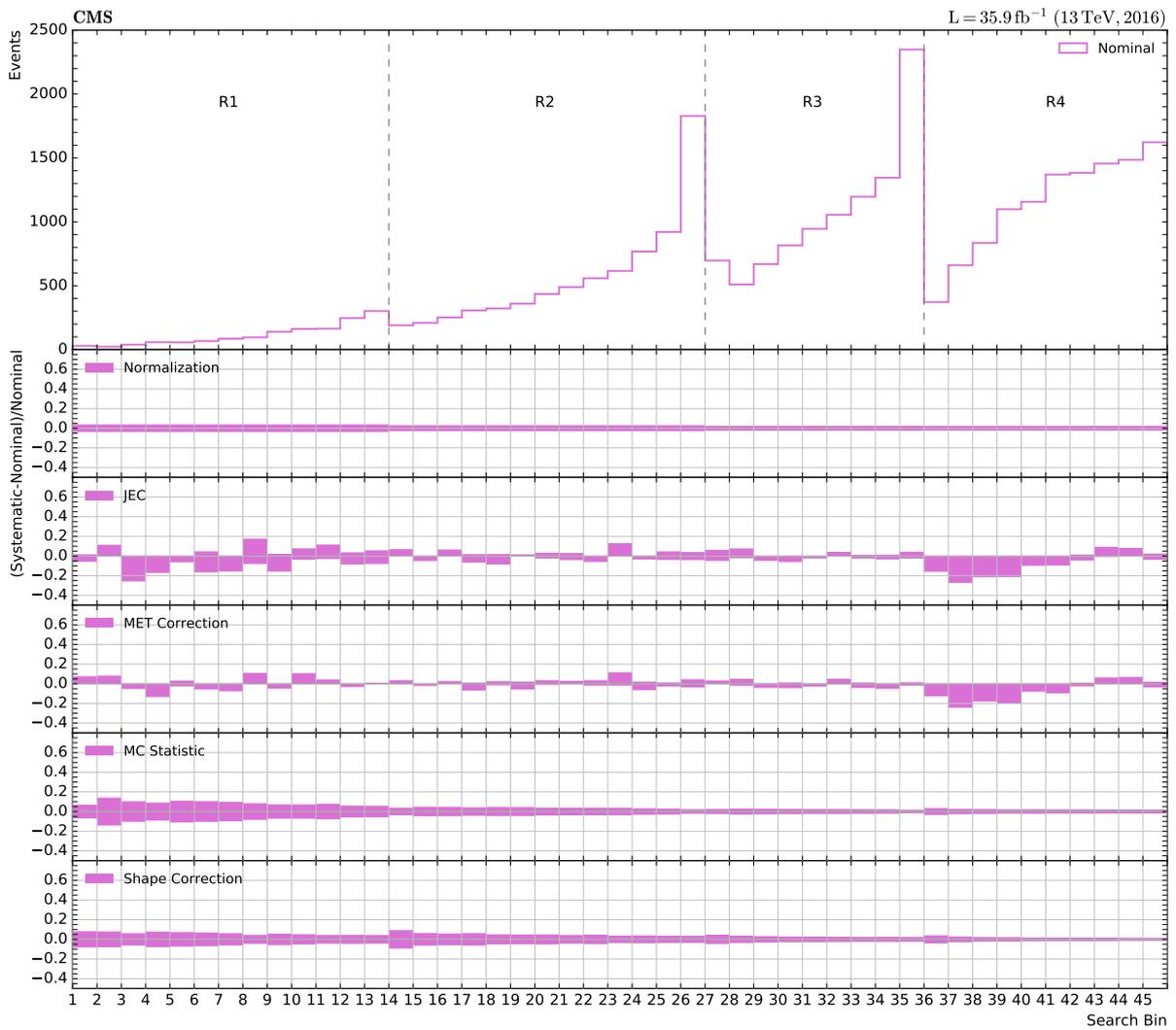


FIGURA 6.3 – Estimativa nominal do número de eventos de $W(l\nu)+\text{Jatos}$ nos *bins* de busca junto com algumas estimativas de suas incertezas sistemáticas.

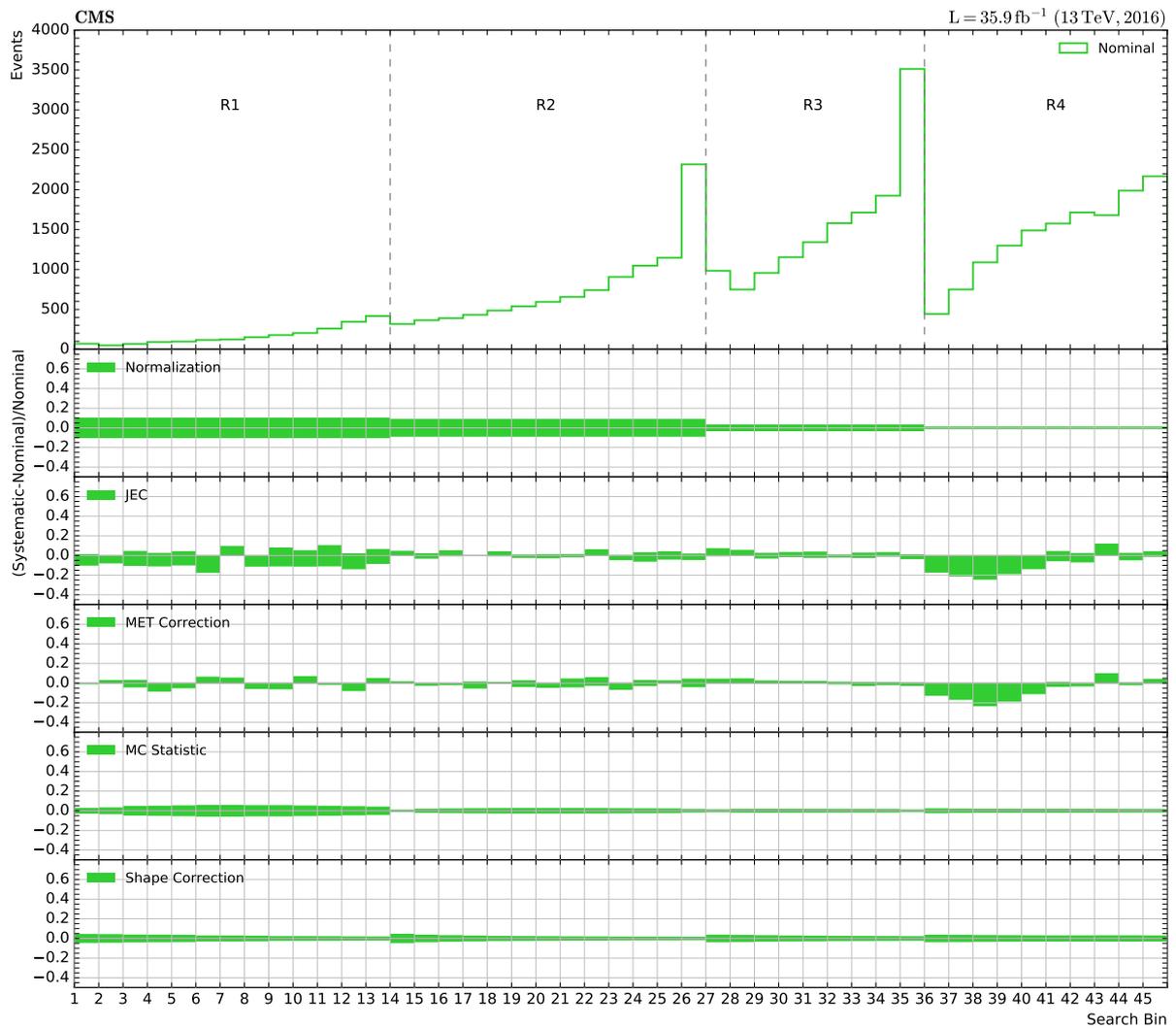


FIGURA 6.4 – Estimativa nominal do número de eventos de $Z(\nu\nu)+\text{Jatos}$ nos *bins* de busca junto com algumas estimativas de suas incertezas sistemáticas.

6.2 Limites de Exclusão

A estimativa dos limites de exclusão é realizada, em cada ponto do *grid* de sinal, utilizando o método de probabilidade de perfil (*profile likelihood method*) [164], que compara a compatibilidade dos dados com as hipóteses de ausência de sinal (apenas *background*) e de presença de sinal (sinal mais *background*), reduzindo o impacto das incertezas sistemáticas. Considera-se uma observação experimental especificada pelo número de entradas observadas $\mathbf{n} = \{n_1, n_2, \dots, n_N\}$ no histograma formado pelos N *bins* de busca. Presume-se que o número n_i de entradas na posição i seja distribuído de acordo com uma distribuição de Poisson e que o número esperado de eventos, E_i , seja dado por

$$E_i = \mu \cdot s_i + b_i, \quad (6.1)$$

onde s_i e b_i correspondem, respectivamente, ao número de eventos de sinal e de *background* esperados no *bin* i . O parâmetro de intensidade de sinal, μ , é definido como a razão entre a seção de choque observada, σ_{obs} , e a esperada na hipótese de presença de sinal, σ_{SUSY} :

$$\mu \equiv \frac{\sigma_{\text{obs}}}{\sigma_{\text{SUSY}}}. \quad (6.2)$$

Desse modo, a hipótese de ausência de sinal é representada por $\mu = 0$ enquanto $\mu = 1$ representa a taxa de sinal esperada, correspondente ao valor nominal da seção de choque de SUSY, σ_{SUSY} .

A análise estatística é baseada na função de probabilidade, \mathcal{L} , construída como o produto dos termos de probabilidade de Poisson de todos os N *bins* de busca. As incertezas sistemáticas nos valores esperados de s_i e b_i são modeladas por um conjunto de P parâmetros incômodos (*nuisance parameters*), $\theta = \{\theta_1, \theta_2, \dots, \theta_P\}$, de modo que o número total de eventos esperados em um determinado *bin* depende de μ e θ . Os parâmetros incômodos são implementados na função de probabilidade como formas funcionais, $\rho(\theta)$, que dependem do tipo de erro sistemático. Portanto a probabilidade total é escrita como:

$$\mathcal{L}(\mu, \theta) = \prod_{i=1}^N \frac{(\mu \cdot s_i(\theta) + b_i(\theta))^{n_i}}{n_i!} \exp(-\mu \cdot s_i(\theta) - b_i(\theta)) \cdot \prod_{k=1}^P \rho(\theta_k). \quad (6.3)$$

A criação de um teste baseado na função de probabilidade tem como objetivo a construção da estatística de teste q_μ [165]:

$$q_\mu = -2 \ln(\mathcal{L}_\mu / \mathcal{L}_{\text{max}}), \quad (6.4)$$

usada para obter os limites superiores de μ para um determinado nível de confiança. \mathcal{L}_{max} é a probabilidade máxima determinada pela variação dos parâmetros μ e θ , e \mathcal{L}_μ

é a probabilidade máxima para uma determinada intensidade de sinal μ . O cálculo dos limites é realizado pelo pacote computacional *Combine*, ferramenta padrão do CMS para estimativa dos limites de exclusão [166,167], que define os limites CL_s , níveis de confiança (*confidence levels*), assintóticos utilizando as fórmulas documentadas em [165].

A análise ainda está na fase cega (*blinded*), ou seja, não são exibidos os dados nos *bins* de busca. Logo, não é possível obter o limite de exclusão observado, porém é possível obter o limite de exclusão esperado e suas faixas de incerteza, utilizando as estimativas do número de processos de *background* e de sinal com suas incertezas sistemáticas consideradas em seção 6.1. O resultado preliminar, considerando parte das incertezas sistemáticas, é exibido na figura 6.5, que mostra a curva do limite de exclusão em função do Δm , para três valores distintos da massa do gluino. Sendo o limite superior em $\mu = \sigma/\sigma_{\text{SUSY}}$ o maior valor de μ não excluído, são excluídos todos os valores de μ acima da curva, com um nível de confiança de 95%. Sabendo-se que $\mu = 1$ representa a taxa de sinal esperada em cada modelo de sinal, são sensíveis ao experimento os modelos com limites de exclusão menores que 1. Esse é o caso dos modelos com $m_{\tilde{g}} \leq 750$ GeV. Nos modelos com $m_{\tilde{g}} = 1000$ GeV, o poder de exclusão da análise varia com o valor de Δm , eles são sensíveis mas se encontram no limite da sensibilidade para valores menores que aproximadamente 50 GeV. Por fim, a análise não apresenta sensibilidade suficiente para excluir os modelos com $m_{\tilde{g}} \geq 1250$ GeV devido ao baixo valor das suas seções de choque.

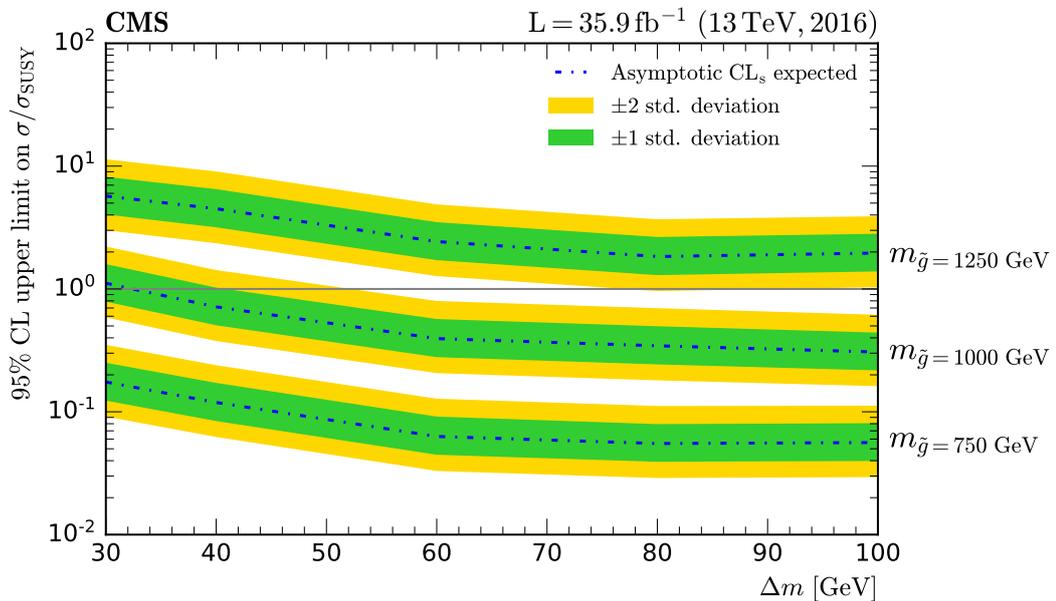


FIGURA 6.5 – Limites de exclusão esperados com 95% C.L. e suas faixas de incerteza para os modelos com massa do gluino igual a 750, 1000 e 1250 GeV.

7 Conclusões e Perspectivas Futuras

O autor se envolveu em todos os aspectos da análise, realizando o tratamento de dados necessário para a execução da análise multivariada e definição dos *bins* de busca. Tendo também avaliado as incertezas sistemáticas disponíveis no momento e obtido uma estimativa dos limites de exclusão. No decorrer da análise, também foi responsável por apresentar os avanços realizados nas reuniões do grupo de SUSY inclusiva do CMS. O autor ficou a cargo da manutenção do código de análise do grupo de pesquisa, escrito em C++, tendo contribuído com diversos melhoramentos, como a criação de um *script* em python para controlar toda a configuração e execução do código, a implementação de processamento paralelo e organização das saídas do código. Concomitantemente ao seu trabalho na análise, o autor também realizou um serviço como contato no banco de dados de alinhamentos e calibrações (*Alignment and Calibration Database - AlCaDB*) do HCAL no CMS, o qual é descrito no apêndice A.

A estrutura da análise foi construída com a definição dos cortes de pré-seleção e dos *triggers* de busca, estabelecimento da metodologia de modelagem dos *backgrounds* dominantes e definição das sub-regiões de busca, que se desdobram nos *bins* de busca com a realização da análise multivariada. Os limites de exclusão esperados mostram que a análise é sensível até massas do gluino entre 1000 e 1250 GeV. São esperadas pequenas modificações nessa expectativa ao serem consideradas todas as incertezas sistemáticas, serem utilizadas as amostras de sinal produzidas centralmente pelo CMS¹ e serem incorporados os dados do CMS armazenados em 2017 e 2018. A dependência da curva de exclusão com respeito a Δm mostra a diminuição gradativa da sensibilidade da análise para valores cada vez menores de Δm e mais acentuada abaixo do limiar em torno de 60 GeV. Isso é reflexo de uma das características principais do sinal buscado, que é a dificuldade cada vez maior na reconstrução dos jatos e SVs originários do decaimento dos gluinos em modelos cada vez mais comprimidos. Esta é a primeira análise sensível na região de alta compressão, $\Delta m \leq 100$ GeV, que leva em conta a propriedade de vida longa dos gluinos. A análise encaminha-se para ser finalizada durante o ano de 2020, fazendo

¹As amostras de sinal utilizadas foram produzidas pelo autor. O grupo do CMS, responsável por produzir as amostras de SUSY oficiais das análises, está produzindo as amostras de MC do sinal requisitadas para esta análise, com um número maior de eventos simulados.

uso completo dos dados armazenados durante o *run 2*, como foi estabelecido de comum acordo com os *conveners* de SUSY do CMS. Discussões sobre pontos específicos da análise e suas perspectivas futuras são realizadas a seguir.

- A definição dos *triggers* de busca ocorreu após muitas considerações e avaliação de todas as possibilidades disponíveis no experimento CMS, chegando-se à configuração de *triggers* de busca descrita na seção 4.4. Essa definição ocorreu recentemente e por isso também é recente a constatação do comportamento anômalo da eficiência dos *triggers* de busca com relação aos dados reais para valores grandes de E_T^{miss} . Esse comportamento será analisado com cuidado para determinar sua origem e, se possível e pertinente, eliminá-lo.
- O resultado obtido pelas redes neurais na análise multivariada depende das amostras de MC utilizadas. Portanto a análise multivariada deverá ser realizada novamente quando estiverem disponíveis as simulações produzidas centralmente dos *backgrounds* e do sinal, referentes aos anos de 2016, 2017 e 2018. Contudo, não é esperado que haja modificações bruscas nas estimativas dos *bins* de busca. Um observação que pode ser feita, é sobre a suposição de que a sub-região R1 seria a mais sensível ao sinal devido à maior quantidade de informação disponível no SV2Jets reconstruído. Com relação aos modelos com $\Delta m = 100 \text{ GeV}$, isso fica bem evidente comparando as curvas de significância entre as sub-regiões. Contudo, deve-se fazer a ressalva de que essa suposição não é necessariamente válida para modelos mais comprimidos, como no caso extremo em que $\Delta m = 30 \text{ GeV}$. Nesses casos de compressão extrema, existe uma supressão na capacidade de reconstrução de eventos com SV2Jets, por causa da baixa energia dos quarks produzidos no decaimento dos gluinos. Por isso, mesmo existindo um número maior de eventos de *background* em R2, existe também um aumento considerável do número de eventos de sinal com compressão extrema nessa sub-região, devido a maior facilidade na reconstrução de SV1Jet's. A reconstrução de dois SV1Jet's também fornece uma quantidade interessante de informação, o resultado é uma maior sensibilidade dos modelos com $\Delta m = 30 \text{ GeV}$ na sub-região R2 em comparação ao observado em R1.
- A estimativa dos eventos de QCD baseou-se no uso da amostra de MC, entretanto, além de tratar-se de um processo de difícil simulação, a estatística da amostra de QCD não é suficiente para que haja uma estimativa razoável do número de eventos nos *bins* de busca. O uso da amostra de MC continuará sendo usado na realização da análise multivariada, o que é aceitável pelo fato dos eventos QCD terem uma contribuição pequena com relação ao *background* total. Todavia, após definidos os *bins* de busca, a contribuição dos eventos de QCD em cada *bin* será estimada fazendo-se uso de um método baseado nos dados reais e não na simulação de Monte

Carlo.

- Por fim, serão avaliadas as incertezas sistemáticas faltantes referentes ao JER, estado de radiação inicial (para as amostras de sinal), *pileup*, PDF, escalas de fatoração e renormalização, marcador de b-jatos e algoritmo de associação. Além de estar sendo avaliada uma alternativa para a avaliação da incerteza sistemática associada ao MLP discriminante, como já foi mencionado.

Referências

- [1] CMS Collaboration, *The CMS experiment at the CERN LHC*, JINST **3** S08004 (2008).
- [2] CMS Collaboration, *Search for supersymmetry in multijet events with missing transverse momentum in proton-proton collisions at 13 TeV*, Phys. Rev. D **96**, 032003 (2017).
- [3] CMS Collaboration, *Search for new phenomena with the M_{T2} variable in the all-hadronic final state produced in proton-proton collisions at $\sqrt{s} = 13$ TeV*, Eur. Phys. J. C **77**, 710 (2017).
- [4] C. Campagnari, M. Franklin, *The Discovery of the top quark*, Rev. Mod. Phys. **69** 137-212 (1997).
- [5] DONUT Collaboration, *Observation of tau neutrino interactions*, Phys. Lett. B **504** 218-224 (2001).
- [6] CMS Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. B **716** (2012).
- [7] CMS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. B **716** (2012).
- [8] M. Tanabashi et al., (*Particle Data Group*), Phys. Rev. D **98**, 030001 (2018).
- [9] C. Burgard, *Example: Standard model of physics*, disponível em <http://www.texample.net/tikz/examples/model-physics/> Acesso em: 20 de maio de 2019.
- [10] C.N. Yang, R.L. Mills, *Conservation of Isotopic Spin and Isotopic Gauge Invariance*, Phys. Rev. **96**, 191-195 (1954).
- [11] R. Shaw, *Invariance under General Isotopic Spin Transformations*, PhD thesis, Cambridge University (1955).
- [12] S. Glashow, *Partial Symmetries of Weak Interactions*, Nucl. Phys. **22**, 579-588 (1961).

-
- [13] A. Salam, J.C. Ward, *Electromagnetic and weak interactions*, Phys. Lett. **13**, 168-171 (1964).
- [14] S. Weinberg, *A Model of Leptons*, Phys. Lett. **19**, 1264-1266 (1967).
- [15] T.D. Lee, C.N. Yang, *Question of Parity Conservation in Weak Interactions*, Phys. Rev. **104**, 254-258 (1956).
- [16] C.S. Wu, E. Ambler, R.W. Hayward, D.D. Hoppes, R.P. Hudson, *Experimental Test of Parity Conservation in Beta Decay*, Phys. Lett. **105**, 1413-1415 (1957).
- [17] UA1 Collaboration, *Experimental Observation of Isolated Large Transverse Energy Electrons with Associated Missing Energy at $s^{1/2} = 540$ GeV*, Phys. Lett. **122**, 103-116 (1983).
- [18] UA2 Collaboration, *Observation of Single Isolated Electrons of High Transverse Momentum in Events with Missing Transverse Energy at the CERN anti-p p Collider*, Phys. Lett. **122**, 476-485 (1983).
- [19] UA1 Collaboration, *Experimental Observation of Lepton Pairs of Invariant Mass Around 95 GeV/c² at the CERN SPS Collider*, Phys. Lett. **126**, 398-410 (1983).
- [20] UA2 Collaboration, *Evidence for $Z^0 \rightarrow e + e^-$ at the CERN anti-p p Collider*, Phys. Lett. **129**, 130-140 (1983).
- [21] F. Englert, R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, Phys. Rev. Lett. **13**, 321-323 (1964).
- [22] G. Guralnik, C. Hagen, T. Kibble, *Global Conservation Laws and Massless Particles*, Phys. Rev. Lett. **13**, 585-587 (1964).
- [23] P.W. Higgs, *Broken symmetries, massless particles and gauge fields*, Phys. Lett. **12**, 132-133 (1964).
- [24] P.W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, Phys. Rev. Lett. **13**, 508-509 (1964).
- [25] LHCb Collaboration, *Observation of $J/\psi p$ Resonances Consistent with Pentaquark States in $\Lambda_b^0 \rightarrow J/\psi K^- p$ Decays*, Phys. Rev. Lett. **115**, 072001 (2015).
- [26] LHCb Collaboration, *The LHCb Detector at the LHC*, JINST **3** S08005 (2008).
- [27] G. Bertone, D. Hooper, J. Silk, *Particle dark matter: Evidence, candidates and constraints*, Phys. Rept. **405** 279-390 (2005).
- [28] D.B. Cline, *Sources of Dark Matter in the Universe: Proceedings, 1st International Symposium, February 16-18, 1994, Bel Air, CA*, World Scientific, (1995).
- [29] W.J.G. de Blok, S.S. McGaugh, A. Bosma et al., *Mass density profiles of LSB galaxies*, Astrophys. J. **552** L23-L26 (2001).

- [30] L.V.E. Koopmans, T. Treu, *The structure and dynamics of luminous and dark matter in the early-type lens galaxy of 0047-281 at $z=0.485$* , *Astrophys. J.* **583** 606-615 (2003).
- [31] H. Hoekstra, H. Yee, M. Gladders, *Current status of weak gravitational lensing*, *New Astron. Rev.* **46** 767-781 (2002).
- [32] Planck Collaboration, *Planck 2013 results. XVI. Cosmological parameters*, *Astron. Astrophys.* **571** A16 (2014).
- [33] E. Kolb, M. Turner, *The Early Universe*, Avalon Publishing (1994).
- [34] L. Bian, *Renormalization group equation, the naturalness problem, and the understanding of the Higgs mass term*, *Phys. Rev.* **D88** (2013).
- [35] S. Coleman, J. Mandula, *All Possible Symmetries of the S-Matrix*, *Phys. Rev.* **159**, 1251 (1967).
- [36] R. Haag, J.T. Lopuszanski, M.F. Sohnius, *All Possible Generators of Supersymmetries of the S-Matrix*, *Nucl. Phys.* **B88** 257 (1975).
- [37] H.J.W. Müller-Kirsten, A. Wiedemann, *Introduction to Supersymmetry*, World Scientific, (2010).
- [38] I.J.R. Aitchison, *Supersymmetry and the MSSM: An Elementary Introduction*, Cambridge University Press, (2007).
- [39] H. Murayama, *Supersymmetry Phenomenology*, arXiv:0002232 (2000).
- [40] C. Csáki, *The Minimal Supersymmetric Standard Model (MSSM)*, arXiv:9606414 (1996).
- [41] R.L. Arnowitt, P. Nath, *SUSY mass spectrum in $SU(5)$ supergravity grand unification*, *Phys. Rev. Lett.* **69**, 725-728 (1992).
- [42] G.L. Kane, C.F. Kolda, L. Roszkowski, J.D. Wells, *Study of constrained minimal supersymmetry*, *Phys. Rev.* **D49**, 6173-6210 (1994).
- [43] G. Giudice, R. Rattazzi, *Theories with gauge mediated supersymmetry breaking*, *Phys. Rept.* **322**, 419-499 (1999).
- [44] B.C. Allanach et al., *The Snowmass Points and Slopes: Benchmarks for SUSY Searches*, *Eur. Phys. J. C* **25**, 113-123 (2002).
- [45] S. Weinberg, *Implications of Dynamical Symmetry Breaking*, *Phys. Rev.* **D13**, 974-996 (1976).
- [46] G. Bhattacharyya, *Electroweak Symmetry Breaking Beyond the Standard Model*, *Pramana - J Phys.* **79**, 675 (2012).

- [47] D.I. Kazakov, *Beyond the Standard Model (In Search of Supersymmetry)*, arXiv:hep-ph/0012288 (2000).
- [48] E. Witten, *Dynamical breaking of supersymmetry*, Nucl. Phys. **B188**, 513 (1981).
- [49] S. Dimopoulos, S. Raby, F. Wilczek, *Supersymmetry and the scale of unification*, Phys. Rev. D **24**, 1681 (1981).
- [50] M.L. Alciati, F. Feruglio, Y. Lin, A. Varagnolo, *Proton lifetime from $SU(5)$ unification in extra dimensions*, J. High Energy Phys. 03, 054 (2005).
- [51] C. Amsler et al. (Particle Data Group), *Review of Particle Physics*, Phys. Lett. B **667**, 1 (2008).
- [52] S. Krippendorff, H.P. Nilles, M. Ratz, M.W. Winkler, *Hidden SUSY from precision gauge unification*, Phys. Rev. D **88**, 035022 (2013).
- [53] K. Choi, A. Falkowski, H. P. Nilles, M. Olechowski, *Soft supersymmetry breaking in KKL T flux compactification*, Nucl. Phys. **B718**, 113 (2005).
- [54] V. Löwen and H.P. Nilles, *Mirage pattern from the heterotic string*, Phys. Rev. D **77**, 106007 (2008).
- [55] B.S. Acharya, G. Kane, P. Kumar, *Compactified string theories - generic predictions for particle physics*, Int. J. Mod. Phys. **A27**, 1230012 (2012).
- [56] H.E. Haber, G.L. Kane, *Gluino decays and experimental signatures*, Nucl. Phys. **B232**, 333 (1984).
- [57] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** S08003 (2008).
- [58] ATLAS Collaboration, *Search for squarks and gluinos with the ATLAS detector using final states with jets and missing transverse momentum at $\sqrt{s} = 8$ TeV*, ATLAS-CONF-2012-109 (2012).
- [59] CMS Collaboration, *Search for gluino mediated bottom- and top-squark production in multijet final states in pp collisions at 8 TeV*, Phys. Lett **725**, 243-270 (2013)
- [60] M. Mühlleitner, A. Djouadi, Y. Mambrini, *SDECAY: a Fortran code for the decays of the supersymmetric particles in the MSSM*, Comput. Phys. Commun. **168**, 46 (2005).
- [61] R. Barbieri, G. Gamberini, G.F. Giudice, G. Ridolfi, *Constraining supergravity models from gluino production*, Nucl. Phys. **B301**, 15 (1988).
- [62] N. Craig, *The State of Supersymmetry after Run I of the LHC*, arXiv/hep-ph:1309.0528 (2013).

- [63] LHC New Physics Working Group Collaboration, *Simplified Models for LHC New Physics Searches*, J. Phys. **G39**, (2012).
- [64] D.S.M. Alves, E. Izaguirre, J.G. Wacker, *Where the Sidewalk Ends: Jets and Missing Energy Search Strategies for the 7 TeV LHC*, JHEP **10**, 012 (2011).
- [65] M.E. Cabrera, J.A. Casas, *Understanding and improving the Effective Mass for LHC searches*, arXiv/hep-ph:1207.0435 (2012).
- [66] CMS Collaboration, *Performance of the CMS missing transverse momentum reconstruction in pp data at $\sqrt{s} = 8$ TeV*, JINST **10**, no. 02, P02006 (2015).
- [67] C.G. Lester, D.J. Summers, *Measuring masses of semiinvisibly decaying particles pair produced at hadron colliders*, Phys. Lett. *B463*, 99-103 (1999).
- [68] A. Barr, C. Lester, P. Stephens, *$m(T2)$: The Truth behind the glamour*, J. Phys. **G29**, 2343-2363 (2003).
- [69] L. Randall, D. Tucker-Smith, *Dijet Searches for Supersymmetry at the LHC*, Phys. Rev. Lett. **101** 221803 (2008).
- [70] CMS Collaboration, *Search strategy for exclusive multi-jet events from supersymmetry at CMS*, CMS-PAS-SUS-09-001 (2009).
- [71] C. Rogan, *Kinematical variables towards new dynamics at the LHC*, arXiv/hep-ph:1006.2727 (2010).
- [72] CMS Collaboration, *Inclusive search for squarks and gluinos in pp collisions at $\sqrt{s} = 7$ TeV*, Phys. Rev. **D85**, 012004 (2012).
- [73] T.J. LeCompte, S.P. Martin, *Large Hadron Collider reach for supersymmetric models with compressed mass spectra*, Phys. Rev. **D84**, 015004 (2011).
- [74] E. Halkiadakis, G. Redlinger, D. Shih, *Status and implications of beyond-the-standard-model searches at the LHC*, Annual Review of Nuclear and Particle Science **64**, 319-342 (2014).
- [75] CMS Collaboration, *CMS Supersymmetry Physics Results*, disponível em <<https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsSUS>> Acesso em: 15 de março de 2019.
- [76] CMS Collaboration, LHCb Collaboration, *Observation of the rare $B_s^0 \rightarrow \mu^+ \mu^-$ decay from the combined analysis of CMS and LHCb data*, Nature **522**, 68-72 (2015).
- [77] ATLAS Collaboration, *Summary plots from the ATLAS Supersymmetry physics group*, disponível em <<http://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/-CombinedSummaryPlots/SUSY/>> Acesso em: 07 de Agosto de 2019.

- [78] H. Baer, C.-h. Chen, F. Paige et al., *Signals for minimal supergravity at the CERN large hadron collider: Multi - jet plus missing energy channel*, Phys. Rev. D **52**, 2746-2759 (1995).
- [79] ALICE Collaboration, *The ALICE experiment at the CERN LHC*, JINST **3** S08002 (2008).
- [80] LEP Collaboration, *LEP Design Report*, Geneva:CERN, 70 (1983).
- [81] M. Benedikt et al., *LHC Design Report*, Geneva:CERN, 356 (2004).
- [82] UKRI, *CERN Accelerator complex*, disponível em <<http://www.stfc.ac.uk/research/particle-physics-and-particle-astronomy/large-hadron-collider/cern-accelerator-complex/>> Acesso em: 23 de junho de 2019.
- [83] CMS Collaboration, *CMS Luminosity - Public Results*, disponível em <https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults#2010-proton-proton_7_TeV_collisi> Acesso em: 06 de agosto de 2019.
- [84] LHC Collaboration, *Longer term LHC schedule*, disponível em <<https://lhc-commissioning.web.cern.ch/lhc-commissioning/schedule/LHC-long-term.htm>> Acesso em: 06 de agosto de 2019.
- [85] G. Altarelli, *Collider Physics within the Standard Model: a Primer*, arXiv/hep-ph:1303.2842 (2013).
- [86] J.M. Butterworth, G. Dissertori, G.P. Salam, *Hard Processes in Proton-Proton Collisions at the Large Hadron Collider*, Ann. Rev. of Nucl. Part. Sci. **62**, 387-405 (2012).
- [87] CMS Collaboration, *SketchUpCMS Gallery*, disponível em <<https://twiki.cern.ch/twiki/bin/view/CMSPublic/SketchUpCMSGallery>> Acesso em: 08 de julho de 2018.
- [88] CMS Collaboration, *Description and performance of track and primary-vertex reconstruction with the CMS tracker*, JINST **9**, no. 10 P10009 (2014).
- [89] B. Martin, G. Shaw, *Particle Physics*, Wiley, 2008.
- [90] CMS Collaboration, *The Phase-1 Upgrade of the CMS pixel detector*, CMS-CR-2016-036 (2016).
- [91] C.E. Wulz, *The CMS experiment at CERN, Congress on Optics and Optoelectronics*, International Society for Optics and Photonics (2005).
- [92] CMS Collaboration, *The CMS electromagnetic calorimeter project : Technical Design Report*, Geneva:CERN, (1997).

- [93] D. Sprenger, M. Weber, R. Adolphi et al., *Validation of Kalman Filter alignment algorithm with cosmic-ray data using a CMS silicon strip tracker endcap*, JINST **5**, P06007 (2010).
- [94] CMS collaboration, *Technical Design Report Volume I: Detector Performance and Software*, Geneva:CERN, (2006).
- [95] CMS Collaboration, *The CMS hadron calorimeter project: Technical Design Report*, Geneva:CERN, (1997).
- [96] S. Chatrchyan et al., *The CMS experiment at the CERN LHC*, Journal of Instrumentation **3** (2008).
- [97] CMS Collaboration, *The CMS magnet project: Technical Design Report*, Geneva:CERN, 327 (1997).
- [98] CMS Collaboration, . *The CMS muon project: Technical Design Report*, Geneva:CERN, (1997).
- [99] M.S. Kim, *CMS reconstruction improvement for the muon tracking by the RPC chambers*, Journal of Instrumentation **8** (2013).
- [100] CMS Collaboration, *Performance of the CMS Drift Tube Chambers with Cosmic Rays*, JINST **5** T03015 (2010).
- [101] CMS Collaboration, *Performance of the CMS Cathode Strip Chambers with Cosmic Rays*, JINST **5** T03018 (2010).
- [102] CMS Collaboration, *The CMS trigger system. The CMS trigger system*, JINST **12** P01020 (2016).
- [103] CMS Collaboration, *Technical design report, vol. 1: The trigger systems*, Geneva:CERN, (2000).
- [104] CMS Collaboration, *Technical design report, Vol. 2: Data acquisition and high-level trigger*, Geneva:CERN, (2002).
- [105] M. Hildreth, *Pileup Tutorial*, disponível em https://twiki.cern.ch/twiki/pub/CMSPublic/WorkBookExercisesHATS_Pileup_-2013/Pileup_Overview_HATS13.pdf Acesso em: 10 de agosto de 2019.
- [106] M. Hildreth, *Pileup in Monte Carlo*, disponível em <https://twiki.cern.ch/twiki/bin/view/CMSPublic/PublicPileupInfo> Acesso em: 10 de agosto de 2019.
- [107] P. Luja, *Utilities for Accessing Pileup Information for Data*, disponível em <https://twiki.cern.ch/twiki/bin/viewauth/CMS/PileupJSONFileforData> Acesso em: 10 de agosto de 2019.

- [108] A. Popov, *Pileup Reweighting Utilities*, disponível em <<https://twiki.cern.ch/twiki/bin/view/CMS/PileupMCReweightingUtilities>> Acesso em: 10 de agosto de 2019.
- [109] CMS Collaboration, *Particle-flow event reconstruction in CMS and performance for jets, taus and MET*, CMS-PAS-PFT-09-001 (2009).
- [110] H. Kirschenmann, *Jets at CMS and the determination of their energy scale*, disponível em <<http://cms.web.cern.ch/news/jets-cms-and-determination-their-energy-scale>> Acesso em: 30 de abril de 2019.
- [111] R. Atkin, *Review of jet reconstruction algorithms*, Journal of Physics: Conference Series **645** (2015).
- [112] M. Cacciari, G.P. Salam, G. Soyez, *The anti-kt jet clustering algorithm*, Journal of High Energy Physics **2008.04** (2008).
- [113] CMS Collaboration, *Jet energy scale and resolution performance with 13 TeV data collected by CMS in 2016*, CMS-DP-2018-028 (2018).
- [114] CMS Collaboration, *Introduction to Jet Energy Corrections at CMS*, disponível em <<https://twiki.cern.ch/twiki/bin/viewauth/CMS/IntroToJEC>> Acesso em: 03 de maio de 2019.
- [115] CMS Collaboration, *Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV*, JINST **12**, P02014 (2017).
- [116] CMS Collaboration, *Determination of jet energy calibration and transverse momentum resolution in CMS*, JINST **6**, (2011).
- [117] CMS Collaboration, *Jet Energy Resolution*, disponível em <<https://twiki.cern.ch/twiki/bin/viewauth/CMS/JetResolution>> Acesso em: 03 de maio de 2019.
- [118] CMS Collaboration, *MET Analysis*, disponível em <<https://twiki.cern.ch/twiki/bin/view/CMSPublic/WorkBookMetAnalysis>> Acesso em: 05 de maio de 2019.
- [119] CMS Collaboration, *Performance of the CMS missing transverse momentum reconstruction in pp data at $\sqrt{s} = 8$ TeV*, JINST **10**, P02006 (2015).
- [120] CMS Collaboration, *MET Filter Recommendations for Run II*, disponível em <<https://twiki.cern.ch/twiki/bin/view/CMS/MissingETOptionalFiltersRun2>> Acesso em: 05 de maio de 2019.
- [121] CMS Collaboration, *Missing Transverse Energy CMS*, disponível em <https://indico.cern.ch/event/726639/contributions/2990361/attachments/1646504/2668752/MET_HATS_ss2018.pdf> Acesso em: 05 de maio de 2019.

- [122] A.B. Galtieri, F. Margaroli, I. Volobouev, *Precision measurements of the top quark mass from the Tevatron in the pre-LHC era*, Reports on Progress in Physics 75(5), 056201 (2012).
- [123] P. Vanlaer et al., *Impact of Tracker Misalignment on Track and Vertex Reconstruction*, CMS Note 2006/xxx, (2006).
- [124] T. Muller et al., *Inclusive secondary vertex reconstruction in jets*, CMS Note 2006/xxx, (2006).
- [125] T. Speer et al., *Vertex Fitting in the CMS Tracker*, CMS Note 2006/xxx, (2006).
- [126] CMS Collaboration, *Identification of b quark jets at the CMS Experiment in the LHC Run 2*, CMS-PAS-BTV-15-001 (2016).
- [127] CMS Collaboration, *CMS: the overall concept*, CERN/LHCC 7-24 (2006). CMS TDR 8.1.
- [128] A. Datta, B. Mukhopadhyaya, A. Raychaudhuri, *Physics at the Large Hadron Collider*, Springer, (2009).
- [129] CMS Collaboration, *The CMS trigger system*, JINST **12** P01020 (2017).
- [130] CMS Collaboration, *Jet Identification for the 13 TeV data Run2016*, disponível em <<https://twiki.cern.ch/twiki/bin/viewauth/CMS/JetID13TeVRun2016>> Acesso em: 05 de maio de 2019.
- [131] CMS Collaboration, *Search for supersymmetry in the all-hadronic final state using top quark tagging in pp collisions at $\sqrt{s} = 13$ TeV*, Phys. Rev. D **96**, 012004 (2017).
- [132] T. Sakuma, H. Flaecher, D. Smith, *Alternative angular variables for QCD multijet background event suppression in SUSY searches at the LHC*, disponível em <<https://indico.tifr.res.in/indico/getFile.py/access?contribId=282&sessionId=13&resId=0&materialId=slides&confId=5736>> Acesso em: 13 de fevereiro de 2019.
- [133] T. Sakuma, H. Flaecher, D. Smith, *Alternative angular variables for suppression of QCD multijet events in new physics searches with missing transverse momentum at the LHC*, arXiv/hep-ph:1803.07942 (2019).
- [134] CMS Collaboration, *Reconstruction of electrons with the Gaussian-sum filter in the CMS tracker at LHC*, J. Phys. G **31**, N9 (2005).
- [135] M.F. Sevilla, A. Ovcharova, *Preliminary ISR reweighting and uncertainty recommendations for Moriond17*, disponível em <https://indico.cern.ch/event/592621/contributions/2398559/attachments/1383909/2105089/16-12-05_ana_manuelf_isr.pdf> Acesso em: 10 de maio de 2019.
- [136] CMS Collaboration, *Search for supersymmetry with razor variables in pp collisions at $\sqrt{s} = 7$ TeV*, Phys. Rev. D **90**, 112001 (2014).

- [137] CMS Collaboration, *Inclusive search for supersymmetry in pp collisions at $\sqrt{s} = 13$ TeV using razor variables and boosted object identification in zero and one lepton final states*, JHEP **03**, 031 (2019).
- [138] A. Coghlan, *A Little Book of R for Multivariate Analysis*, Cambridge, 2017.
- [139] P. Cohen, S.G. West, L.S. Aiken, *Applied multiple regression/correlation analysis for the behavioral sciences*, Psychology Press, 2014.
- [140] A. Samuel, *Some Studies in Machine Learning Using the Game of Checkers*, IBM Journal of Research and Development **3**, 210-229 (1959).
- [141] I.N. Silva, D.H. Spatti, R.A. Flauzino, *Redes Neurais Artificiais para engenharia e ciências aplicadas*, Artliber Editora, 2010.
- [142] A. Hoecker, P. Speckmayer, J. Stelzer, J. Therhaag, E. von Toerne, and H. Voss, *TMVA: Toolkit for Multivariate Data Analysis*, PoS A CAT 040 (2007) [physics/0703039].
- [143] P.C. Bhat, *Multivariate Analysis Methods in Particle Physics*, Annual Review of Nuclear and Particle Science **61**, 281-309 (2011).
- [144] B.P. Roe, H.J. Yang, J. Zhu, I. Stancu, G. McGregor, *Boosted decision trees as an alternative to artificial neural networks for particle identification*, Nuclear Instruments and Methods in Physics Research A **543**, 577-584 (2005).
- [145] L.N. Castro, *Fundamentals of Natural Computing: Basic Concepts, Algorithms, and Applications*, CRC Press, (2006).
- [146] S. Haykin, *Neural Networks and Learning Machines*, Prentice Hall, (2008).
- [147] W.S. McCulloch, W. Pitts, *A logical calculus of the ideas immanent in nervous activity*, The bulletin of mathematical biophysics **5**, 115-133 (1943).
- [148] D.O. Hebb, *The organization of behavior: a neuropsychological theory*, Wiley, New York, USA (1949).
- [149] F. Rosenblatt, *The perceptron: a probabilistic model for information storage and organization in the brain*, Psychological Review **65**, 386-408 (1958).
- [150] D.E. Rumelhart, G.E. Hilton, R.J. Williams, *Learning internal representation by error propagation*, Parallel Distributed Processing **1**, 318-362 (1986).
- [151] B.B. Benuwa et al., *A Review of Deep Machine Learning*, International Journal of Engineering Research in Africa **24**, 124-136 (2016).
- [152] A. Krizhevsky, I. Sutskever, G.E. Hinton, *Imagenet classification with deep convolutional neural networks*, Advances in neural information processing systems, 1097-1105 (2012).

- [153] A. Karpathy, *CS231n Convolutional Neural Networks for Visual Recognition*, disponível em <<http://cs231n.github.io/neural-networks-1/>> Acesso em: 12 de julho de 2019.
- [154] A.L. Hodgkin, A.F. Huxley, *A quantitative description of membrane current and its application to conduction and excitation in nerve*, *Journal of Physiology* **117**, 500-544 (1952).
- [155] E.R. Kandel, J.H. Schwartz, T.M. Jenseil, *Principles of neural science*, McGraw-Hill, New York, USA (2000).
- [156] B. Widrow, M.E. Hoff, *Adaptative switching circuits*, *Proceedings of the IRE Wescon Convention Record*, 96-104 (1960).
- [157] D.P. Kingma, J. Ba, *Adam: A Method for Stochastic Optimization*, arXiv/cs:1412.6980 (2017).
- [158] F. Chollet et al., *Keras*, disponível em <<https://keras.io>> Acesso em: 15 de julho de 2019.
- [159] C. Bini, *Data analysis in Particle Physics*, disponível em <<https://pdfs.semanticscholar.org/ae9a/e2e27521cf647f3d31c833d751d66bd8e283.pdf>> Acesso em: 22 de setembro de 2019.
- [160] S. Alekhin et al., *The PDF4LHC Working Group Interim Report*, arXiv/hep-ph:1101.0536. (2011).
- [161] M. Cacciari et al., *The $t\bar{t}$ cross-section at 1.8 TeV and 1.96 TeV: A Study of the systematics due to parton densities and scale dependence*, *JHEP* **04**, 068 (2004).
- [162] S. Catani, D. de Florian, M. Grazzini, P. Nason, *Soft gluon resummation for Higgs boson production at hadron colliders*, *JHEP* **07**, 028 (2003).
- [163] CMS B-tag POG, *Methods to apply b-tagging efficiency scale factors*, disponível em <<https://twiki.cern.ch/twiki/bin/viewauth/CMS/BTagSFMMethods>> Acesso em: 15 de outubro de 2019.
- [164] W.R. Rolke, A.M. Lopez, J. Conrad, *Limits and confidence intervals in the presence of nuisance parameters*, *Nucl. Instrum. Methods Phys. Res. Section A* **551**, 493-503 (2005).
- [165] G. Cowan, K. Cranmer, E. Gross, O. Vitells, *Asymptotic formulae for likelihood-based tests of new physics*, *Eur. Phys. J. C* **71**, 1554 (2011).
- [166] CMS Collaboration, *Combine*, disponível em <<http://cms-analysis.github.io/HiggsAnalysis-CombinedLimit/>> Acesso em: 22 de outubro de 2019.
- [167] N. Wardle, *Documentation of the RooStats-based statistics tools for Higgs PAG*, disponível em <<https://twiki.cern.ch/twiki/bin/viewauth/CMS/SWGuideHiggsAnalysisCombinedLimit>> Acesso em: 22 de outubro de 2019.

-
- [168] J. Alwall, R. Frederix, S. Frixione et al., *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, J. High Energ. Phys. **79** (2014).
- [169] S. Agostinelli, J. Allison, K. Amako et al., *Geant4 - a simulation toolkit*, Nuclear Instruments and Methods in Physics Research A **506**, 250-303 (2003).
- [170] T. Sjöstrand, S. Ask, J.R. Christiansen et al., *An introduction to PYTHIA 8.2*, Computer Physics Communications **191**, 159-177 (2015).

Apêndice A - Serviço no Banco de Dados do HCAL no CMS

Durante um período de aproximadamente dois anos (início de 2017 até início de 2019), o autor realizou um serviço no experimento CMS, ocupando a posição de especialista de banco de dados do grupo responsável pelo calorímetro hadrônico (*HCAL DB expert*). Os aspectos do trabalho realizado nesse período são descritos a seguir.

A.1 Trabalho no banco de dados

O CMS possui um banco de dados que armazena as informações relativas aos alinhamentos e calibrações de todas as partes do detector. A unidade básica de informação presente no banco de dados é denominada *payload*. Um *payload* representa o conjunto de parâmetros, associados a um tipo específico de condição do detector, consumidos no fluxo de trabalho do processamento de dados físicos. Ele é tipicamente armazenado em arquivos de texto cuja a estrutura interna depende do tipo de condição especificada. No *software* oficial do experimento (CMSSW), cada tipo de condição está associada a uma classe em C++ projetada para importar o conteúdo do *payload*. Tais classes são chamadas de *records*.

Uma *tag* especifica um conjunto de *payloads* referentes ao mesmo tipo de condição do detector. Os *payloads* que compõem a *tag* são válidos em diferentes intervalos de tempo especificados pelo parâmetro conhecido como IOV (*interval of validity*), geralmente representado pelo número do *run*¹ em que o *payload* começou a valer. As tags são armazenadas dentro de arquivos DB, arquivos genéricos de banco de dados que armazenam os dados em um formato estruturado. Os *payloads* são adicionados às *tags* utilizando *scripts* produzidos centralmente pela colaboração do CMS. Abaixo é exibido um exemplo do conteúdo de uma *tag* com 8 inserções de *payloads* do tipo “HcalRespCorrs”. A primeira coluna mostra o IOV, a segunda informa o momento em que o *payload* foi inserido, a terceira exibe um código de identificação denominado *payload hash* (único para cada *payload*) e a quarta

¹Intervalo discreto da tomada de dados.

coluna informa o tipo do objeto (tipo da condição). De acordo com o que foi exposto, pode-se dizer que o penúltimo *payload* é válido do *run* 190456 até o *run* 190644 e que o mesmo *payload* foi inserido nos *runs* 175860 e 190645, visto que possuem o mesmo *payload hash*.

Since: Run	Insertion Time	Payload	Object Type
1	2019-07-19 05:03:45	6776d6c69ab38619d010ef2769b6e2b22f98d1f5	HcalRespCorrs
148656	2019-07-19 05:03:45	d6c553c8d94c20428d163d495be5c1c9662af16c	HcalRespCorrs
155418	2019-07-19 05:03:45	460f200ca8fa335996c8e97ad5be8f0d07eb5919	HcalRespCorrs
160431	2019-07-19 05:03:45	af40407dc988604cd6f41b022974b60cbc271c04	HcalRespCorrs
175860	2019-07-19 05:03:45	5a2c08d112398243ead99d741ae1f3f80de7b0dd	HcalRespCorrs
189778	2019-07-19 05:03:45	726c9d2bb661c1d7375414ce050af934f06dcb8f	HcalRespCorrs
190456	2019-07-19 05:04:51	23baa20129f7c14fda09ec7fbc5a1ba51aaa052b	HcalRespCorrs
190645	2019-07-19 05:06:02	5a2c08d112398243ead99d741ae1f3f80de7b0dd	HcalRespCorrs

As *global tags* são rótulos que identificam um conjunto de *tags* associadas aos diferentes tipos de condições do detector. Uma *global tag* pode conter, por exemplo, o conjunto completo de alinhamentos e calibrações necessários para a produção das simulações de MC referentes ao ano de 2018. Uma *queue* é uma entidade mutável composta por um conjunto de *tags*. A partir das *queues*, usuários podem criar candidatas a *global tags* para serem utilizadas na realização de testes. É também a partir das *queues*, que os especialistas em alinhamentos e calibrações (*ALCa experts*) criam as *global tags*.

O conjunto completo de alinhamentos e calibrações associados ao calorímetro hadrônico estão divididos em 26 tipos diferentes de condições. Entre as principais condições pode-se citar:

- **HcalGains:** composta por resultados de medições de dados fornecidos por fios e várias correções aplicadas ao brilho magnético em cintiladores, variações de alta tensão, etc;
- **HcalRespCorrs:** coeficientes sem unidade que contêm as correções que devem ser aplicadas à resposta do HCAL;
- **HcalChannelQuality:** objeto de 32-*bits* que consiste em *bits* acionados ou não, de acordo com condições predefinidas para uma célula específica do HCAL;
- **HcalL1TriggerObjects:** objeto projetado para armazenar informações de cada canal do HCAL no momento de geração das HCAL LUTs (*Look-Up Tables*²).

Enquanto ocupava a posição de especialista de banco de dados do HCAL, o autor era responsável pela criação, documentação e movimentação, dentro do banco de dados

²Matrizes utilizadas na substituição de computações em tempo de execução por simples operações de indexação de matrizes.

do CMS, de todas as *tags* associadas às condições do HCAL. Sendo também o mediador das interações entre o grupo do HCAL e os *AlCa experts*. O fluxo de trabalho normal começava pelo recebimento dos arquivos de texto com *payloads* de diferentes tipos de condições, originários de análises realizadas pelos especialistas em calibração das condições do HCAL. Tendo os *payloads*, o autor produzia as *tags* que eram armazenadas dentro de arquivos DB. Posteriormente, essas *tags* eram inseridas dentro do banco de dados do CMS, podendo ter os mais diversos fins. Usualmente, as *tags* inseridas no banco de dados eram testadas pelos especialistas do HCAL antes de sua inserção nas *global tags*, sendo frequentemente necessário que o autor criasse candidatas a *global tags* para serem usadas nos testes. Todas as *tags* inseridas no banco de dados foram devidamente documentadas pelo autor na página interna do CMS na internet, referente às condições do HCAL.

As *tags* e *global tags* são separadas em diferentes categorias que definem a utilidade dos *payloads* no experimento, por exemplo *global tags* de MC são criadas para serem usadas na produção de amostras de simulação, *global tags offline* são usadas no reprocessamento de dados reais armazenados e *global tags HLT* são utilizadas pelos *HLT triggers* nas reconstrução de eventos *online*. Os *payloads* produzidos são apresentados pelos especialistas em calibração nas reuniões do grupo alinhamentos e calibrações (*AlCa meetings*). Uma vez recebido o consentimento dos *AlCa experts*, o autor era responsável pela inserção das *tags* criadas dentro das *global tags* apropriadas e deveria informar a comunidade interna do CMS sobre a atualização, no fórum de discussões do grupo de alinhamentos e calibração.

A.2 Trabalho no projeto de automatização

A implantação das condições do HCAL é um processo que envolve muitas pessoas: especialistas em calibração, especialistas do banco de dados, validadores de diferentes subgrupos do HCAL (HLT, L1) e especialistas em operações. Atualmente, o processo completo exige muitas etapas e a comunicação entre os especialistas, são suscetíveis a erros e podem gerar atrasos desnecessários. Considerando que a correção dos danos causados pela radiação e a corrente escura dos fotomultiplicadores de silício (*silicon photomultipliers* - SiPM) são atualizadas no HCAL a cada 10 fb^{-1} (atualização das condições do tipo HcalGains), e a existência de outras condições atualizadas com menos frequência, mas exigindo o mesmo processo de implantação, é imperativo tornar o processo mais rápido e robusto. A equipe do HCAL está se esforçando para automatizar tarefas repetitivas comuns.

Atualizações das condições HcalPedestals, HcalGains, HcalRespCorrs e HcalChannelQuality implicam em modificações na condição do HcalL1TriggerObjects. Desse modo, quando o autor recebia um *payload* para atualizar uma ou mais das condições citadas, após

a criação das *tags* correspondentes, ele deveria criar uma candidata a *global tag* incluindo as novas *tags* criadas e entrar em contato com os especialistas de L1 *Trigger* do HCAL para que eles gerassem as LUTs e a nova condição para HcalL1TriggerObjects a partir da candidata a *global tag* fornecida. Então, os especialistas de L1 *Trigger* forneciam o arquivo de texto com a nova condição do HcalL1TriggerObjects para o autor e ele criava a *tag* correspondente. Na última etapa do processo, a nova *tag* de HcalL1TriggerObjects era indicada aos especialistas de L1 *Trigger* para que eles gerassem os gráficos de validação do L1 *trigger*. Visando eliminar toda essa interação entre os especialistas do banco de dados e do L1 *trigger* do HCAL, durante seu último semestre como especialista de banco de dados do HCAL, o autor foi designado para automatizar esse processo.

No projeto de automatização, o autor fez uso de um servidor de integração contínua (CI), de código aberto e escrito em Java, conhecido como Jenkins. Integração Contínua é uma prática de desenvolvimento, na qual os desenvolvedores podem integrar o código em um repositório compartilhado de maneira mais dinâmica. Jenkins ajuda a automatizar o processo de desenvolvimento de *software*, facilitando os aspectos técnicos da entrega contínua. Toda atualização é verificada por um compilador automatizado, permitindo que se verifique a integrabilidade do código, detectando problemas com antecedência. O Jenkins possui um *plug-in* que permite sua integração com um código armazenado em um repositório do GitHub, companhia que oferece hospedagem de *códigos* em sua página na internet para que sejam atualizados utilizando seu *software* de desenvolvimento e controle de versões, chamado GIT.

O autor produziu os *scripts* de automação e os hospedou em HcalConditionsAutomation. O diretório do GitHub onde estão os *scripts* é monitorado por um servidor Jenkins do CMS, configurado com a ajuda do especialista em Jenkins do CMS, Malik Shahzad Muzaffar. O servidor Jenkins foi configurado para rodar os *scripts* de automação em um computador do CMS toda vez que houver uma modificação em algum dos arquivos dentro do diretório monitorado. Entretanto, o processo automatizado só é iniciado se a opção “validate” for configurada com “yes” no arquivo de configuração NewValidation.txt cujo conteúdo é exposto a seguir.

```
#-----  
# VALIDATION INFO  
#-----  
validate      no  
  
week          11  
year          2018  
  
outdir        /eos/user/g/gcorreia/Hcal_project/Outputs  
  
#-----  
# LUT GENERATION
```

```

#-----
NewLUTtag    Physics2018v11Test
NewGT        101X_dataRun2_HLT_Candidate_2018_10_24_18_32_13
NewRun       325500

OldLUTtag    Physics2018v10
OldGT        101X_dataRun2_HLT_Candidate_2018_09_04_15_22_38
OldRun       322800

HOAsciiInput HO_ped9_inputLUTcoderDec.txt
release_LUT   CMSSW_10_4_0_pre1
arch_LUT      slc6_amd64_gcc700

#-----
# TAG GENERATION
#-----
geometry      2018

#-----
# L1 RATE VALIDATION
#-----
run           325170
lumi_start    1
lumi_end      3
dataset       /ZeroBias/Run2018D-v1/RAW
tier2         T2_CH_CERN

release_L1    CMSSW_10_3_1
arch_L1       slc6_amd64_gcc700

```

Na parte “VALIDATION INFO” é informado se será realizada a validação, o mês e o ano em que está sendo executada e o diretório para onde serão enviados os arquivos de saída. Na parte “LUT GENERATION” são fornecidas informações para a geração das novas LUTs e das LUTs anteriores para que sejam geradas as LUTs, gráficos de comparação e o novo *payload* de HcalL1TriggerObjects. As informações requeridas são os nomes que serão atribuídos aos arquivos com as LUTs, junto com a *global tag* e o *run* utilizados para a extração das condições usadas na geração das LUTs. No final são incluídas as configurações relacionadas a versão do *software* CMSSW utilizada. Na parte “TAG GENERATION” é informada a geometria dos subdetectores do experimento que é passível de ser modificada toda vez que o experimento sofre atualizações. Essa informação é necessária para a criação da nova *tag* de HcalL1TriggerObjects. Por fim, na parte “L1 RATE VALIDATION” é fornecido o nome do arquivo com os dados utilizados na a geração dos gráficos de validação do L1 *trigger* do HCAL, sua localização no *grid* do CMS, junto com o *run* e a seções de luminosidade que devem ser consideradas. No final também são fornecidas informações sobre a versão do *software* CMSSW que deve ser utilizada.

A figura A.1 mostra o fluxo de ações do processo automatizado. O processo é iniciado

pelo especialista de banco de dados do HCAL após ter sido criada a candidata a *global tag*. Foi configurado no Jenkins o envio de emails para todos os membros do organização do GitHub, que hospeda os *scripts* de automação, no início e no fim do processo. O diretório de saída é compartilhado com todas as pessoas do grupo do HCAL que utilizam os arquivos gerados no processo.

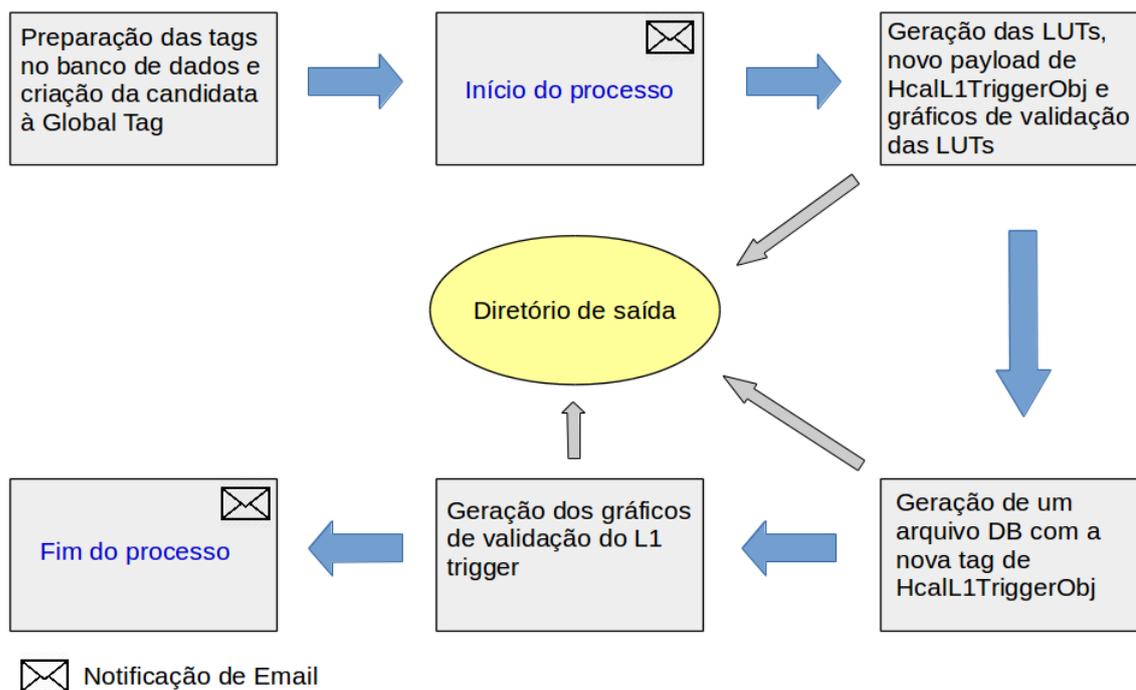


FIGURA A.1 – Figura esquemática mostrando o fluxo de ações do processo automatizado.

Apêndice B - Amostras de Monte Carlo e Dados

B.1 Amostras de Monte Carlo dos Backgrounds

Amostras reconstruídas com o software oficial do CMS (CMSSW), versão 8_0_X (Summer16).

Nome da amostra	σ [pb]	L [fb ⁻¹]
Amostras de W+Jatos		
WJetsToLNu_Pt-100To250_[2]/_ext1/_ext4	627,1	284179
WJetsToLNu_Pt-250To400_[2]/_ext1/_ext4	21,83	28299
WJetsToLNu_Pt-400To600_[2]/_ext1	2,635	4437
WJetsToLNu_Pt-600ToInf_[2]/_ext1	0,4102	4462
Amostras de $t\bar{t}$ +Jatos		
TTJets_[1]	831,76	12,191
Amostras de QCD		
QCD_HT100to200_[1]	27990000	0,003
QCD_HT200to300_[1]	1712000	0,011
QCD_HT300to500_[1]	347700	0,049
QCD_HT500to700_[1]	32100	0,59
QCD_HT700to1000_[1]	6831	2,29
QCD_HT1000to1500_[1]	1207	3,95
QCD_HT1500to2000_[1]	119,9	33,12
QCD_HT2000toInf_[1]	25,24	78,91
Amostras de $Z(\nu\nu)$ +Jatos		
ZJetsToNuNu_HT-100To200_[3]	344,83	15,21
ZJetsToNuNu_HT-200To400_[3]	95,53	53,73
ZJetsToNuNu_HT-400To600_[3]	13,198	77,31
ZJetsToNuNu_HT-600To800_[3]	3,148	1831
ZJetsToNuNu_HT-800To1200_[3]	1,451	1496
ZJetsToNuNu_HT-1200To2500_[3]	0,3546	1042
ZJetsToNuNu_HT-2500ToInf_[3]	0,00854	47427

TABELA B.1 – Amostras de Monte Carlo dos Backgrounds.

Nome da amostra	σ [pb]	L [fb ⁻¹]
Amostras de γ +Jatos	LO	
GJets_HT-100To200_[1]	9238	0,56
GJets_HT-200To400_[1]	2305,0	4,35
GJets_HT-400To600_[1]	274,4	9,22
GJets_HT-600ToInf_[1]	93,46	26,36
Amostras de DY+Jatos	LO(*NNLO)	
DYJetsToLL_M-1To5_HT-150to200_[1]	1124	17,45
DYJetsToLL_M-1To5_HT-200to400_[1]	789,8	15,45
DYJetsToLL_M-1To5_HT-400to600_[1]	65	75,08
DYJetsToLL_M-1To5_HT-600toInf_[1]	16,72	118,3
DYJetsToLL_M-5to50_HT-100to200_[1]	224,2	4,50
DYJetsToLL_M-5to50_HT-200to400_[1]	37,2	28,09
DYJetsToLL_M-5to50_HT-400to600_[1]	3,581	283,9
DYJetsToLL_M-5to50_HT-600toInf_[1]	1,124	908,4
DYJetsToLL_M-50_HT-100to200_[1]	181,3*	15,17
DYJetsToLL_M-50_HT-200to400_[1]	50,42*	19,08
DYJetsToLL_M-50_HT-400to600_[1]	6,984*	153,3
DYJetsToLL_M-50_HT-600to800_[1]	1,681*	4933
DYJetsToLL_M-50_HT-800to1200_[1]	0,7754*	3442
DYJetsToLL_M-50_HT-1200to2500_[1]	0,1862*	3201
DYJetsToLL_M-50_HT-2500toInf_[1]	0,004385*	91104
Amostras de Top único	NLO	
ST_s-channel_4f_InclusiveDecays_[4]	10,32	2935
ST_t-channel_antitop_4f_inclusiveDecays_[5]	80,95	479,4
ST_t-channel_top_4f_inclusiveDecays_[5]	136,02	494,3
ST_tW_top_5f_inclusiveDecays_[6]	35,85	27,67
ST_tW_antitop_5f_inclusiveDecays_[6]	35,85	27,85
Amostras de Dibósons	NLO(*NNLO)	
ZZ_[7]	16,523	59,92
WZ_[7]	47,13	21,22
WW_[7]	118,7*	8,37

TABELA B.2 – Amostras de Monte Carlo dos Backgrounds. (Continuação)

- [1] TuneCUETP8M1_13TeV-madgraphMLM-pythia8
- [2] TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8
- [3] 13TeV-madgraph
- [4] 13TeV-amcatnlo-pythia8
- [5] 13TeV-powhegV2-madspin-pythia8_TuneCUETP8M1
- [6] 13TeV-powheg-pythia8_TuneCUETP8M2T4
- [7] TuneCUETP8M1_13TeV-pythia8

B.2 Amostras de Monte Carlo do Sinal

As amostras de sinal foram reconstruídas com o software oficial do CMS (CMSSW), versão 8_0_20 (Summer16). A geração de eventos foi realizada pelo MadGraph5_aMC@NLO [168] e Pythia 8 [170] para o estágio de decaimento dos gluinos e hadronização dos jatos. A simulação da resposta do detector foi feita pelo *software* GEANT-4 [169].

Nome da amostra	σ [pb]	L [fb ⁻¹]
Amostras de Sinal	LO	
Signal_250_30_Summer16	1501	0,025
Signal_250_40_Summer16	1501	0,025
Signal_250_60_Summer16	1501	0,025
Signal_250_80_Summer16	1501	0,025
Signal_250_100_Summer16	1501	0,025
Signal_500_30_Summer16	36,22	0,788
Signal_500_40_Summer16	36,22	0,788
Signal_500_60_Summer16	36,22	0,786
Signal_500_80_Summer16	36,22	0,661
Signal_500_100_Summer16	36,22	0,792
Signal_750_30_Summer16	2,822	8,62
Signal_750_40_Summer16	2,822	9,13
Signal_750_60_Summer16	2,822	9,04
Signal_750_80_Summer16	2,822	9,09
Signal_750_100_Summer16	2,822	9,15
Signal_1000_30_Summer16	0,359	67,5
Signal_1000_40_Summer16	0,359	60,6
Signal_1000_60_Summer16	0,359	67,5
Signal_1000_80_Summer16	0,359	65,0
Signal_1000_100_Summer16	0,359	67,0
Signal_1250_30_Summer16	0,0587	379
Signal_1250_40_Summer16	0,0587	396
Signal_1250_60_Summer16	0,0587	398
Signal_1250_80_Summer16	0,0587	400
Signal_1250_100_Summer16	0,0587	401

TABELA B.3 – Amostras de Monte Carlo do Sinal.

B.3 Amostras de Dados Reais

Nome da amostra	L [fb ⁻¹]
Amostras com energia transversa faltante	35,92
Data_MET_Run2016B-23Sep2016	5,750
Data_MET_Run2016C-23Sep2016	2,573
Data_MET_Run2016D-23Sep2016	4,242
Data_MET_Run2016E-23Sep2016	4,025
Data_MET_Run2016F-23Sep2016	3,105
Data_MET_Run2016G-23Sep2016	7,576
Data_MET_Run2016H-Prompt	8,651
Amostras com elétrons	35,92
Data_SingleElectron_Run2016B-23Sep2016	5,750
Data_SingleElectron_Run2016C-23Sep2016	2,573
Data_SingleElectron_Run2016D-23Sep2016	4,242
Data_SingleElectron_Run2016E-23Sep2016	4,025
Data_SingleElectron_Run2016F-23Sep2016	3,105
Data_SingleElectron_Run2016G-23Sep2016	7,576
Data_SingleElectron_Run2016H-Prompt	8,651
Amostras com múons	35,92
Data_SingleMuon_Run2016B-23Sep2016	5,750
Data_SingleMuon_Run2016C-23Sep2016	2,573
Data_SingleMuon_Run2016D-23Sep2016	4,242
Data_SingleMuon_Run2016E-23Sep2016	4,025
Data_SingleMuon_Run2016F-23Sep2016	3,105
Data_SingleMuon_Run2016G-23Sep2016	7,576
Data_SingleMuon_Run2016H-Prompt	8,651

TABELA B.4 – Amostras de Dados.

Apêndice C - Validação das variáveis de busca

C.1 Região Leptônica

C.1.1 Variáveis de CR1

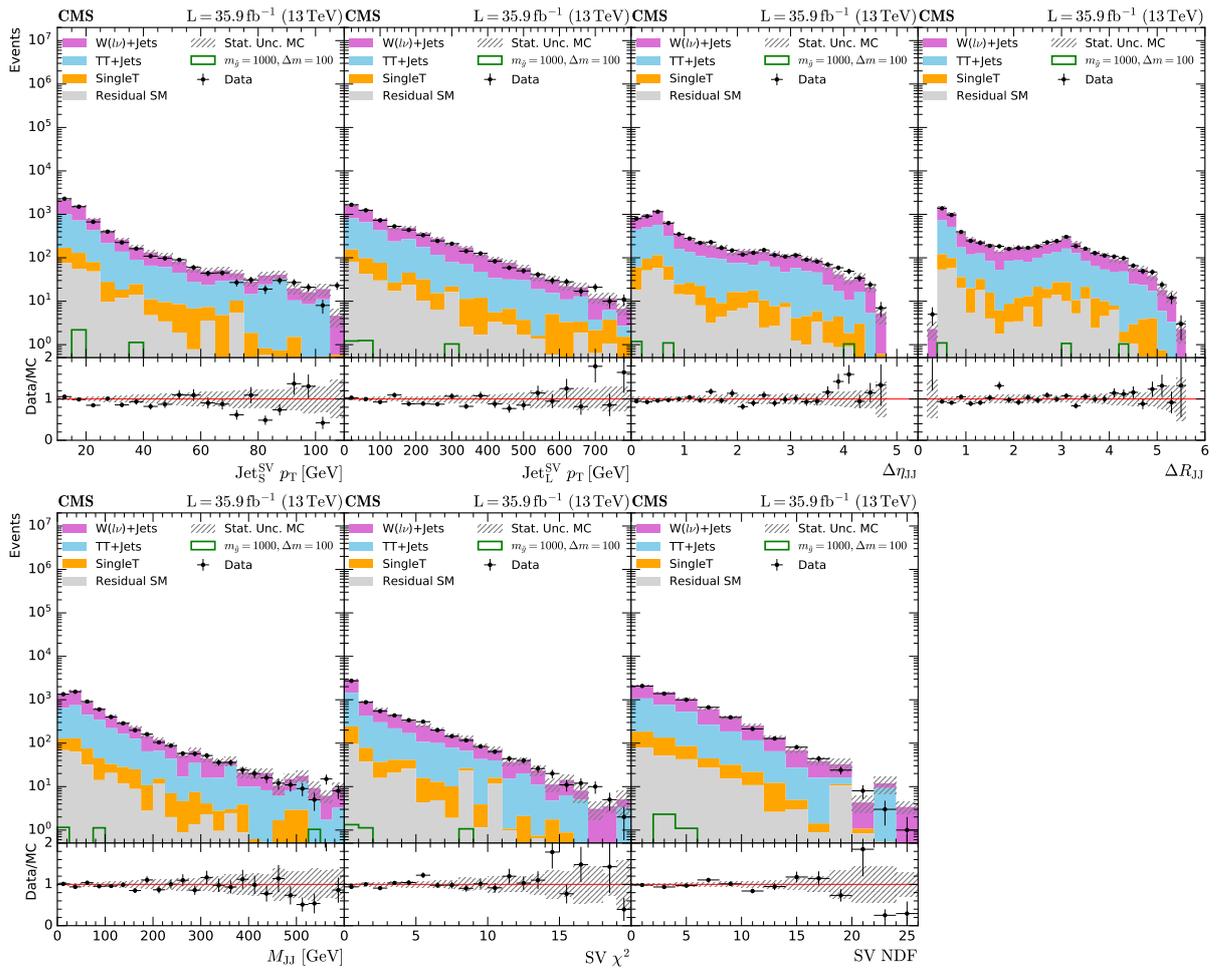


FIGURA C.1 – Validação das variáveis regionais na sub-região CR1 da região leptônica.

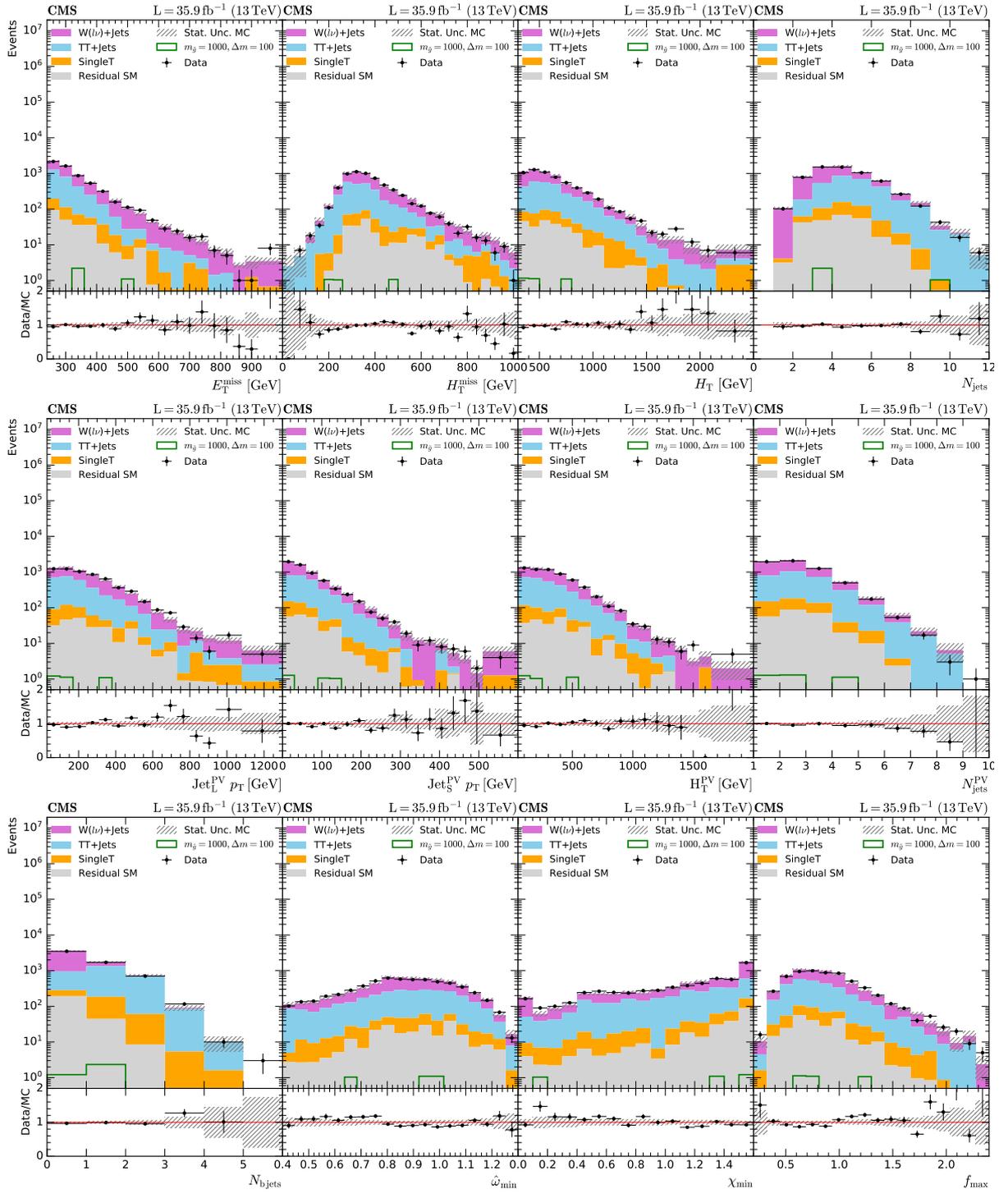


FIGURA C.2 – Validação das variáveis globais na sub-região CR1 da região leptônica.

C.1.2 Variáveis de CR2

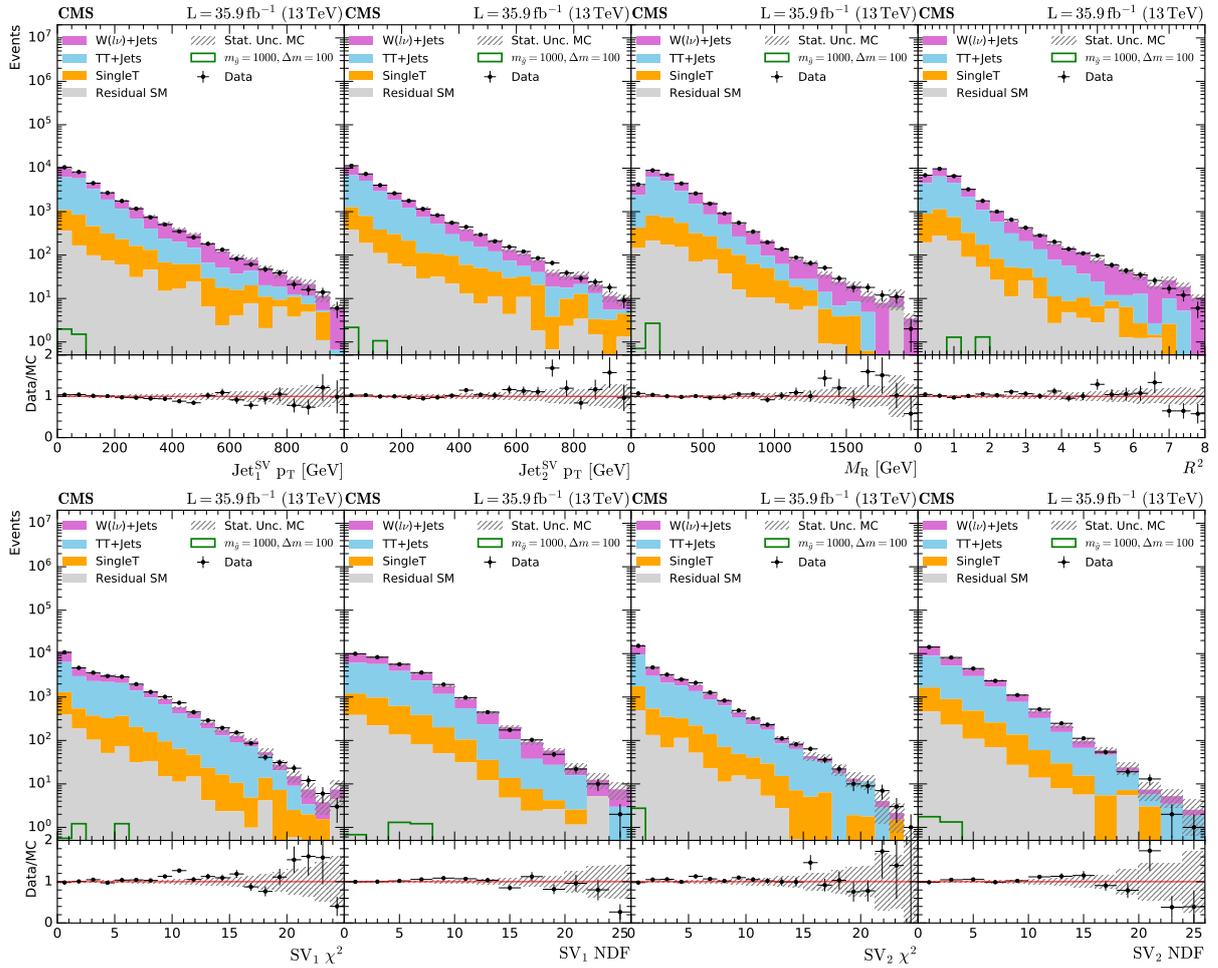


FIGURA C.3 – Validação das variáveis regionais na sub-região CR2 da região leptônica.

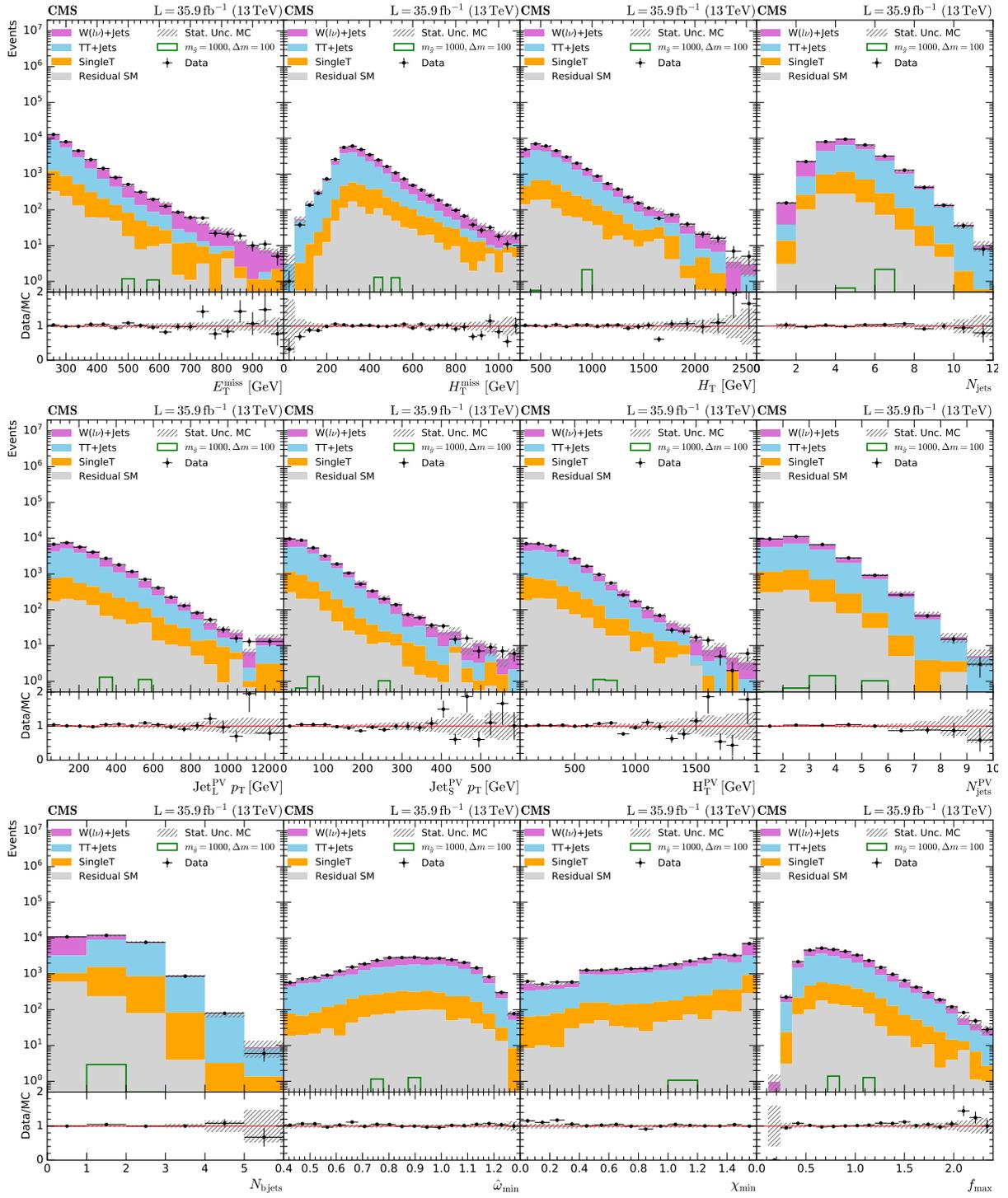


FIGURA C.4 – Validação das variáveis globais na sub-região CR2 da região leptônica.

C.1.3 Variáveis de CR3

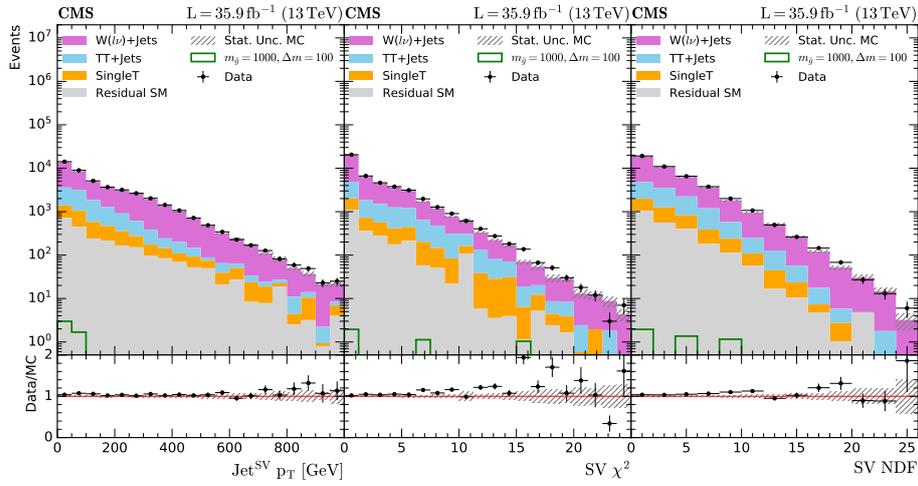


FIGURA C.5 – Validação das variáveis regionais na sub-região CR3 da região leptônica.

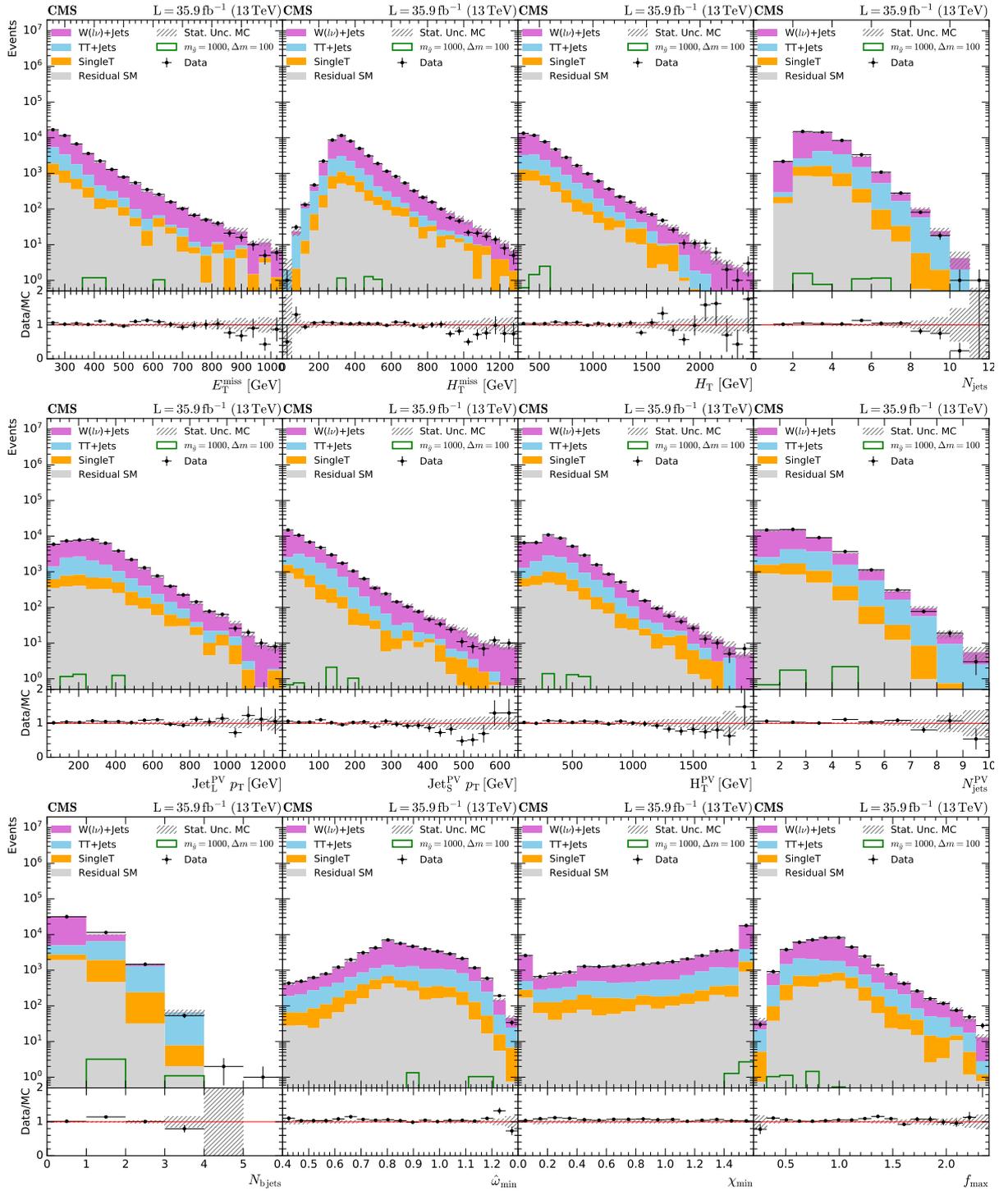


FIGURA C.6 – Validação das variáveis globais na sub-região CR3 da região leptônica.

C.1.4 Variáveis de CR4

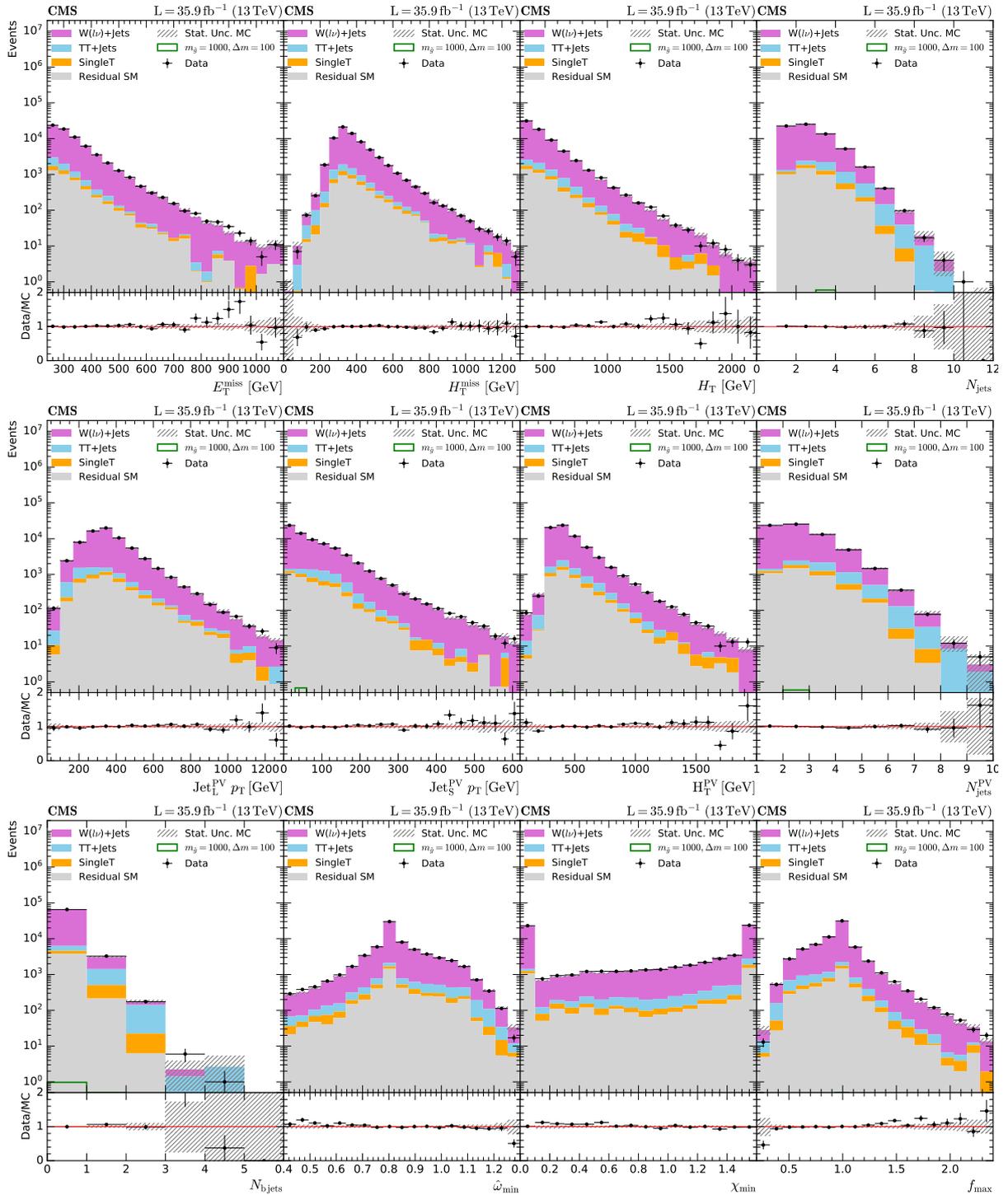


FIGURA C.7 – Validação das variáveis globais na sub-região CR4 da região leptônica.

C.2 Região de $Z(\nu\nu)+\text{Jatos}$

C.2.1 Variáveis de CR1

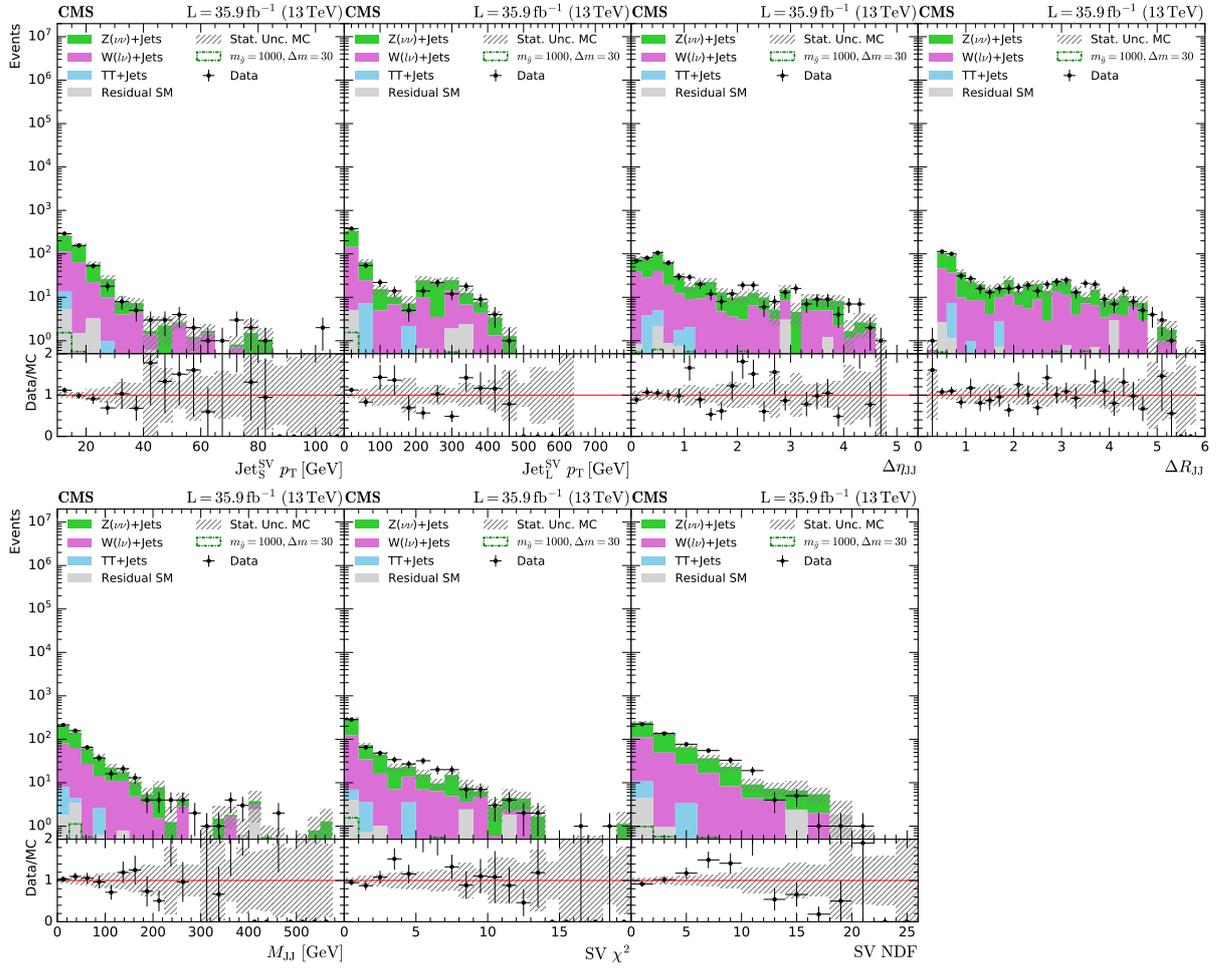


FIGURA C.8 – Validação das variáveis regionais na sub-região CR1 da região de $Z(\nu\nu)+\text{Jatos}$.

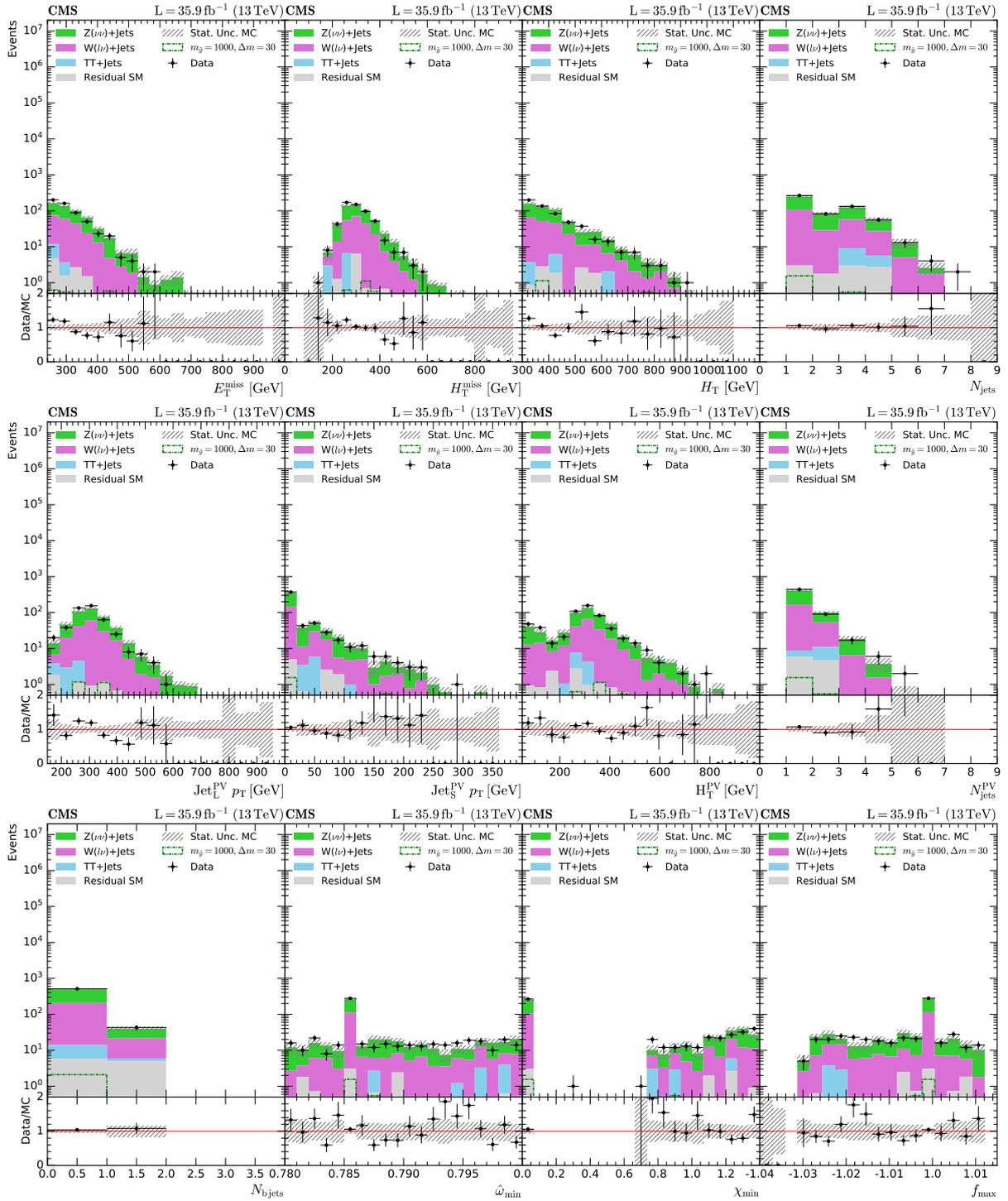
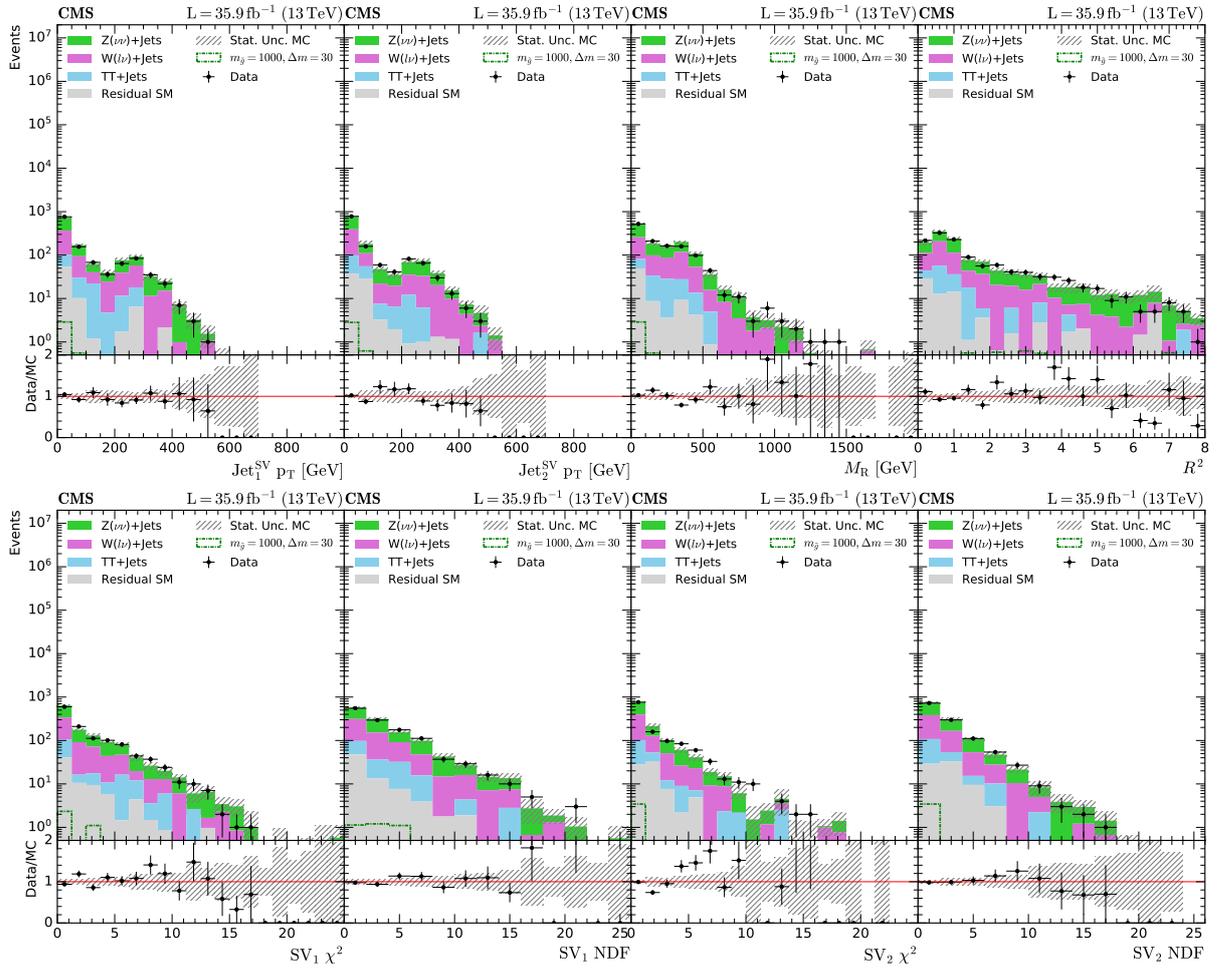
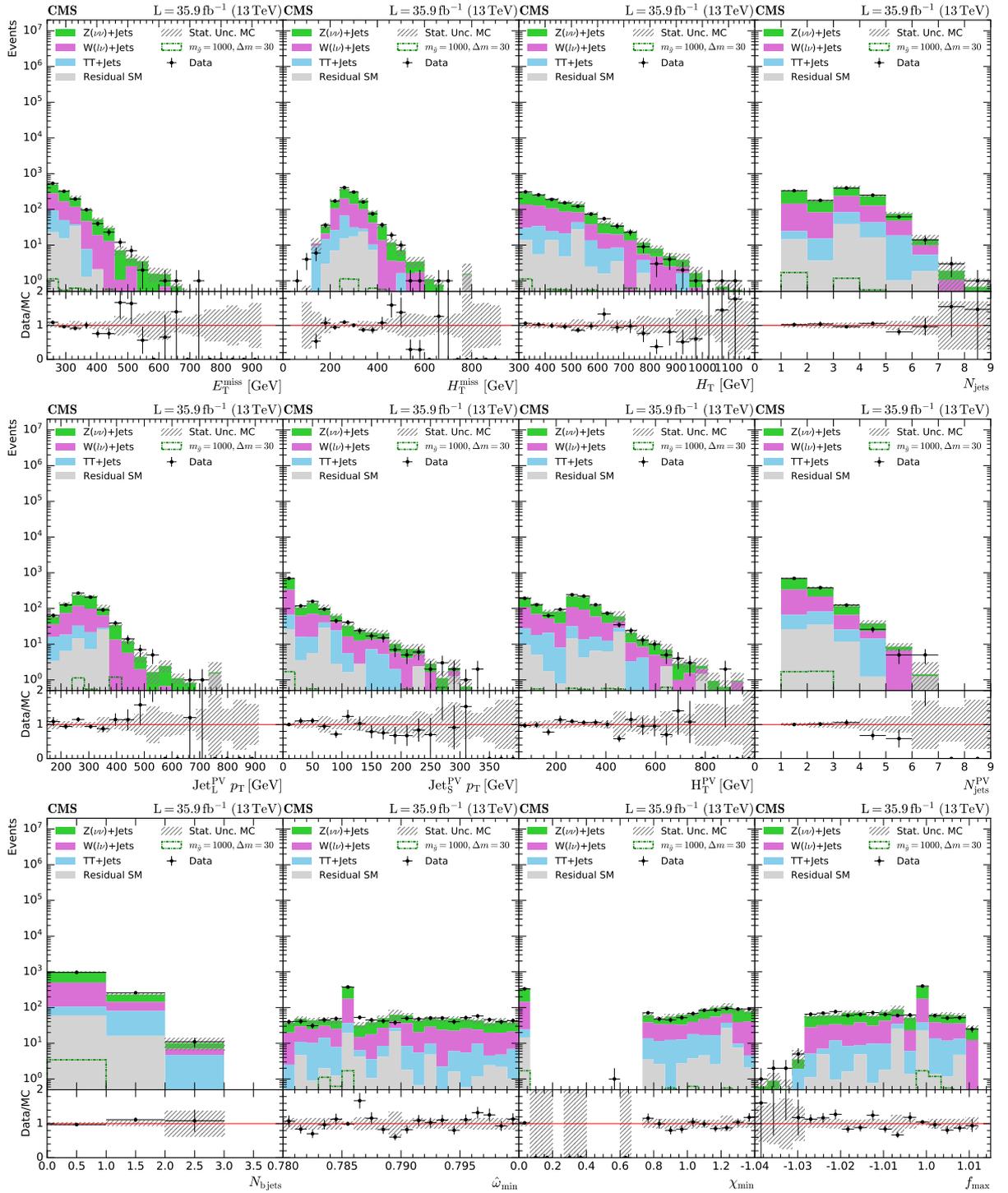


FIGURA C.9 – Validação das variáveis globais na sub-região CR1 da região de $Z(\nu\nu)+\text{Jatos}$.

C.2.2 Variáveis de CR2


 FIGURA C.10 – Validação das variáveis regionais na sub-região CR2 da região de $Z(\nu\nu)+\text{Jatos}$.


 FIGURA C.11 – Validação das variáveis globais na sub-região CR2 da região de $Z(\nu\nu)+\text{Jatos}$.

C.2.3 Variáveis de CR3

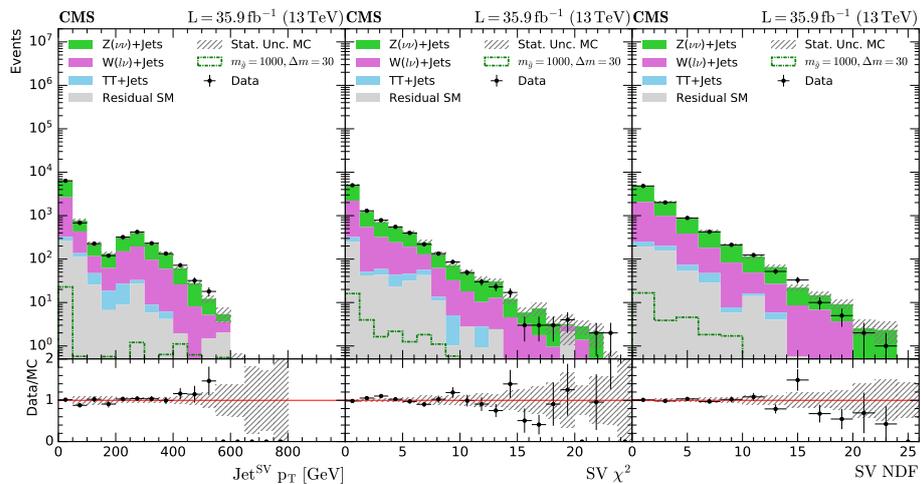
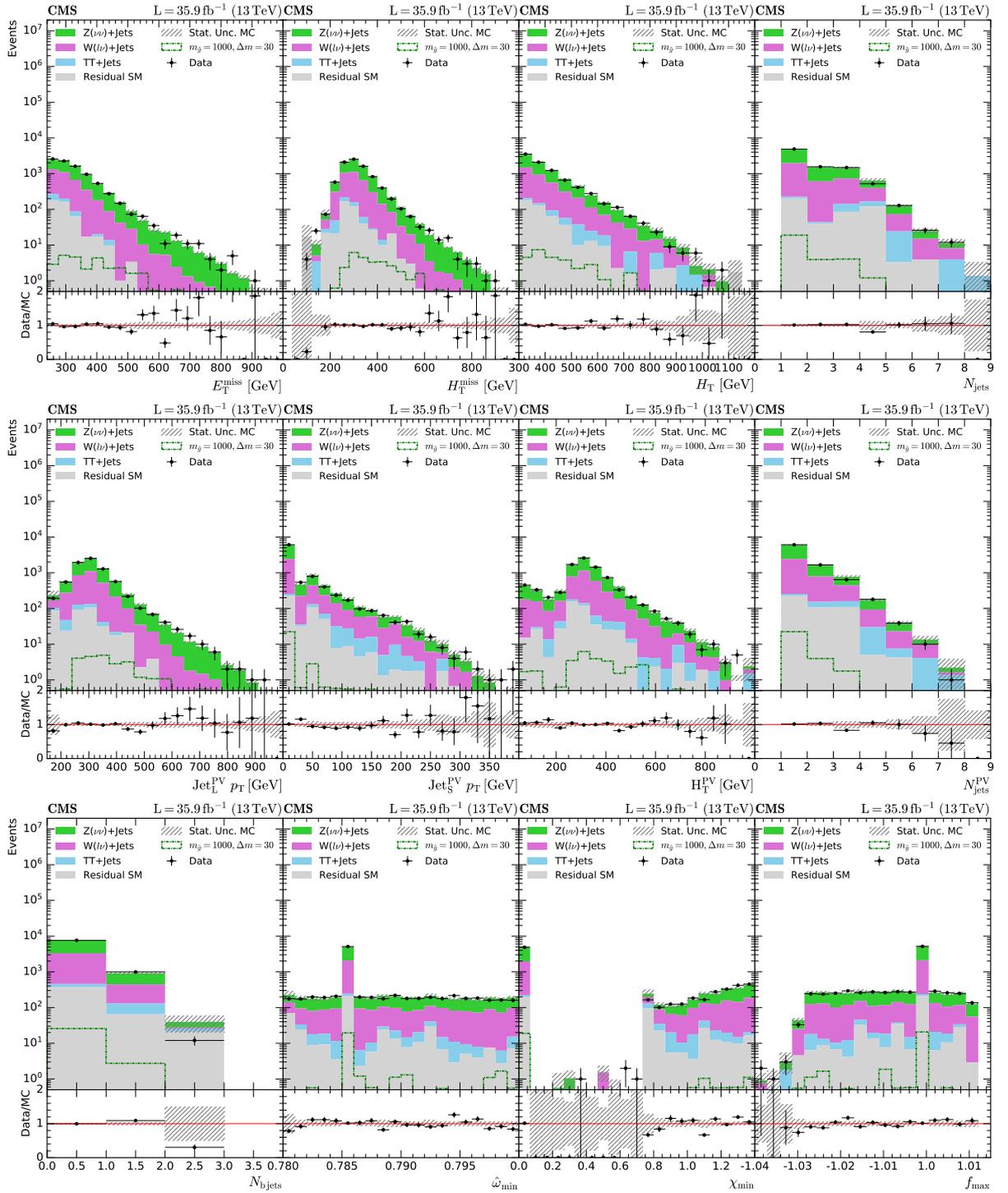


FIGURA C.12 – Validação das variáveis regionais na sub-região CR3 da região de $Z(\nu\nu)+\text{Jatos}$.


 FIGURA C.13 – Validação das variáveis globais na sub-região CR3 da região de $Z(\nu\nu)+\text{Jatos}$.

C.2.4 Variáveis de CR4

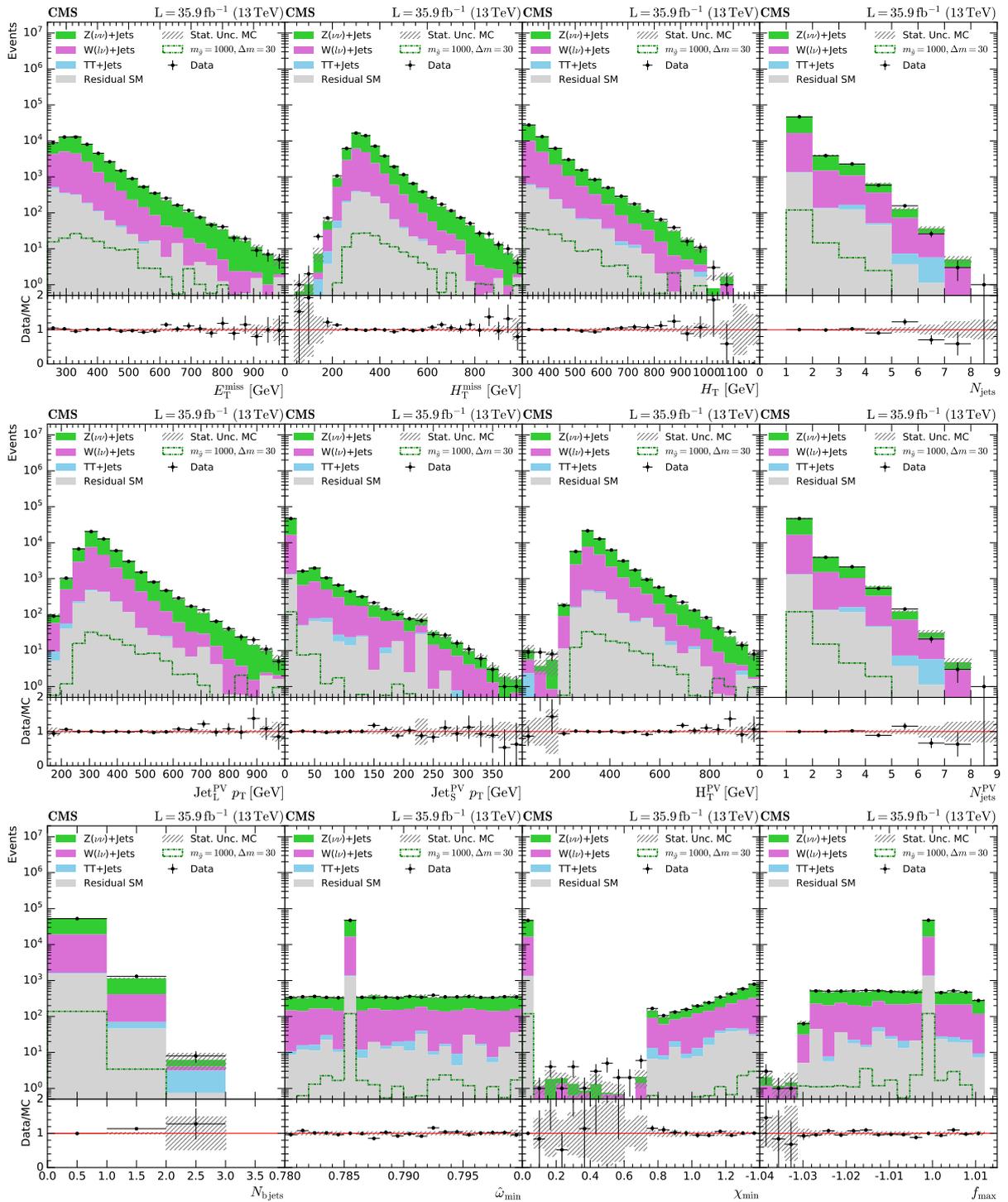


FIGURA C.14 – Validação das variáveis globais na sub-região CR4 da região de $Z(\nu\nu)+\text{Jatos}$.

C.3 Região de QCD

C.3.1 Variáveis de CR1

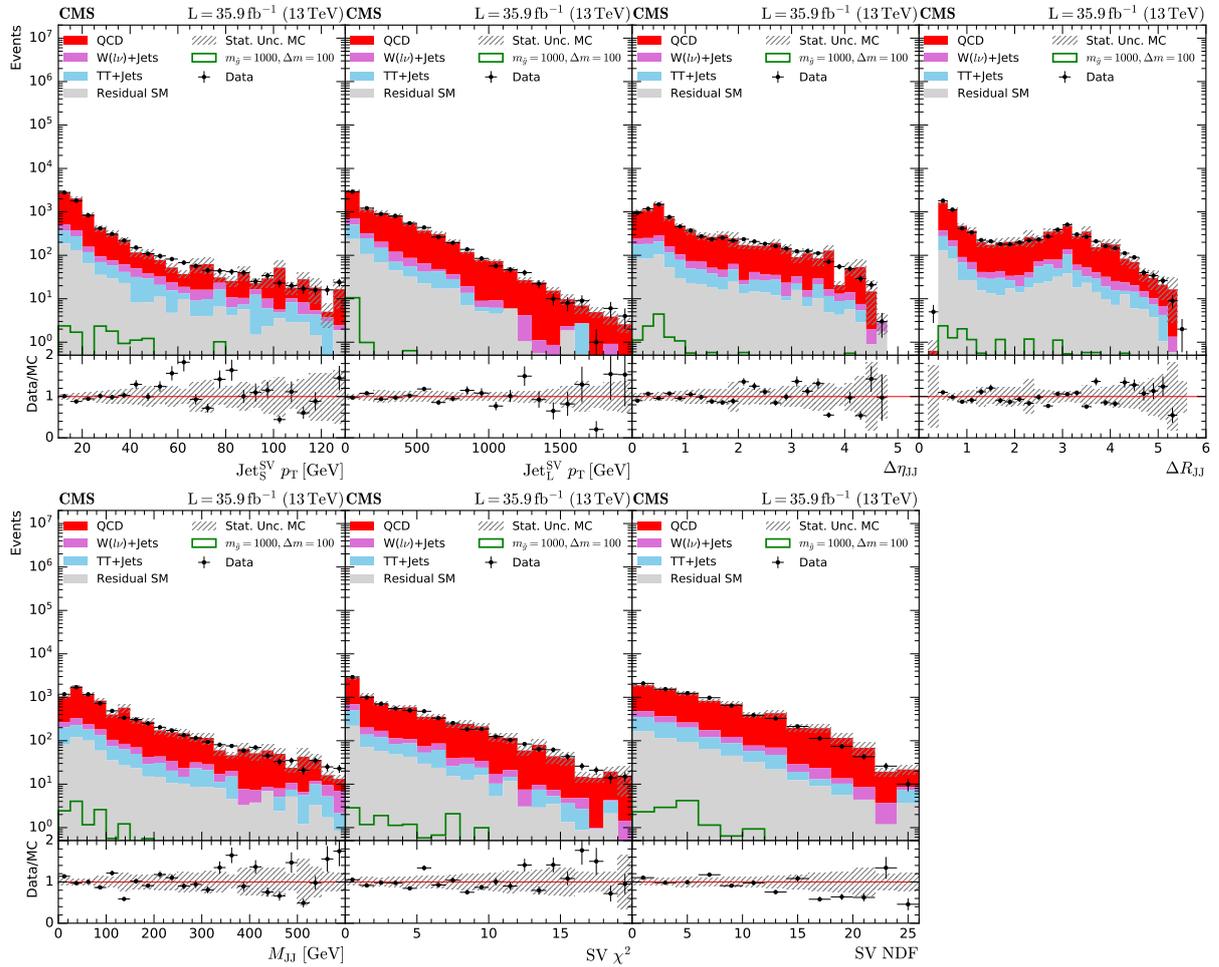


FIGURA C.15 – Validação das variáveis regionais na sub-região CR1 da região de QCD.

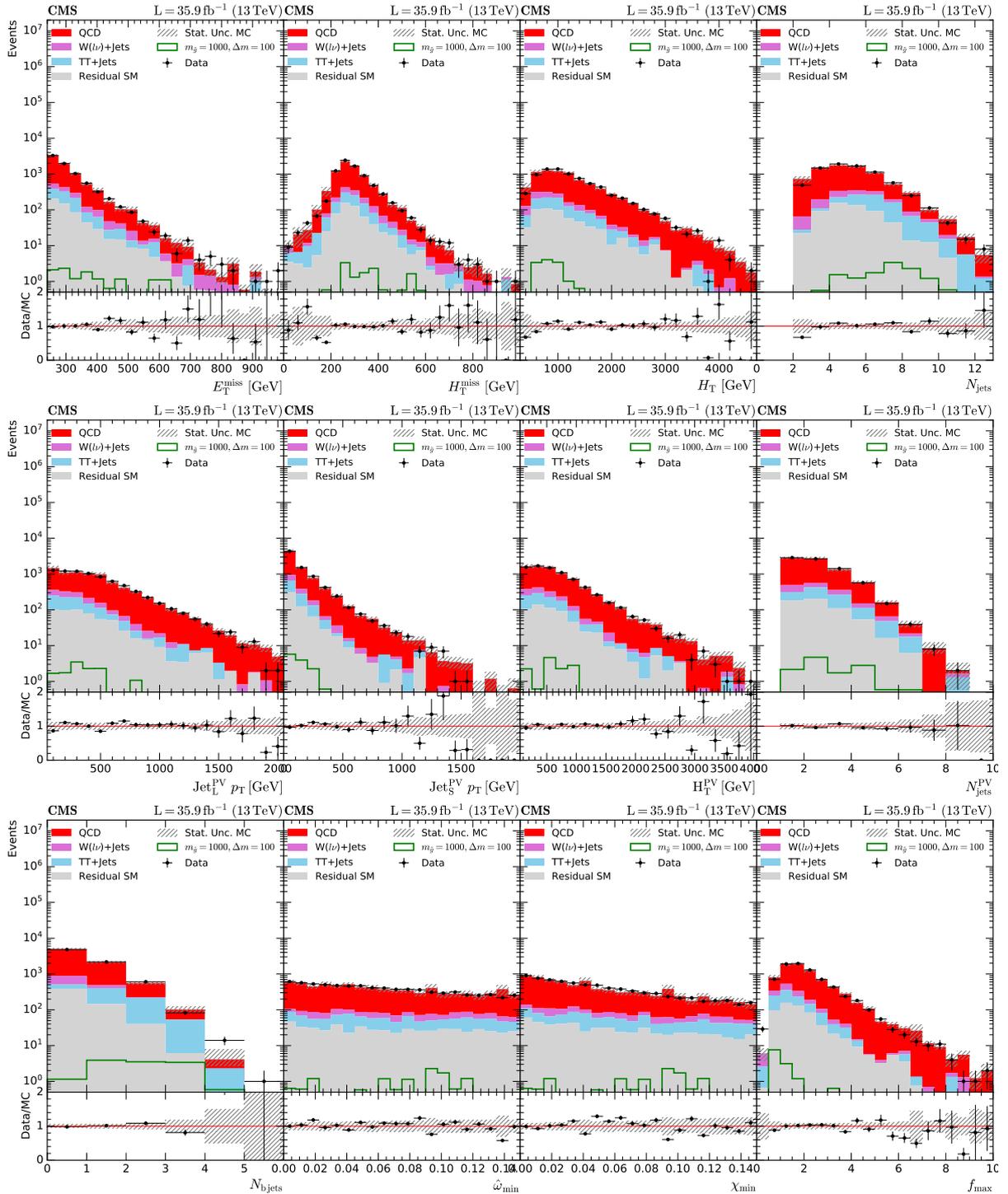


FIGURA C.16 – Validação das variáveis globais na sub-região CR1 da região de QCD.

C.3.2 Variáveis de CR2

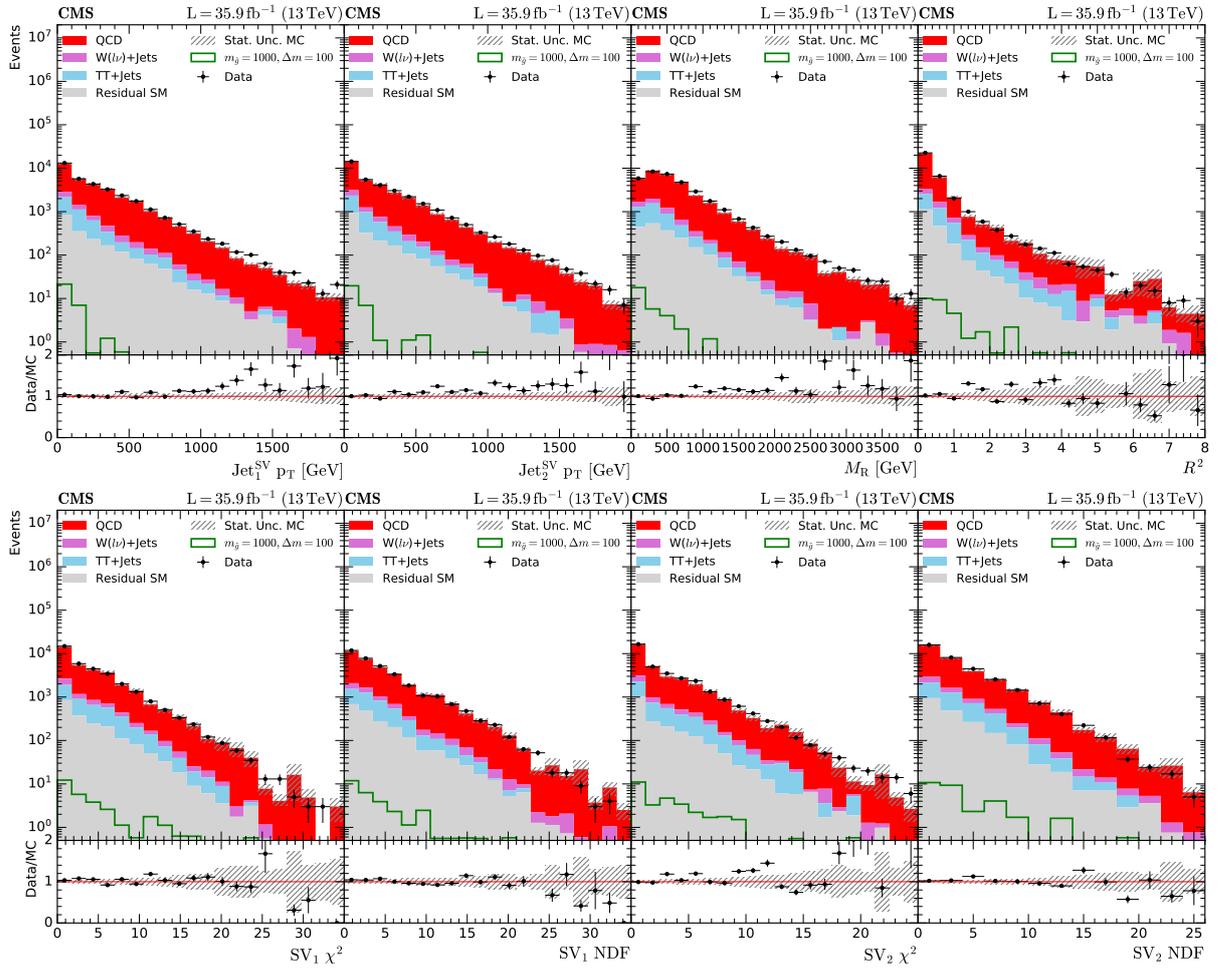


FIGURA C.17 – Validação das variáveis regionais na sub-região CR2 da região de QCD.

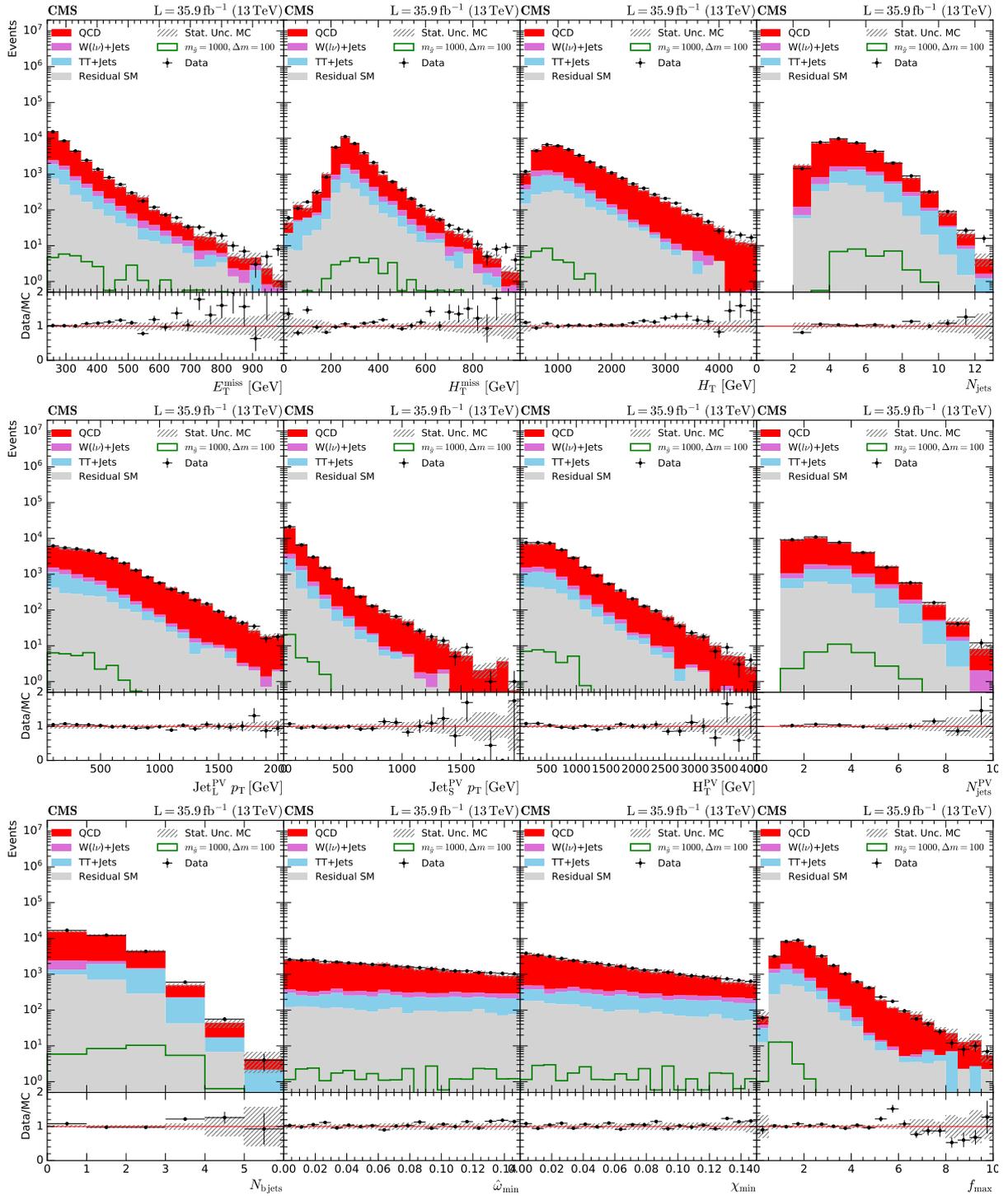


FIGURA C.18 – Validação das variáveis globais na sub-região CR2 da região de QCD.

C.3.3 Variáveis de CR3

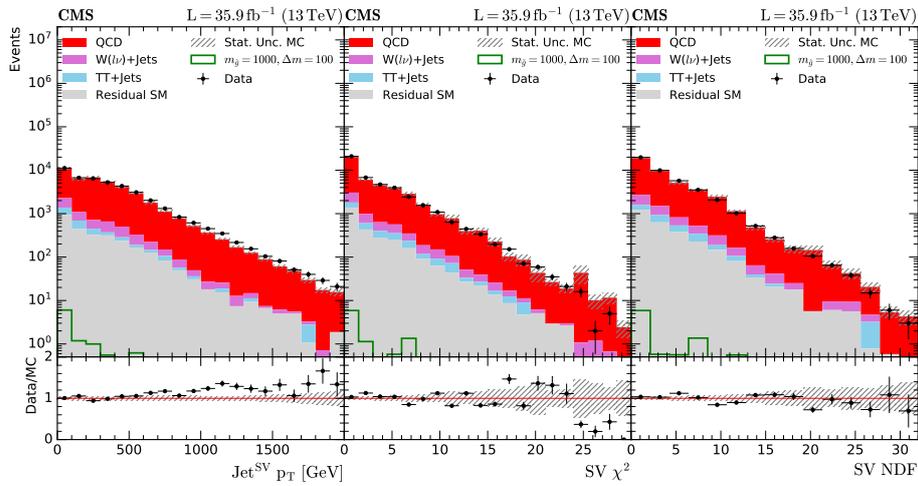


FIGURA C.19 – Validação das variáveis regionais na sub-região CR3 da região de QCD.

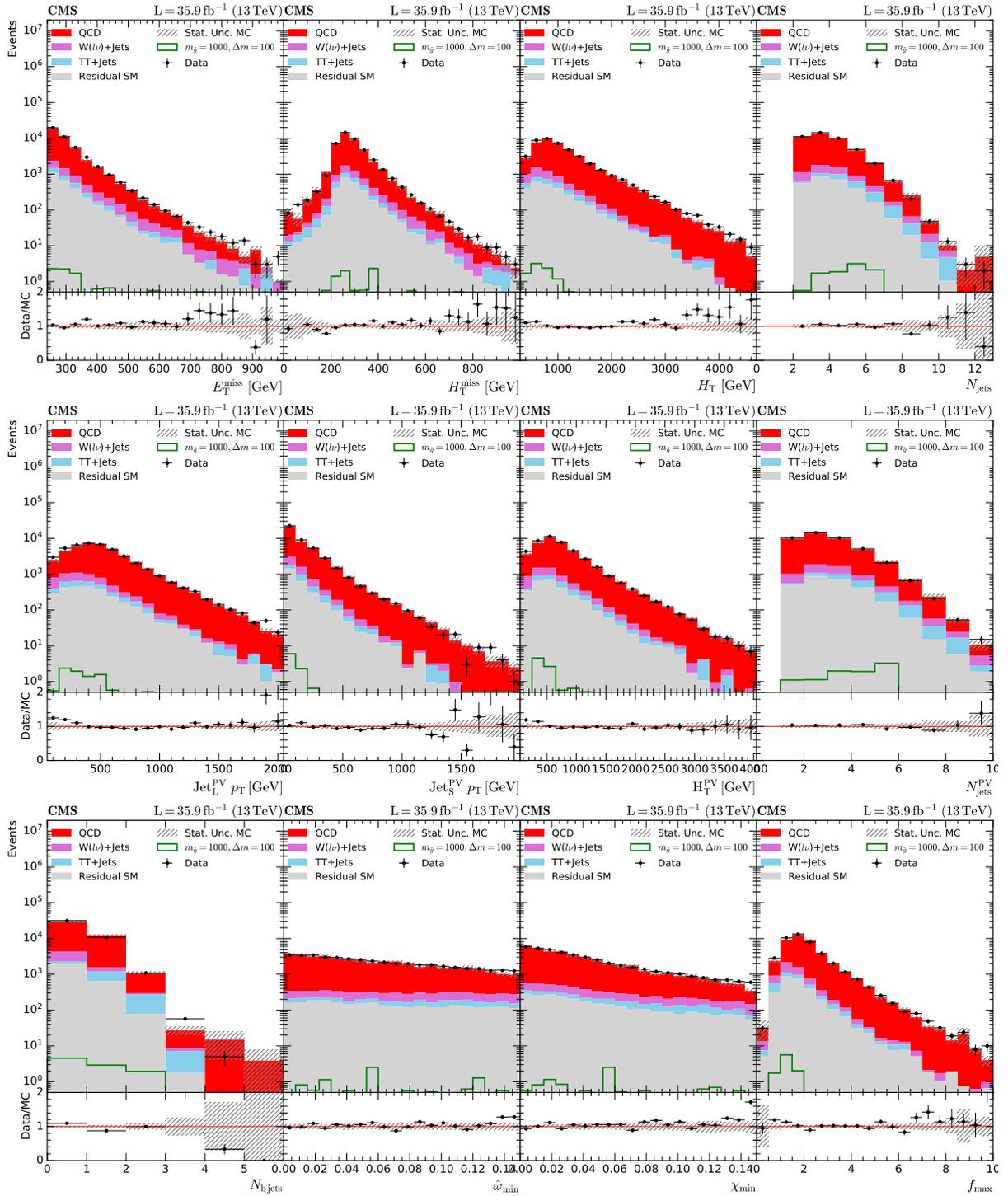


FIGURA C.20 – Validação das variáveis globais na sub-região CR3 da região de QCD.

C.3.4 Variáveis de CR4

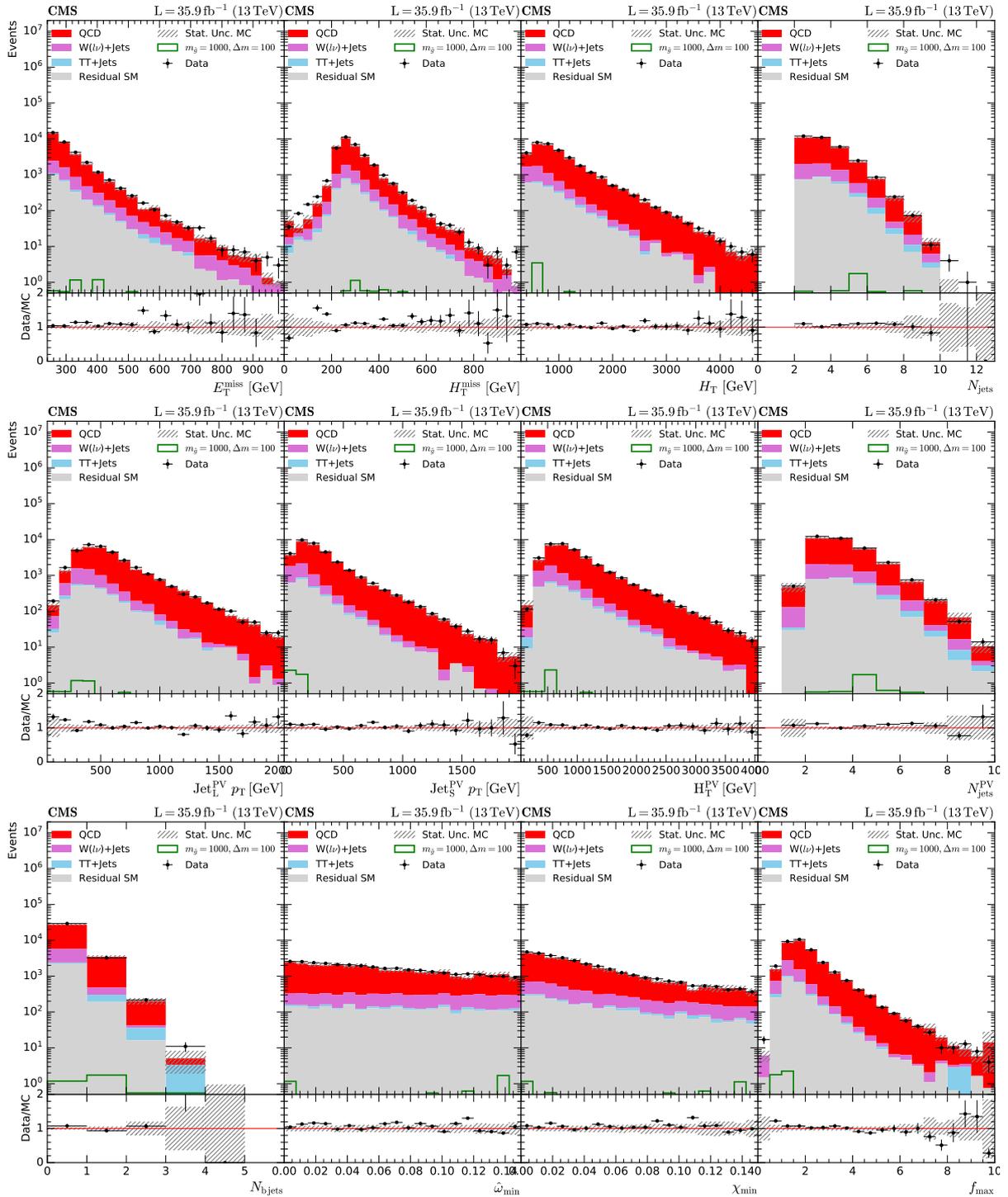


FIGURA C.21 – Validação das variáveis globais na sub-região CR4 da região de QCD.

Apêndice D - MVA Gráficos

Coletânea de gráficos de controle das análises multivariadas realizadas nas sub-regiões de busca. Os gráficos consistem de matrizes de correlação linear e distribuições¹ das variáveis de entrada para as classes de sinal e *background*, gráfico de evolução da função de perda, checagem de *overtraining* e avaliação da performance.

¹ Os nomes das variáveis que aparecem nos gráficos das distribuições das variáveis de entrada são aqueles utilizados no código da análise. A correspondência entre os nomes das variáveis globais é descrita na seção 5.4.1. A correspondência entre as variáveis regionais é dada a seguir:

- **R1:** $M_{JJ} \Leftrightarrow MJJ$, $Jet_L^{SV} p_T \Leftrightarrow MaxJetPt$, $Jet_S^{SV} p_T \Leftrightarrow MinJetPt$, $SV \chi^2 \Leftrightarrow Chi2$, $SVNDF \Leftrightarrow Ndof$, $\Delta\eta_{JJ} \Leftrightarrow DEtaJJ$ e $\Delta R_{JJ} \Leftrightarrow DRJJ$;
- **R2:** $Jet_1^{SV} p_T \Leftrightarrow JetPt1$, $Jet_2^{SV} p_T \Leftrightarrow JetPt2$, $M_R \Leftrightarrow MR$, $R^2 \Leftrightarrow MRT$, $SV_1 \chi^2 \Leftrightarrow Chi21$, $SV_1NDF \Leftrightarrow Ndof1$, $SV_2 \chi^2 \Leftrightarrow Chi22$ e $SV_2NDF \Leftrightarrow Ndof2$;
- **R3:** $Jet^{SV} p_T \Leftrightarrow JetPt$, $SV \chi^2 \Leftrightarrow Chi2$ e $SVNDF \Leftrightarrow Ndof$.

D.1 Gráficos de R1

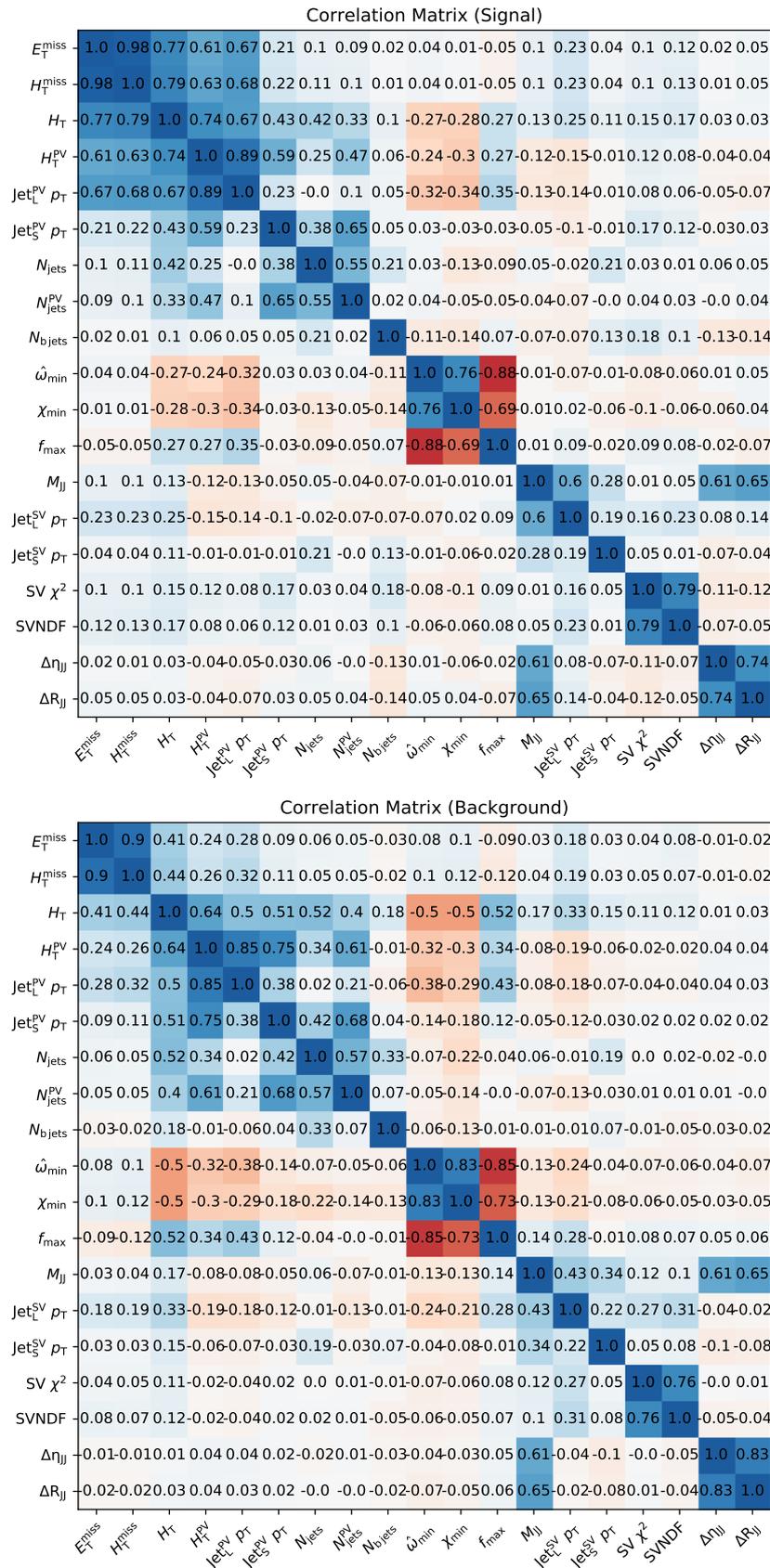


FIGURA D.1 – Matrizes de correlação linear das variáveis de entrada em R1.

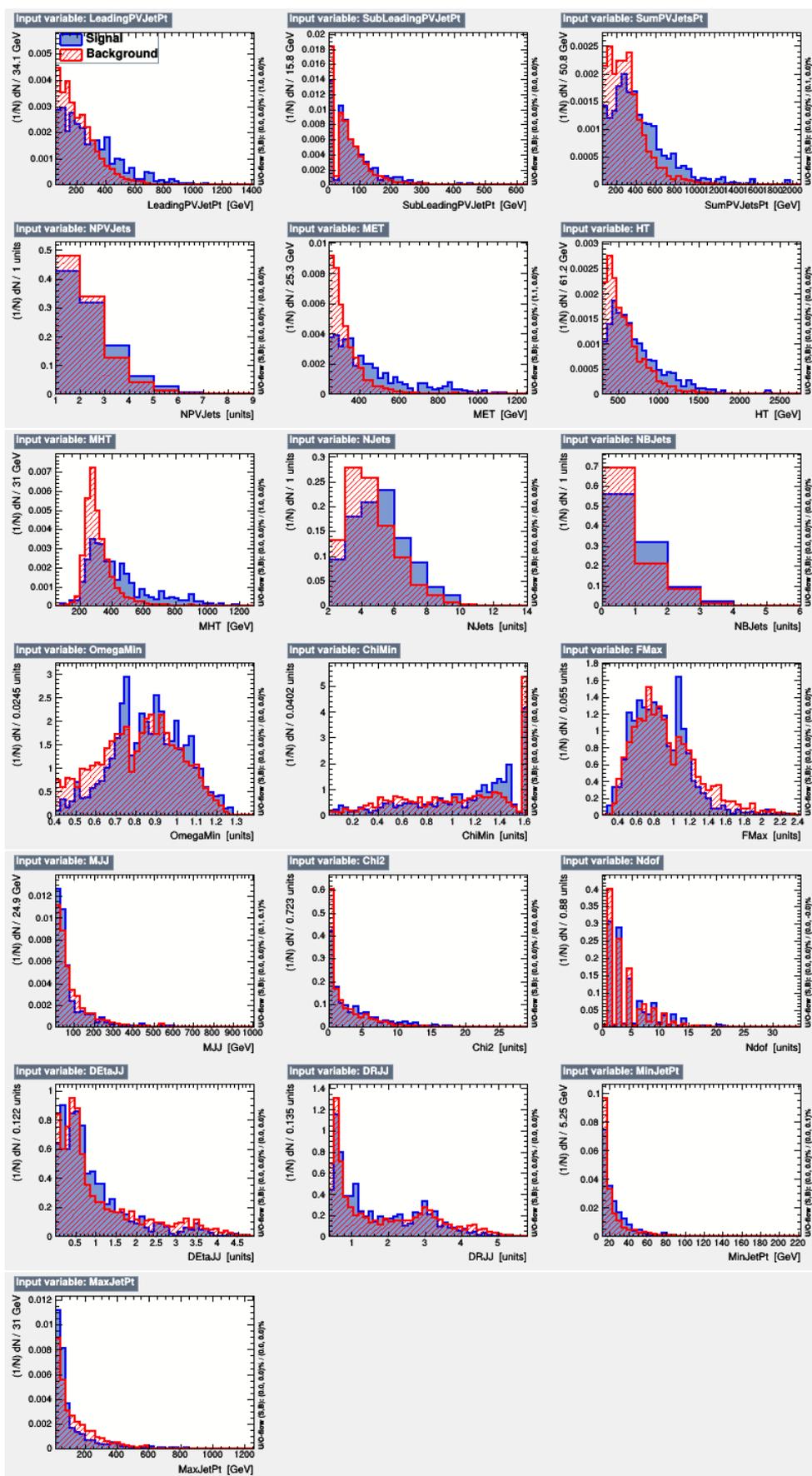


FIGURA D.2 – Distribuições das variáveis de entrada em R1.

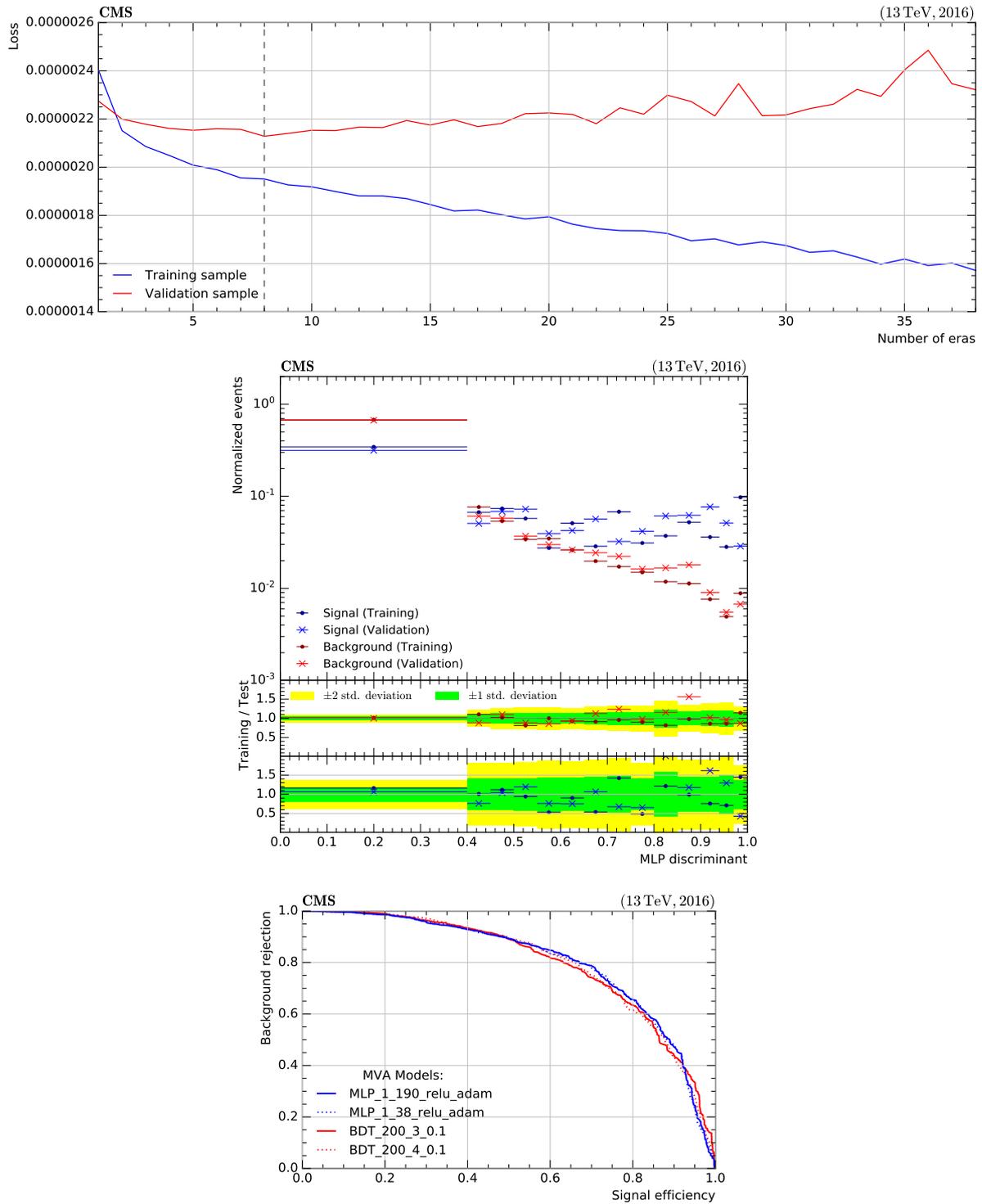


FIGURA D.3 – Gráfico de evolução da função de perda (acima), checagem de *overtraining* (centro) e avaliação da performance (abaixo) da RNA usada em R1.

D.2 Gráficos de R2

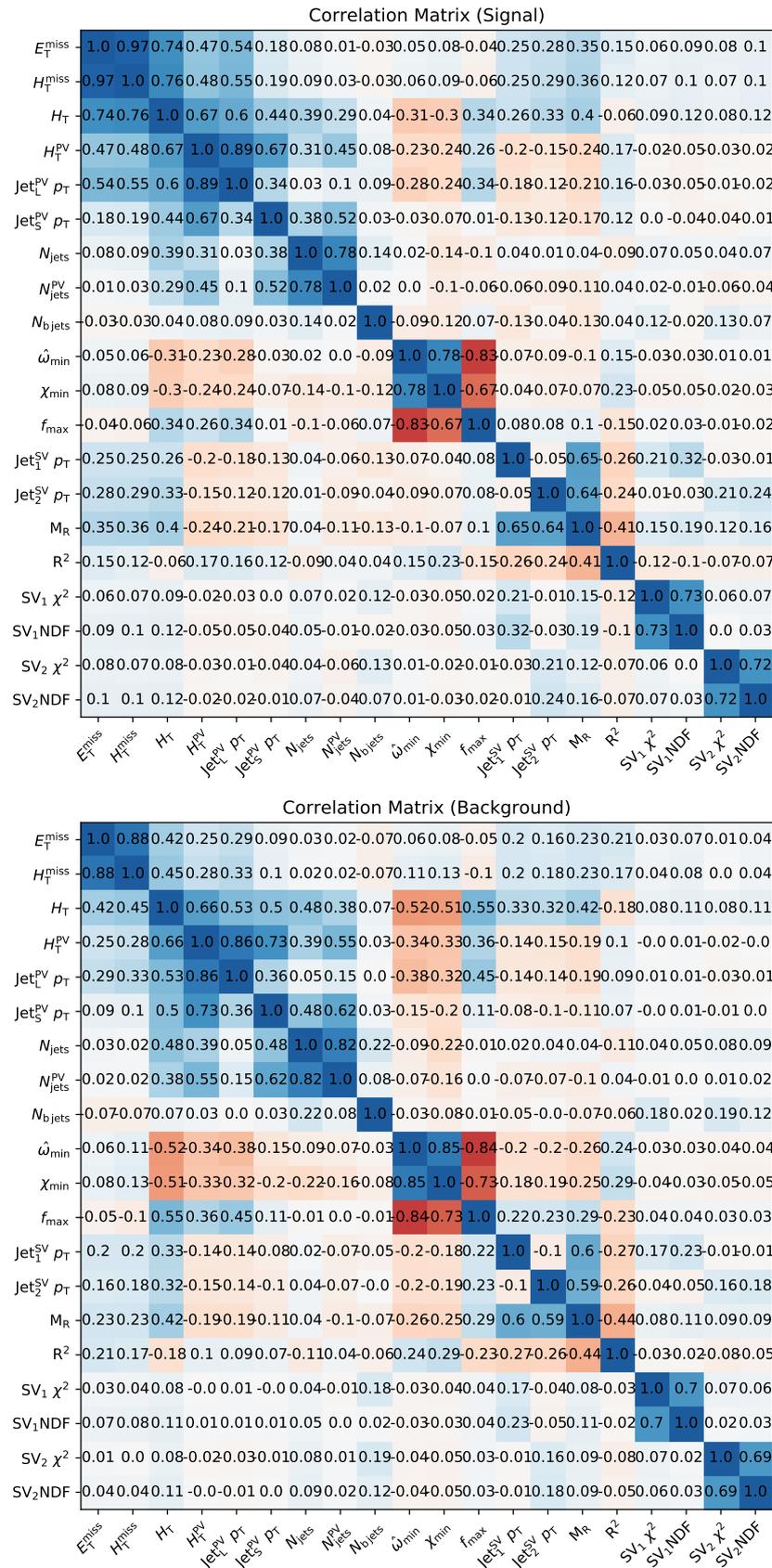


FIGURA D.4 – Matrizes de correlação linear das variáveis de entrada em R2.

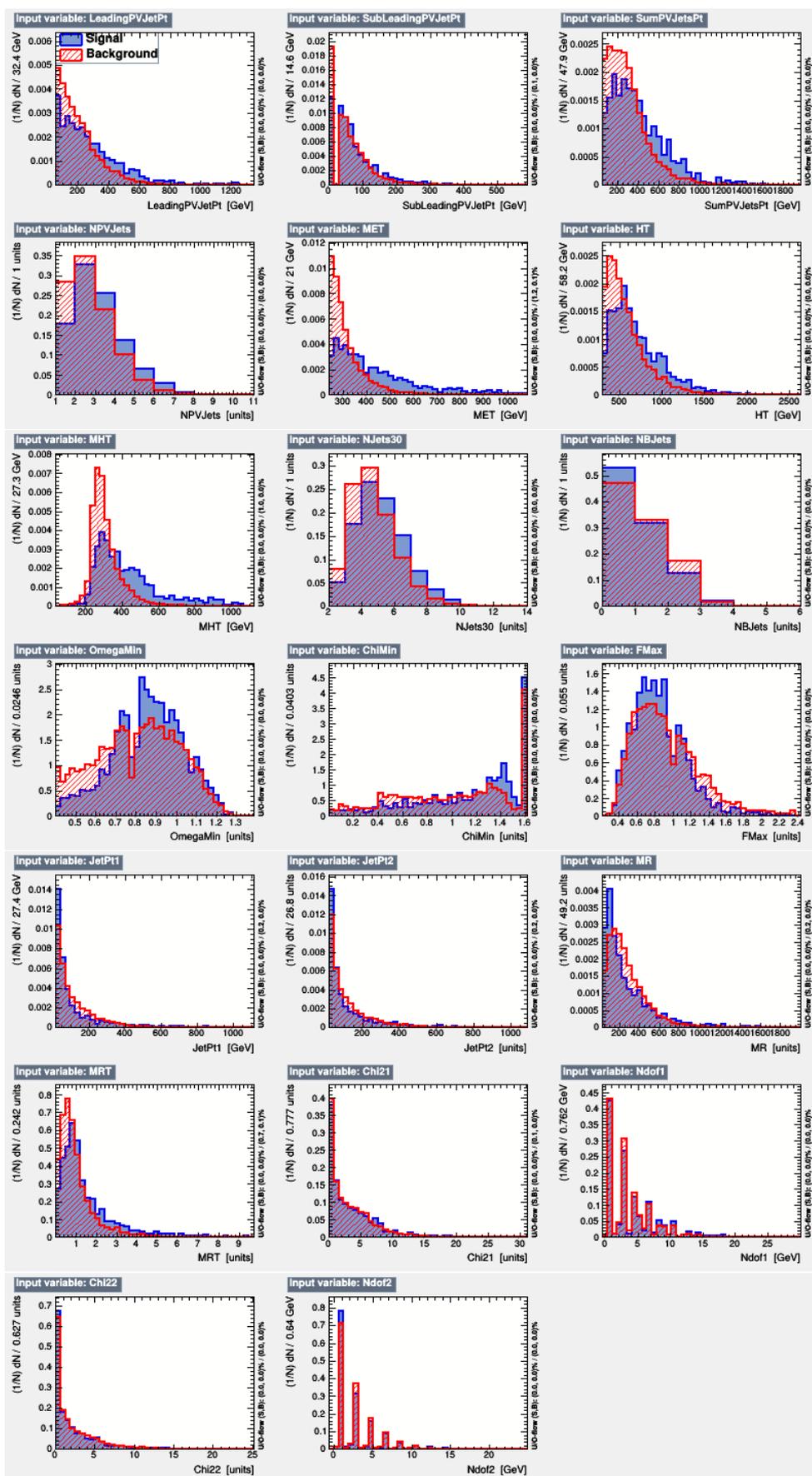


FIGURA D.5 – Distribuições das variáveis de entrada em R2.

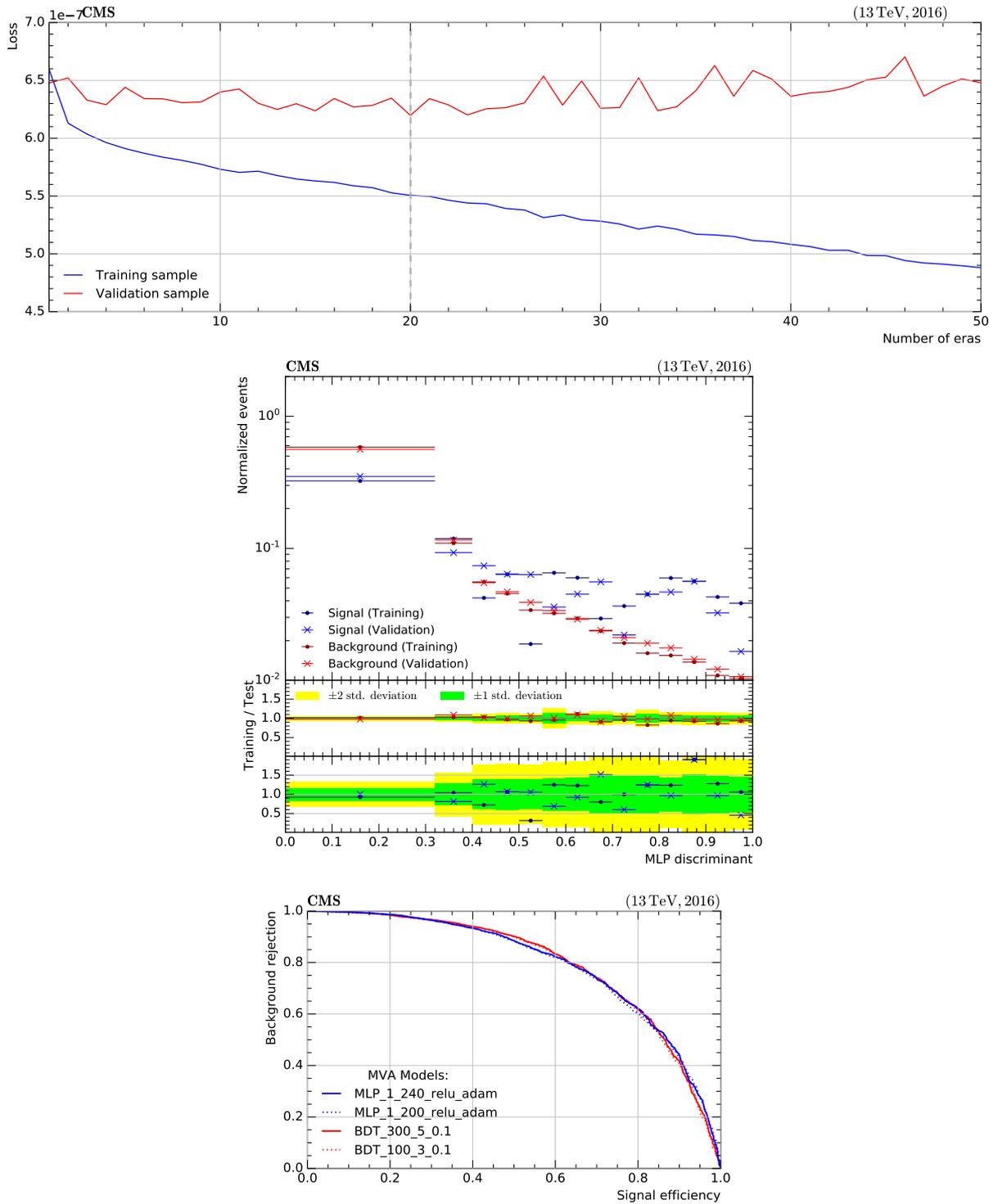


FIGURA D.6 – Gráfico de evolução da função de perda (acima), checagem de *overtraining* (centro) e avaliação da performance (abaixo) da RNA usada em R2.

D.3 Gráficos de R3

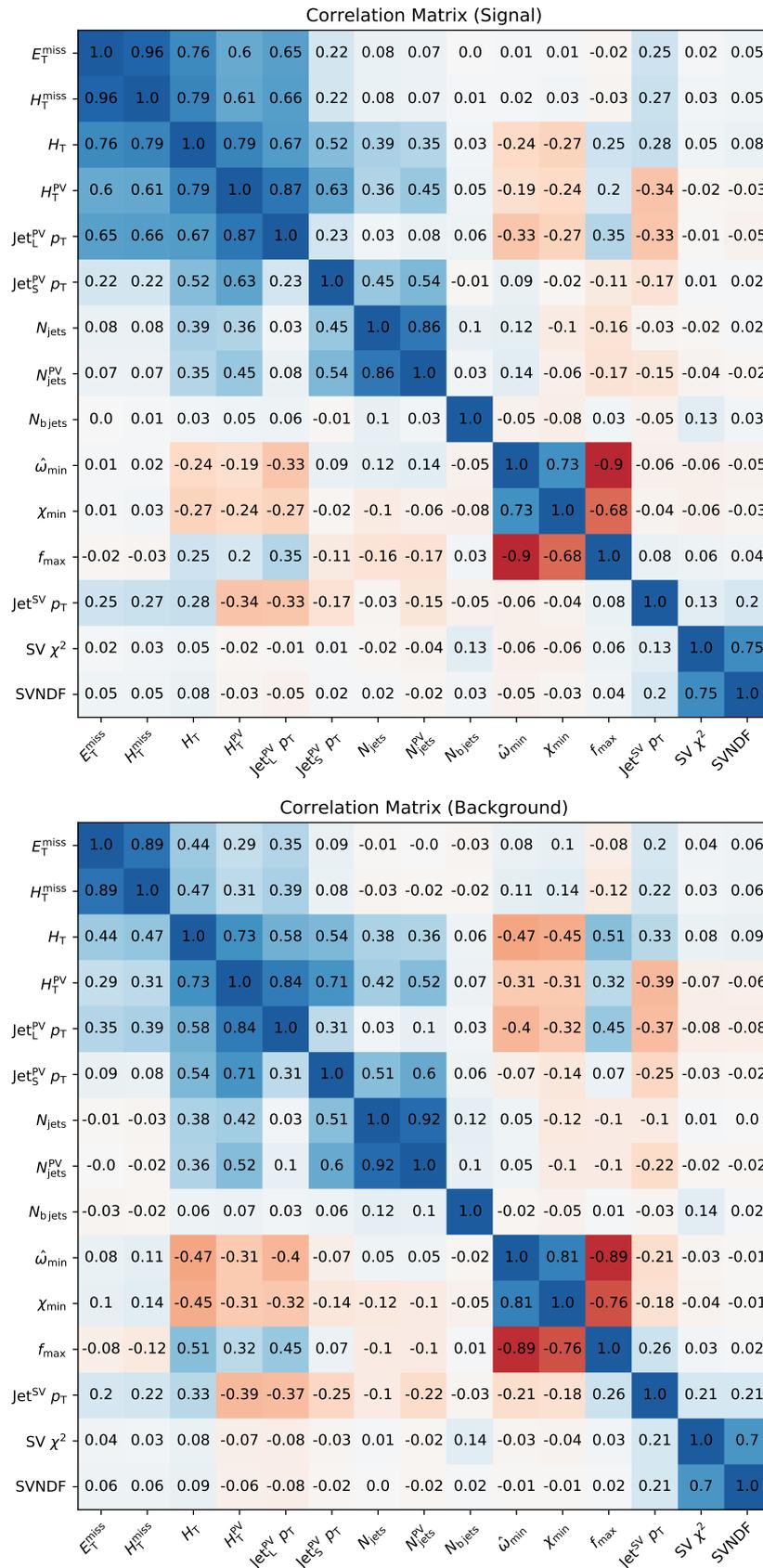


FIGURA D.7 – Matrizes de correlação linear das variáveis de entrada em R3.

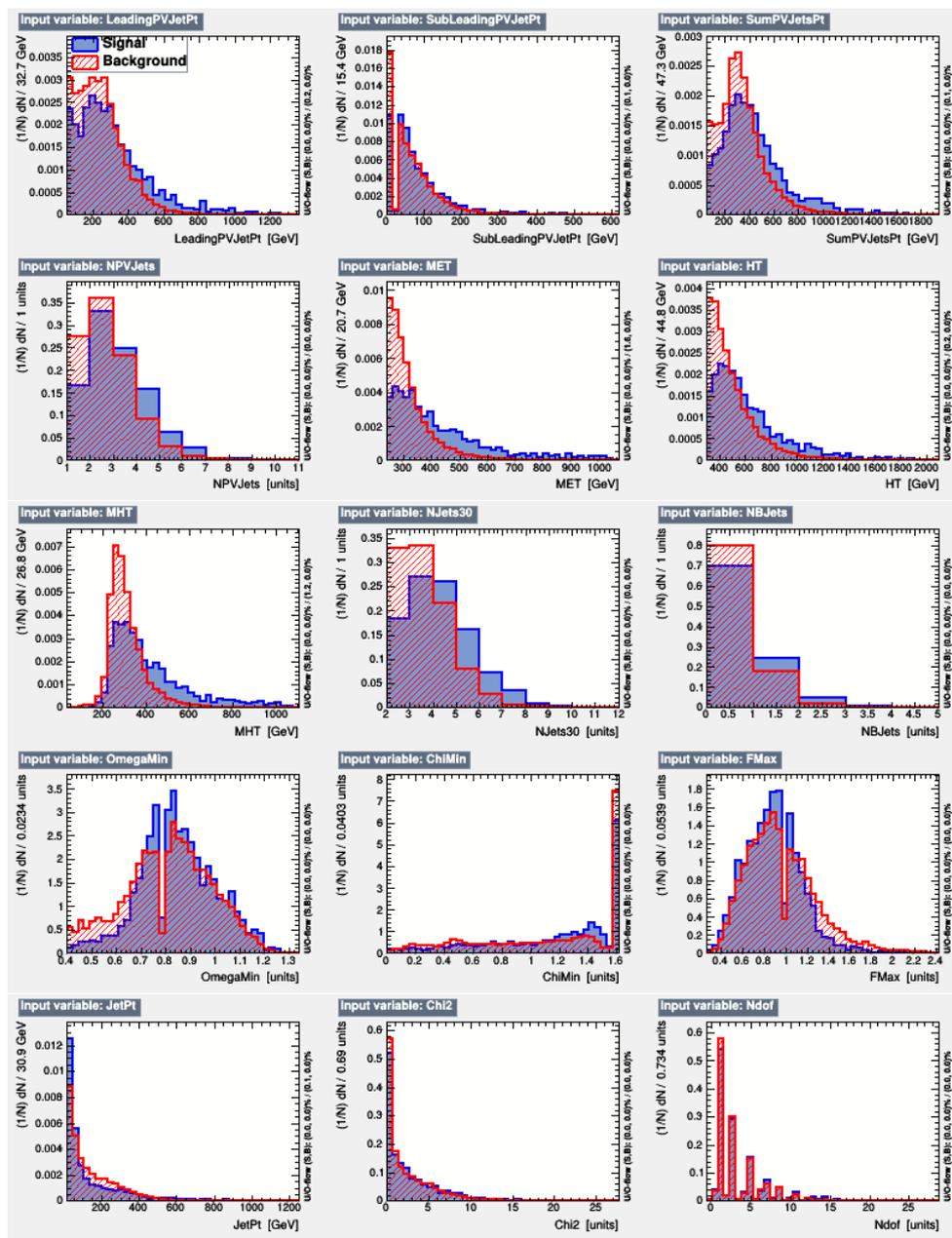


FIGURA D.8 – Distribuições das variáveis de entrada em R3.

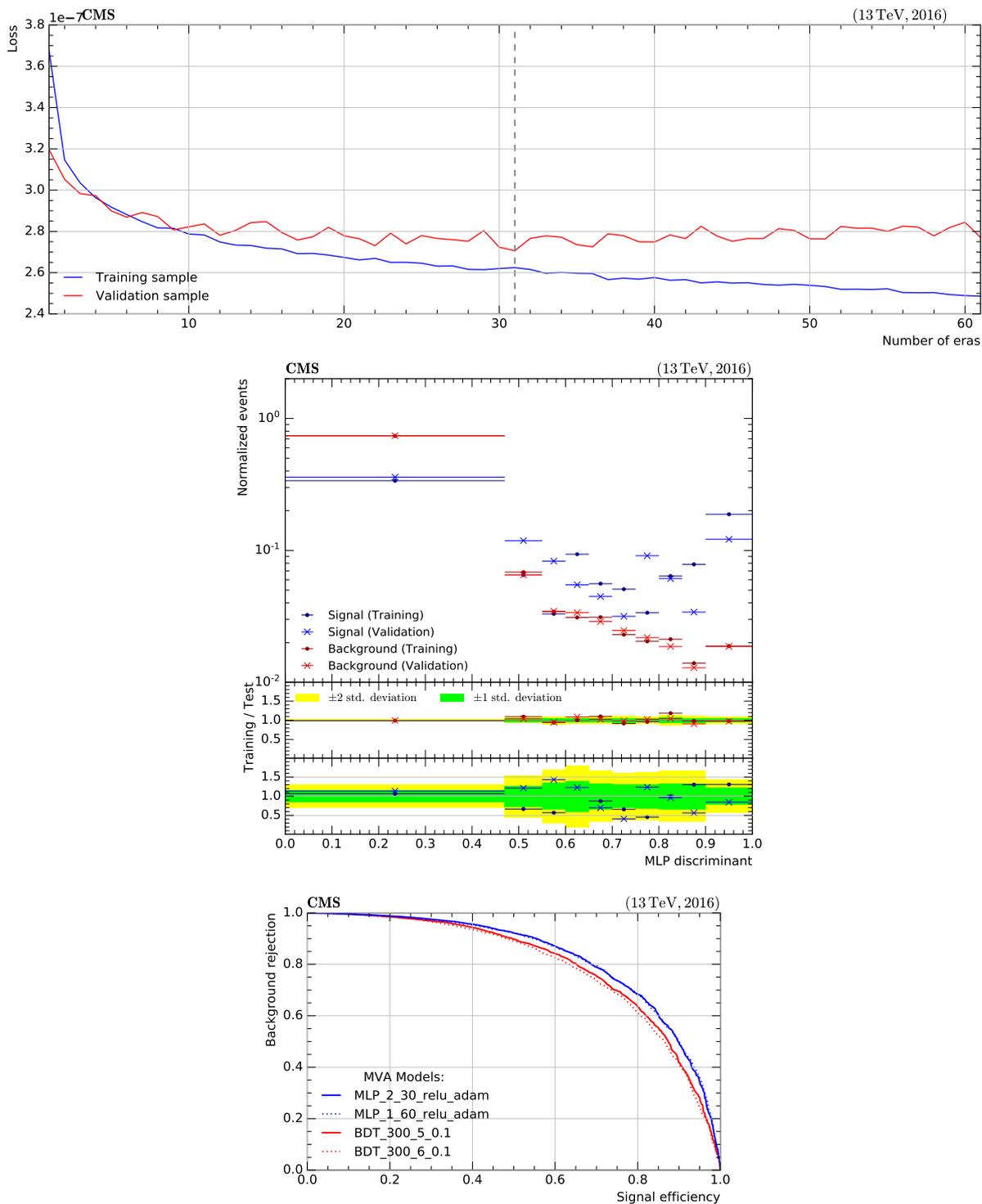


FIGURA D.9 – Gráfico de evolução da função de perda (acima), checagem de *overtraining* (centro) e avaliação da performance (abaixo) da RNA usada em R3.

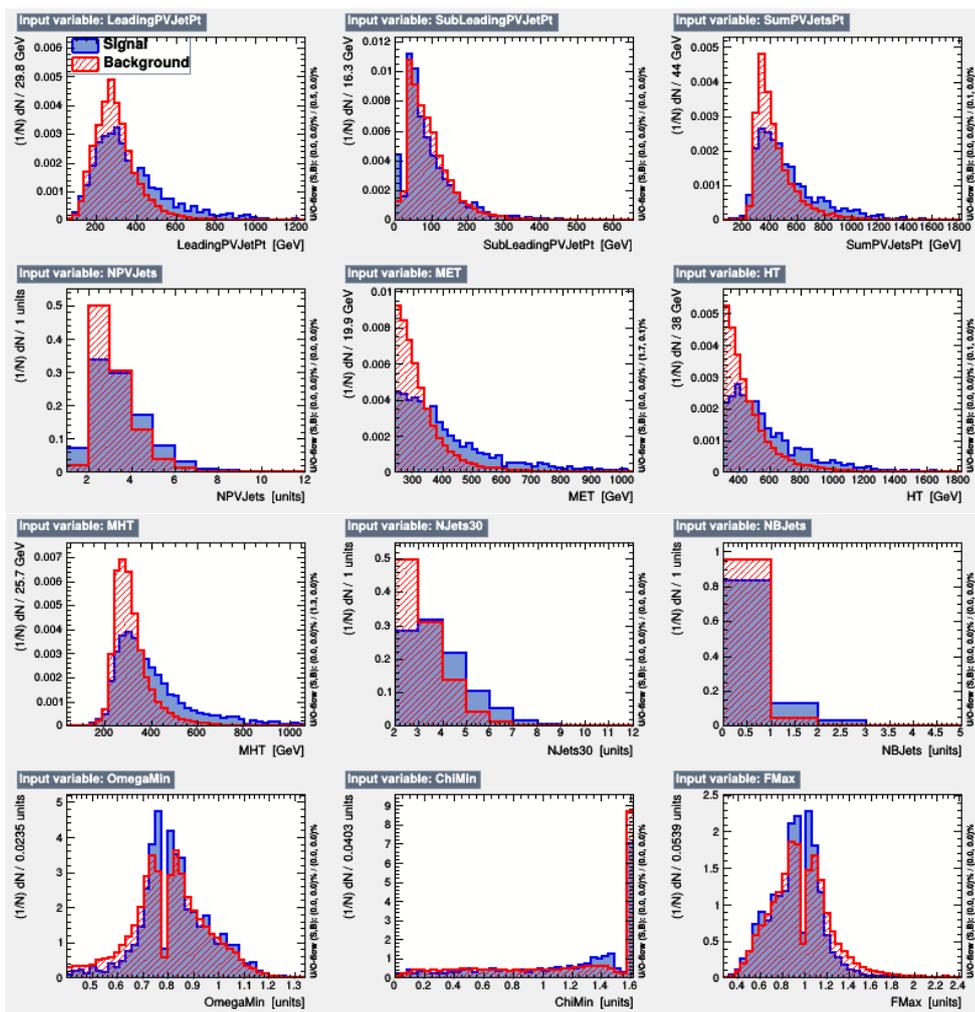


FIGURA D.11 – Distribuições das variáveis de entrada em R4.

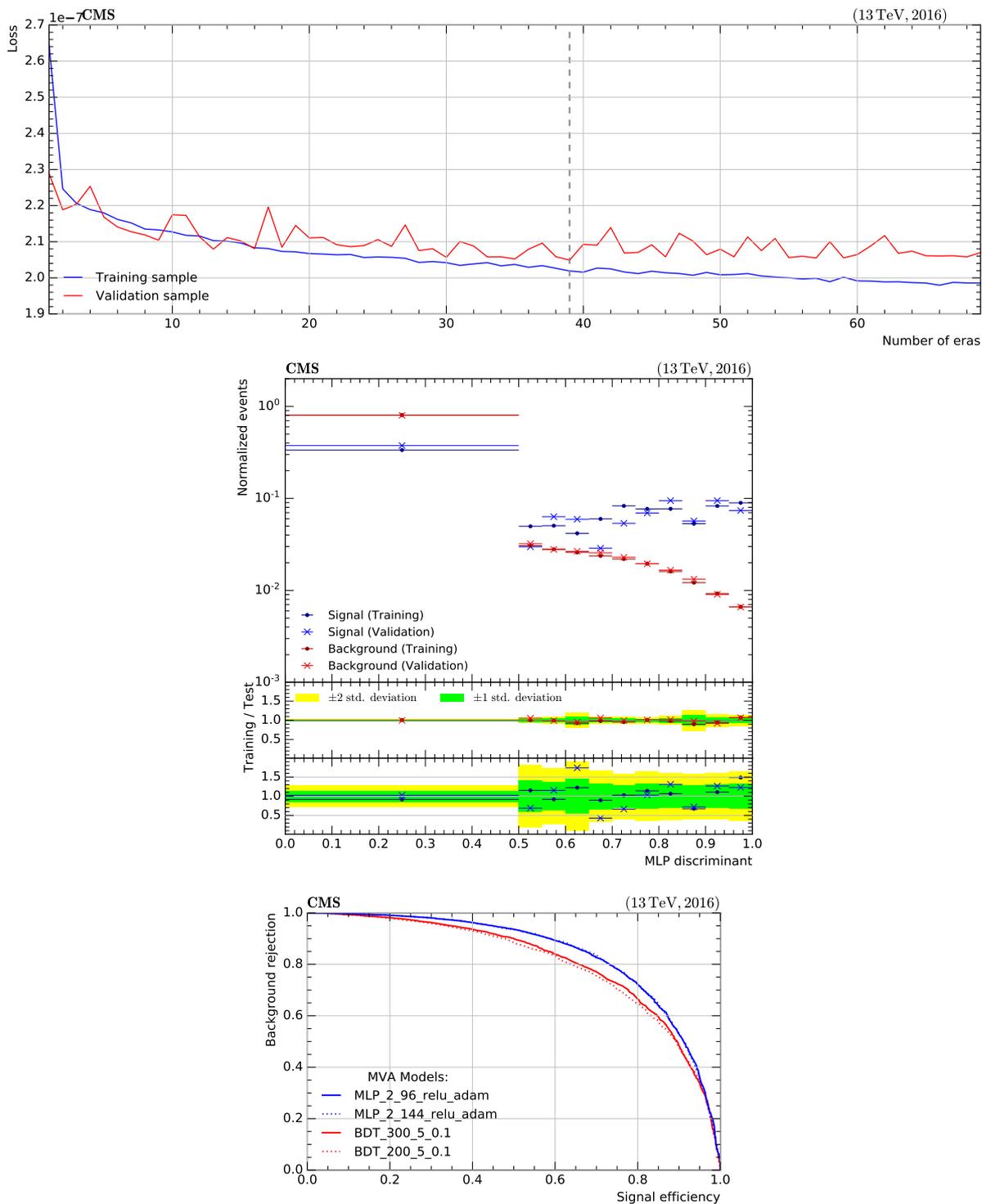


FIGURA D.12 – Gráfico de evolução da função de perda (acima), checagem de *overtraining* (centro) e avaliação da performance (abaixo) da RNA usada em R4.