Model Unspecific Search for New Physics with High p_T Photons in CMS

von

Stefan Antonius Schmitz

Diplomarbeit in Physik

vorgelegt der

Fakultät für Mathematik, Informatik und Naturwissenschaften
der Rheinisch-Westfälischen Technischen Hochschule Aachen

im Oktober 2009

angefertigt im

III. Physikalischen Institut A
Prof. Dr. Thomas Hebbeker

Abstract

In 2009 the LHC [1] collider at the European center of particle physics CERN [2] will start operations, colliding protons with a center of mass energy of up to 14 TeV. Designed as a large multi purpose detector CMS [3] will then start taking collision data. CMS will perform precision measurements within the Standard Model of particle physics and expand the search for new physical phenomena into regions that have not yet been probed by previous experiments.

Many theories about what physics beyond the Standard Model at the TeV scale might look like have been proposed. Together these models leave room for a broad spectrum of possible experimental signatures that one might look for in the data.

Various analyses focus on processing the available information with the aim of finding evidence for a specific model of choice. MUSiC as a **M**odel **U**nspecific **S**earch **i**n **C**MS provides a complementary approach by scanning the data for noteworthy deviations from the Standard Model expectation while making only basic assumptions about the nature of new physics. This is achieved by employing dedicated algorithms which evaluate several distributions within a number of specified event classes. Taking into account both statistical and systematic uncertainties potential deviations are analyzed and quantified in terms of significance.

A more detailed description of the necessary analysis steps is presented in this work. Aspects of event reconstruction and the processing of the data are outlined with a special focus on photon selection. The search algorithms, which estimate significances by the application of p-values, are explained and the treatment of systematic uncertainties is discussed. Statistical aspects of the search algorithms are evaluated emphasizing the importance of coverage properties of the employed p-values. Test scenarios with high energetic photons as a potential signature of new physics are presented in order to demonstrate the functionality and operational reliability of the chosen approach.

[1] Large Hadron Collider
[2] European Organization for Nuclear Research
[3] Compact Muon Solenoid

Zusammenfassung

Noch im Verlauf dieses Jahres wird der LHC [4]-Beschleuniger am CERN [5], dem europäischen Forschungszentrum für Teilchenphysik, in Betrieb genommen werden und dann Protonen mit einer Schwerpunktsenergie von bis zu 14 TeV zur Kollision bringen. Der CMS [6]-Detektor ist ein multifunktionaler Großdetektor, der dazu entwickelt wurde, anhand der aus den Kollisionen gewonnenen Informationen Messungen im Rahmen des Standardmodells der Teilchenphysik durchzuführen und gleichzeitig intensiv nach Signalen neuer Physik in Bereichen zu suchen, die bislang experimentell nicht zugänglich waren.

Es existieren verschiedene Modelle darüber, wie Physik jenseits des Standardmodells im TeV-Bereich aussehen könnte. Die Gesamtheit solcher Theorien bietet Raum für eine sehr große Zahl möglicher Signaltypen in den vom CMS-Experiment aufgezeichneten Daten. Während verschiedene Einzelanalysen in möglichst effizienter Weise nach Hinweisen auf ein einzelnes Modell oder bestimmte Parameterbereiche innerhalb eines Modelltyps fahnden, versucht MUSiC [7] als modellunabhängige Suche nach neuer Physik in CMS mit nur möglichst grundlegenden Annahmen die Daten auf statistisch signifikantes Auftreten unbekannter Effekte hin zu analysieren. Dies geschieht mit Hilfe von Suchalgorithmen die, basierend auf Kenntnissen über das Standardmodell, Verteilungen innerhalb bestimmter Klassen von Ereignissen untersuchen und unter Berücksichtigung statistischer und systematischer Unsicherheiten auf Abweichungen hin überprüfen.

Konzept und geplante Durchführung des Projektes werden aus verschiedener Sichtweise erläutert. Es werden Aspekte der Rekonstruktion von physikalischen Objekten behandelt, wobei hier ein besonderer Schwerpunkt auf der Selektion von Photonen liegt. Die verwendeten Suchalgorithmen, die Signifikanzen auf Basis der Berechnung von p-Werten bestimmen, werden vorgestellt. Auf die in die Analyse eingeschlossene Behandlung einzelner systematischer Unsicherheiten wird dabei gesondert eingegangen. Es werden statistische Aspekte im Kontext des verwendeten Suchalgorithmus behandelt, wobei insbesondere die coverage-Eigenschaften der verwendeten p-Werte diskutiert werden. Testszenarien mit hochenergetischen Photonen als möglicher Signatur neuer Physik werden vorgestellt um die Funktionalität des Ansatzes zu demonstrieren.

[4] Large Hadron Collider
[5] European Organization for Nuclear Research
[6] Compact Muon Solenoid
[7] **M**odel **U**nspecific **S**earch **i**n **C**MS

Contents

1 Introduction

Though the Standard Model of physics which describes the world in terms of interacting particle fields provides today's physicists with a highly successful theory, many fundamental questions remain unanswered. On the one hand the "Higgs" boson, an important ingredient of the theory, has not yet been discovered. On the other hand there are both experimental and theoretical arguments suggesting that the Standard Model is certainly far from being a consistent theory of everything. Some of the known problems and models that have been proposed to address them include:

- There is convincing evidence from the interpretation of astrophysical data in terms of Einstein's theory of general relativity suggesting that our current knowledge of the universe seems to be restricted to a small fraction of its energy content. It is found that in terms of gravity a substantial part of the unknown energy seems to be interacting like the matter we know but is non-baryonic and has so far eluded any direct observation in dedicated experiments. Furthermore it is completely unknown in terms of non-gravitational interactions. Several theories predict weakly interacting massive particles as the origin of this "dark matter".

- Some of the more abstract problems encountered in the Standard Model can be addressed starting from the postulation of an additional symmetry between bosons (particles with integer spin) and fermions (particles with half integer spin). The resulting "supersymmetric" models create a highly complex theoretical playground including possibilities of discovery at the LHC [1].

- It is an important feature of the Standard Model that it does not include gravity. At which energy scale this shortcoming will start to be experimentally accessible is not clear but there are theories which allow for observable effects at the TeV scale. Models with added extra dimensions are one prominent example incorporating such possibilities.

- From today's perspective the values of fundamental parameters like masses and charges are found to be set to rather arbitrary values seemingly lacking any underlying principle. Presuming that there is scientific value in arguments which might be considered to be of mainly aesthetic origin one may conclude that this situation is somehow unsatisfactory.

Of course what can be said in few sentences is only capable of giving a superficial and highly fragmentary impression of the many possibilities ahead of us, but should suffice to convey that there is still a lot of undiscovered territory in particle physics. And indeed we are confronted with a vast landscape of competing ideas and models for physics beyond the Standard Model allowing for a large spectrum of conceivable experimental signatures. Of course one should also leave room for the possibility that at very high energies nature might behave differently from everything what physicists have been able to hypothesize.

After years of planning and construction the LHC will soon start operations and begin to collide beams of protons with a center of mass energy of up to 14 TeV. The CMS [2] experiment which is situated at one of the collision points will then contribute to new precision tests within the realm of known physics and hopefully provide hints towards new physics beyond the scope

[1] Large Hadron Collider
[2] Compact Muon Solenoid

of the Standard Model. Within the CMS project enormous amounts of data about potentially interesting proton-proton collisions will be stored. On the basis of this data a multitude of dedicated analyses will start to look for evidence of experimental signatures motivated by their respective model of choice. It has been conjectured that the "entire mass of data would thus be likely to be analyzed only in the particular way that addresses questions that they are concerned with. It is certainly possible that there are many clues to Nature's ways hidden in such data, even if we do not properly read them yet" [1, page 1025].

While Roger Penrose was probably thinking in a somewhat different direction when formulating this statement one should consider that the same state of technology that allows for an automatized storage of massive amounts of data also provides the means for an automatized approach to data analysis relying on the availability of high computational capacities. Hence unexpected hints towards new physics do not necessarily need to be triggered by new ingenious models. Instead they might be discovered by clever ways of systematically scanning large fractions of the recorded data in order to find deviations from the expectations on the basis of the Standard Model. Through the automation of many steps in this process it is possible to address a range of possible signatures that is beyond what can reasonably be covered by dedicated analyses. It is this idea of a "Model Unspecific Search" to aspects of which this work is dedicated. The MUSiC [3] analysis [2] comprises the necessary efforts related to the realization of such a model unspecific analysis in the context of the CMS experiment.

A first proposal of such a kind of search strategy has already been formulated in the context of the L3 experiment [3]. Since then detectors at different collider experiments have conducted model unspecific searches (e.g. [4] and [5], corresponding recent reports are presented in [6] and [7]). Though some interesting deviations were found none of these efforts has resulted in convincing evidence of physics beyond the Standard Model.

But at CMS the existing arguments for possible discoveries at the TeV scale combined with the rather fuzzy expectations of what exactly should be looked for set the stage for an experimental scenario which might be in favor of model unspecific enterprises.

It should also be mentioned that this is the first time that such a search strategy will be a part of the analysis process within a large scale experiment right from the beginning of data taking. During the challenging period of early data a broad scan for unexpected deviations as conducted within the MUSiC framework may contribute to the global understanding of the detector and help to understand aspects of the relation between event simulation and data. This aspect is however not within the scope of what is discussed in the following chapters but has been addressed in some detail elsewhere [2].

In order to arrive at a legitimate assessment of the significance of observed deviations a sound evaluation of the statistical properties of the applied algorithms is indispensable. It is also crucial to develop some understanding of potential shortcomings and pathological situations that might either lead to an overestimation of significances or obfuscate potentially interesting effects. Thus several topics related to statistical aspects are addressed in this work including coverage tests for the applied p-values. A variation of the original search algorithm using a different estimator of significance is discussed and tested. Subsequent to the necessary conceptual and practical preparations simulated test scenarios are analyzed focusing on experimental signatures with high energy photons. Signatures motivated by the theoretical frameworks of Gauge Mediated Supersymmetry Breaking, Compositeness and Large Extra Dimensions are evaluated to provide a proof of concept.

[3] Model Unspecific Search in CMS

2 A Synopsis of the Standard Model of Particle Physics

It is known that the Standard Model of particle physics, which is outlined in this chapter, cannot provide a valid description of our universe at all energy scales. But up to now no unambiguous deviations from the theory have been observed[1]. Accordingly, it is the Standard Model against which any potential claim of "new physics" needs to be tested. Of course, this also holds true for a model unspecific analysis approach. A selection of important concepts of the Standard Model is outlined in the following sections (2.1 to 2.5). It should be mentioned that the presented approach is not the only possibility to introduce the Standard Model[2].

Each particle in a Quantum Field Theory corresponds to a (quantum) field. The particle content of the Standard Model is summarized in figure 2.1. All particles can be subdivided into fermions (half-integer spin) and bosons (integer spin). The fermionic content is given by three families of quark and lepton fields ψ with spin $\frac{1}{2}$. Additionally there are gauge vector bosons A_μ, which can be seen as force mediators, and a scalar field ϕ, the "Higgs" field, which is responsible for the generation of mass terms.

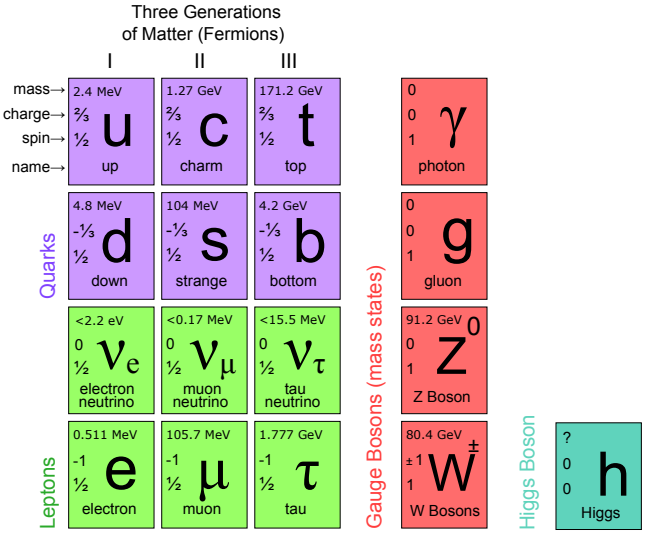

Figure 2.1 Particle content of the Standard Model [10, modified version]

2.1 Gauge Field Theory

The concept of gauge invariance plays a crucial part in the construction of the Standard Model. A Lagrangian can be used to define the particle content and the interactions of a quantum field theory. In the following sections the strong and electroweak interactions of the Standard Model are constructed by applying the principle of local gauge invariance to an initial Lagrangian.

[1] Here it is assumed that neutrino oscillations (see e.g. [8]) can be incorporated into the Standard Model and do thus not pass for physics beyond the Standard Model.

[2] An alternative approach would be to use the path integral formalism (see e.g. [9]).

In practice, one starts from the free fermion Lagrangian, which describes propagating fermions without interactions. It is given by

$$\mathscr{L} = \bar{\psi}\left(i\partial\!\!\!/ - m\right)\psi \tag{2.1}$$

with $\bar{\psi} = \psi^\dagger\gamma_0$ and $\partial\!\!\!/ = \gamma_\mu\partial^\mu$. This ansatz can be motivated by the resulting field equations for $\bar{\psi}$ which are the Dirac equation for ψ.

Now one considers unitary local gauge transformations $U(x)$ of the form

$$\psi(x) \to \psi'(x) = U(x)\psi(x) \quad . \tag{2.2}$$

For a given symmetry group G with generators T^a, $U(X)$ is given by [11]

$$U = e^{i\alpha_a(x)T^a} \quad , \quad a = 1\ldots N \quad \text{with} \quad \left[T^a, T^b\right] = if^{abc}T^c \quad , \tag{2.3}$$

where f^{abc} are structure constants of G. N is the number of degrees of freedom of the group. If the generators T^a commute, the gauge symmetry is called "Abelian". Otherwise the gauge symmetry is called "non-Abelian". Now one requires $\mathscr{L}$ to be invariant under transformations $U(X)$. This criterion is not met by the free Lagrangian (2.1) but can be restored by replacing ∂_μ with [12]

$$\partial_\mu \to \mathbf{D}_\mu = \partial_\mu - ig\mathbf{A}_\mu \quad , \quad \mathbf{A}_\mu = A^a_\mu T^a \quad . \tag{2.4}$$

Here we have introduced additional fields A^a_μ. During the quantization of the theory these will become the "gauge bosons" of the Standard Model which describe the electroweak and strong interactions. If one fixes the transformation properties of A^a_μ via

$$\mathbf{A}'_\mu = U\mathbf{A}_\mu U^{-1} - \frac{i}{g}\left(\partial\mu U\right)U^{-1} \quad , \tag{2.5}$$

the Lagrangian

$$\mathscr{L}' = \bar{\psi}\left(i\not{\mathbf{D}} - m\right)\psi \tag{2.6}$$

is invariant under the transformations $U(x)$. In this way the dynamics of the model is described by the gauge fields A^a_μ which are induced by the postulation of local gauge invariance with respect to the group G. If one wants to describe the propagation of gauge fields, a kinetic term of the form $\mathscr{L}_{kin} = 1/2 \cdot tr\left\{F_{\mu\nu}F^{\mu\nu}\right\}$ needs to be added [12]. Including this term does not break the gauge symmetry. The field strength tensor $F_{\mu\nu}$ for the gauge fields is given by

$$F_{\mu\nu} = -\frac{i}{g}\left[D_\mu, D_\nu\right] \quad . \tag{2.7}$$

The resulting full Lagrangian respecting the gauge symmetry is the gauge invariant "Yang-Mills" Lagrangian

$$\mathscr{L}_{\mathrm{YM}} = -\frac{1}{2}tr\left\{F_{\mu\nu}F^{\mu\nu}\right\} + \bar{\psi}\left(i\not{\mathbf{D}} - m\right)\psi \quad . \tag{2.8}$$

Lagrangians of the form $\mathscr{L}_{\mathrm{YM}}$ can be used to describe all electroweak and strong interactions in the Standard Model. The corresponding gauge symmetries are $SU(3)_{\mathrm{color,C}}$ for the strong interactions (Quantum Chromodynamics) and $SU(2)_{\mathrm{weak\,isospin,W}} \times U(1)_{\mathrm{hypercharge,Y}}$ for the electroweak interactions. Hence the full gauge symmetry of the Standard Model is

$$SU(3)_C \times SU(2)_W \times U(1)_Y \quad . \tag{2.9}$$

2.2 Theories with Quantized Fields

In section 2.1 it has been described how gauge symmetry can be used to introduce the interactions of the strong and electroweak forces. However, the Lagrangian is defined in terms of classical fields. If the model is meant to have characteristics of a quantum theory, the fields need to be quantized. This can be achieved by elevating the classical fields to field operators. The field operators are defined in terms of creation- and annihilation operators which act on the Fock space of the model. For example a real scalar field is quantized via

$$\hat{\phi}(x) = \int \frac{d^3k}{(2\pi)^3 2k_0} \left[\hat{a}(k)\, e^{i\vec{k}\,\vec{x}-ik_0t} + \hat{a}^\dagger(k)\, e^{-i\vec{k}\,\vec{x}+ik_0t}\right] \tag{2.10}$$

with creation operators $\hat{a}^\dagger(k)$ (create particle with momentum k) and annihilation operators $\hat{a}(k)$ (annihilate particle with momentum k). Details concerning the quantization of fields and the formalism of creation and annihilation operators can be found, for example, in [13]. At equal times one postulates canonical commutation relations for scalar fields χ_b and canonical anticommutation relations for fermion fields χ_f :

$$\begin{aligned}
\{\pi_{f\,\alpha}(t,\vec{x}), \chi_{f\,\beta}(t,\vec{x}\,')\} &= -i\delta_{\alpha\beta}\delta^3(\vec{x}-\vec{x}\,') \\
\{\chi_{f\,\alpha}(t,\vec{x}), \chi_{f\,\beta}(t,\vec{x}\,')\} &= \{\pi_{f\,\alpha}(t,\vec{x}), \pi_{f\,\beta}(t,\vec{x}\,')\} = 0 \\
[\pi_b(t,\vec{x}), \chi_b(t,\vec{x}\,')] &= -i\delta^3(\vec{x}-\vec{x}\,') \\
[\chi_b(t,\vec{x}), \chi_b(t,\vec{x}\,')] &= [\pi_b(t,\vec{x}), \pi_b(t,\vec{x}\,')] = 0
\end{aligned} \tag{2.11}$$

with the canonical momentum

$$\pi_\alpha = \frac{\partial \mathscr{L}}{\partial \dot{\chi}_\alpha} \quad . \tag{2.12}$$

The commutation relations are postulated in such a way that the important spin-statistics theorem is a direct consequence.

In collider experiments one is interested in the cross sections of processes relating initial states with n particles $|p_1,\ldots,p_n\rangle_\mathrm{i}$ to final states $|p'_1,\ldots,p'_m\rangle_\mathrm{f}$ with m particles. The respective cross sections and decay widths are closely linked to the "S-matrix" which (in a somewhat simplified notation) relates initial and final states via [11, 13]

$$\begin{aligned}
S_{fi} = {}_\mathrm{f}\langle p'_1,\ldots,p'_m \,|\, p_1,\ldots,p_n\rangle_\mathrm{i} &= {}_\mathrm{i}\langle p'_1,\ldots,p'_m\,|S|\,p_1,\ldots,p_n\rangle_\mathrm{i} \\
&= (2\pi)^4\,\delta^4\left(\sum_\mathrm{f} p'_f - \sum_\mathrm{i} p_i\right) i\mathscr{M} \quad .
\end{aligned} \tag{2.13}$$

Starting from the quantized fields and making use of the Wick theorem and the LSZ-formalism (see e.g. [13, 14]) one can derive the Feynman rules which allow for a perturbative expansion of the matrix element $\mathscr{M}$ in the coupling constant g of the gauge group. The strategy is to find all possible Feynman diagrams up to a certain order in the coupling constant and to derive the corresponding matrix element $\mathscr{M}$ with the help of the Feynman rules. Then one can calculate cross sections and decay widths which are proportional to $|S_{fi}|^2$ (or equivalently $|\mathscr{M}|^2$).

Unfortunately it turns out that for the quantum field theory describing the Standard Model there are Feynman diagrams adding divergent contributions to $\mathscr{M}$ if one tries to calculate higher order corrections. For the Standard Model the divergences can be balanced by the process of "renormalization". The idea is to isolate all divergent contributions and absorb them into a scale-dependent redefinition of masses, charges and wave functions [12]. This is possible, because the respective parameters in the Lagrangian are not observable [11]. Consequently, one arrives at finite results for physical observables like cross sections. For a coupling constant g the running of g in dependence on the scale parameter μ and a cutoff scale Λ is governed by its respective beta function. The theory is "asymptotically free" if the beta function has an ultraviolet fixed

point meeting certain criteria. This implies that even if the value of the coupling constant is large at low energies, there is a range of high energies for which one can calculate results in a perturbative regime.

The Standard Model is a "renormalizable" quantum field theory as the procedure of renormalization can be proven to work at all orders of perturbation theory. However, one may wonder how fundamental this feature really is for the construction of sensible quantum field theories (see e.g. [15]). Notably, all the specified models of physics beyond the Standard Model that are discussed in chapter 3 are not renormalizable.

2.3 Quantum Chromodynamics

The gauge group (see section 2.1) describing the strong interactions of Quantum Chromodynamics (QCD) is $SU(3)_C$. The $SU(3)_C$ transformations act on fermion fields with an additional color index. According to the gauge group one has to introduce eight gauge fields, the so-called "gluons", corresponding to the $3^2 - 1 = 8$ generators of $SU(3)$. The gluons have color charge and do thus not only couple to the quarks, but also to themselves. This is a behavior that is unique to quantum field theories with non-Abelian gauge symmetries. The fermions with the additional degree of freedom related to color are the "quarks". There are three generations of quarks. Each generation contains two quarks which differ both in their electromagnetic charge Q and their respective mass parameters. This results in the overall number of six quark "flavors" (see figure 2.1). During strong interactions the color of quarks is changed, but the quark flavor is kept invariant. QCD has the property of asymptotic freedom, which has been introduced in section 2.2. This implies that the coupling constant decreases for high energies. However, QCD is non-perturbative in the low energy regime.

Another important aspect of QCD is color confinement. Quarks and gluons cannot be observed as single particles but are confined in colorless states with either two or three (valence) quarks, the mesons and baryons. Though this property has not yet been rigorously proven, color confinement is in agreement with all experimental results.

Direct observational evidence of the concept of color is for example given by the measurement of the ratio between the hadronic cross section and the di-muon cross section at a $e^+\,e^-$ collider, because at leading order one can use the relation

$$R = \frac{\sigma\left(e^+\,e^- \to \text{hadrons}\right)}{\sigma\left(e^+\,e^- \to \mu^+\,\mu^-\right)} = N_C \cdot \sum_{\substack{q \text{ with} \\ 2m_q < E_{\text{CM}}}} Q_q^2 \quad . \tag{2.14}$$

N_C is the number of quark colors. The measurements are in good agreement with the Standard Model assumption that $N_C = 3$ (see e.g. [16]).

2.4 Electroweak Theory

In contrast to QCD, the electroweak theory is chiral. This means that electroweak interactions differentiate between left- and right-handed fermions. The left- and right-handedness of fermions is defined by the projection operators

$$P_{\text{L}} = \frac{1}{2}\left(1 - \gamma^5\right) \quad \text{and} \quad P_{\text{R}} = \frac{1}{2}\left(1 + \gamma^5\right) \quad . \tag{2.15}$$

The gauge group of the electroweak model is $SU(2)_W \times U(1)_Y$, where the respective quantum numbers are given by the weak isospin W and the hypercharge Y. The Standard Model fermions are arranged in $SU(2)_W$ singlets and doublets. Experimentally it is known that neutral current interactions via a photon or massive Z boson couple both to right- and left-handed fermions

and conserve flavor. On the other hand charged currents via the massive vector bosons $W^{\pm}$ are flavor-changing but couple only to left-handed fermions. Phenomenological and symmetry based considerations (see [11]) suggest to arrange the fermions of the first generation in $SU(2)_W$ singlets and doublets as shown in equation (2.16).

$$q_L = \begin{pmatrix} u_L \\ d_L \end{pmatrix} \quad l_L = \begin{pmatrix} \nu_L \\ e_L \end{pmatrix} \quad u_R \quad d_R \quad e_R \tag{2.16}$$

Here right-handed neutrinos, which would be singlets both under $SU(2)_W$ and $U(1)_Y$, have been omitted. The other two generations are arranged in an analogous way. Now one introduces the massless gauge bosons W_μ^a and B. The covariant derivative D_μ is defined differently for left- and right-handed fermions (e.g. [17]):

$$\begin{aligned} D_\mu \psi_L &= \left(\partial_\mu + \frac{i}{2}\sigma^a W_\mu^a + \frac{i}{2} g' Y B_\mu\right) \psi_L \\ D_\mu \psi_R &= \left(\partial_\mu + \frac{i}{2} g' Y B_\mu\right) \psi_R \end{aligned} \tag{2.17}$$

The coupling constants of $SU(2)_W$ and $U(1)_Y$ are g and g'. Here σ^a labels the Pauli matrices. Because of the transformations (2.17) the electroweak model violates parity [17].

In order to achieve $SU(2)_W \times U(1)_Y$ gauge invariance for the chiral electroweak gauge symmetry one has to omit the fermion mass terms in equation (2.1) (e.g. [18]). This issue is addressed along with the masses of the vector bosons in section 2.5.

2.5 Higgs Mechanism

Electroweak physics as it has been introduced in section 2.4 describes fermions and vector bosons without masses. This is in blatant conflict with observations and one needs to wonder how the four electroweak gauge bosons relate to the observed massive vector bosons $W^{\pm}$and Z and the massless photon γ.

The missing mass terms in the Standard Model can be generated with the help of the "Higgs mechanism". One introduces an additional $SU(2)_W$ doublet in the form of a complex scalar field ϕ with four degrees of freedom. Then an additional term

$$\mathscr{L}_{\text{higgs}} = (D_\mu \phi)^\dagger (D^\mu \phi) + \mu^2 \phi^\dagger \phi - \frac{1}{2}\lambda \left(\phi^\dagger \phi\right)^2 \quad \text{with} \quad \mu^2 > 0 \tag{2.18}$$

is added to the Standard Model Lagrangian. The respective potential has a minimum at $\phi^\dagger \phi = v^2 = \mu^2/\lambda$. In order to use perturbative methods the field ϕ needs to be translated to a selected minimum. This procedure is called spontaneous symmetry breaking. After symmetry breaking three massless Goldstone bosons appear in the Lagrangian. One can eliminate these degrees of freedom by applying appropriate gauge transformations (see e.g. [18]). Accordingly it is possible to write ϕ in the unitary gauge given by

$$\phi = \begin{pmatrix} 0 \\ v + \frac{1}{\sqrt{2}} h(x) \end{pmatrix} \tag{2.19}$$

with a real scalar field $h(x)$. With this choice one can rewrite $\mathscr{L}_{\text{higgs}}$ and the real scalar field $h(x)$, the "Higgs boson", is obtained with its self interactions and a contribution [11]

$$\mathscr{L}'_{\text{higgs}} = \dots \begin{pmatrix} W_\mu^1, W_\mu^2, W_\mu^3, B_\mu \end{pmatrix} \begin{pmatrix} g^2 & 0 & 0 & 0 \\ 0 & g^2 & 0 & 0 \\ 0 & 0 & g^2 & gg' \\ 0 & 0 & gg' & g'^2 \end{pmatrix} \begin{pmatrix} W_\mu^1 \\ W_\mu^2 \\ W_\mu^3 \\ B_\mu \end{pmatrix} \quad . \tag{2.20}$$

One can see that this expression introduces mass terms for the vector bosons of $SU(2)_W \times U(1)_Y$. The matrix is of rank three with eigenvalues 0, g^2, g^2 and $g^2 + g'^2$. Thus the mass eigenstates are combinations of the fields W_μ, W_μ^2, W_μ^3 and B_μ which can be interpreted as the massless photon, the two bosons $W^\pm$ of equal mass and the heavier Z boson.

While direct fermion mass terms are not invariant under electroweak gauge transformations it is possible to introduce additional Yukawa couplings between the Higgs doublet and the Standard Model fermions. These contributions can generate the fermion mass terms in the process of symmetry breaking. The same Yukawa couplings also introduce the phenomenon of quark flavor mixing in charged current interactions which is described by the CKM-matrix (see e.g. [12]).

Concludingly, one can now write down the full Lagrangian $\mathscr{L}_{\mathrm{SM}}$ of the Standard Model before symmetry breaking, which is

$$\mathscr{L}_{\mathrm{SM}} = \mathscr{L}_{\mathrm{YM,\,SU(3)_C \times SU(2)_W \times U(1)_Y}} + \mathscr{L}_{\mathrm{higgs}} + \mathscr{L}_{\mathrm{Yukawa}} \quad . \tag{2.21}$$

3 Aspects of Physics beyond the Standard Model

While the Standard Model is able to make predictions that have been verified with an impressive level of accuracy, there are many open questions and theoretical difficulties like the ones mentioned in section 1. A multitude of theories have been proposed to address such problems. In the context of this work it is an important proof of concept to test the presented concept of a model unspecific search with respect to some popular theories of beyond Standard Model physics.

The models and ideas that are outlined in the following paragraphs correspond to the test scenarios used in chapter 10. Here the focus is on signatures with one or more high p_T [1] photons in the final state. The presented models have been selected because of the availability of simulated events and due to their complementary signatures.

Each of the discussed theories is sufficiently sophisticated for being the subject of a large amount of papers and textbooks. All of them are covered by dedicated analyses within the CMS collaboration. Here it is only possible to introduce some basic ideas and concepts. The important parameters for the benchmark scenarios in chapter 10 are introduced for each model.

3.1 Supersymmetric Models with Gauge Mediated Supersymmetry Breaking

From today's perspective the Poincaré algebra seems to reflect a fundamental symmetry of nature. One may ask if for a relativistic quantum field theory the Poincaré group can be extended in a non-trivial[2] way. The Coleman-Mandula no-go theorem states that such an extension is not possible with bosonic[3] generators [19]. However one can consider additional fermionic generators which change the spin of a state by 1/2, so that

$$Q_\alpha \left|\text{bosonic state}\right\rangle = \left|\text{fermionic state}\right\rangle_\alpha \tag{3.1}$$

and

$$Q_\alpha \left|\text{fermionic state}\right\rangle^\alpha = \left|\text{bosonic state}\right\rangle \quad . \tag{3.2}$$

This kind of additional symmetry is called "Supersymmetry" (SUSY). Though it is discussed to introduce $N > 1$ sets of operators Q_a we will only consider the case $N = 1$. The extended commutation relations of the Poincaré algebra are [19, (2.17) to (2.20)]

$$\begin{aligned} [Q_\alpha, P^\rho] &= 0 \\ \left\{Q_\alpha, \overline{Q}_{\dot\beta}\right\} &= 2\,(\sigma^\rho)_{\alpha\dot\beta}\, P_\rho \\ [M^{\rho\sigma}, Q_\alpha] &= -i\,(\sigma^{\rho\sigma})_\alpha^{\ \beta}\, Q_\beta \\ \{Q_\alpha, Q_\beta\} &= \left\{\overline{Q}_{\dot\alpha}, \overline{Q}_{\dot\beta}\right\} = 0 \quad . \end{aligned} \tag{3.3}$$

[1] transverse momentum (with respect to the beamline)

[2] with new generators that do not commute with the generators of the Poincaré algebra

[3] preserving the bosonic or fermionic nature of a state

In these equations the Weyl spinor formalism is used in which four component spinors Ψ , $\bar{\Psi}$ are replaced by

$$\Psi = \begin{pmatrix} \psi_\alpha \\ \bar{\chi}^{\dot{\alpha}} \end{pmatrix} \quad \text{and} \quad \bar{\Psi} = \begin{pmatrix} \chi^\alpha & \bar{\psi}_{\dot{\alpha}} \end{pmatrix} \quad . \tag{3.4}$$

Correspondingly undotted (dotted) indices label left-handed (right-handed) Weyl spinors. For a renormalizable supersymmetric extension of a four-dimensional field theory all particle fields can be arranged in chiral and vector supermultiplets. Each of them has an equal number of fermionic and bosonic degrees of freedom [20, pages 6-7]. While a chiral supermultiplet contains a single spin-1/2 Weyl fermion ψ and two real scalars ϕ, a vector supermultiplet contains a spin-1 vector boson A_μ and a spin-1/2 Weyl fermion λ. All Standard Model fermions must reside in chiral supermultiplets [20, page 7]. To include gravity one may add a supermultiplet featuring a spin-2 field with a spin-3/2 superpartner[4]. The price for this is the loss of renormalizability.

Now it is the aim to construct a Lagrangian that respects the U(1) × SU(2) × SU(3) gauge symmetries of the Standard Model but is also invariant under supersymmetry transformations. A general ansatz for a corresponding Lagrangian density is (see [20, page 25])

$$\begin{aligned} \mathscr{L} \;&=\; \mathscr{L}_{\text{chiral}}\left(\psi,\phi\right) + \mathscr{L}_{\text{vector}}\left(A^\mu,\lambda\right) \\ &\quad -\sqrt{2}g\left(\phi^* T^a \psi\right)\lambda^a - \sqrt{2}g\lambda^{\dagger a}\left(\psi^\dagger T^a \phi\right) + g\left(\phi^* T^a \phi\right) D^a \quad . \end{aligned} \tag{3.5}$$

T^a, $a = 1 \ldots n$ labels the n generators of the gauge group. D^a are gauge auxiliary fields. In this concise notation the indices labeling the multiplicity of the fields have been partially suppressed. Obviously the notation of $\mathscr{L}_{\text{chiral}}$ and $\mathscr{L}_{\text{vector}}$ are rather symbolic as those terms also contain the hermitian conjugate fields and partial derivatives. A derivation of equation (3.5) and the expansions of the terms $\mathscr{L}_{\text{chiral}}$ and $\mathscr{L}_{\text{vector}}$ can be found in [20]. The term $\mathscr{L}_{\text{chiral}}$ can be written with the help of the "superpotential" W, which is an analytic function of the scalar fields ϕ_i and plays an important part in the definition of supersymmetric models.

One of the main arguments supporting the credibility of supersymmetric theories is that they can present an elegant solution to the so-called "hierarchy problem". The hierarchy problem is related to the quadratic divergences of the one-loop corrections to the higgs mass. Those turn out to be higher than the mass term $-\mu^2$ of the higgs boson on tree level by many orders of magnitude, if we chose the cut-off parameter Λ at the scale of the Planck mass. Consequently getting back to a phenomenologically viable scale of the effective higgs mass parameter at one-loop level would require an excessive amount of finetuning for $-\mu^2$. A more detailed discussion of the hierachy problem and how it is addressed by SUSY theories can for example be found in [21].

MSSM

The Minimal Supersymmetric Standard Model (MSSM) contains the smallest number of supermultiplets that is consistent with the particle content of the Standard Model and respects the gauge symmetries of the Standard Model. To obtain masses for all quarks and leptons it is necessary to introduce two higgs chiral supermultiplets [22, page 22]. Once one has chosen the gauge group and the particle content of the supersymmetric theory the only remaining freedom remains in the choice of the superpotential W [19, page 23].

Some results in the context of supersymmetry can be expressed more elegantly in a formalism in which the particle fields are degrees of freedom of mathematical objects commonly known as "superfields", which closely correspond to the respective supermultiplets (see e.g. [19] for an introduction). The particle content and the necessary notation for the MSSM superfields is given in table (3.1). If one requires R-parity, which will be described in the next paragraph, the

[4] this is a necessary extension for models in which supersymmetry is elevated to a local symmetry

superfield	particle	spin	superpartner	spin
Q	quark (u_L, d_L)	$\frac{1}{2}$	squark $\left(\tilde{u}_L, \tilde{d}_L\right)$	0
$\bar{u}$	quark $u_R^\dagger$	$\frac{1}{2}$	squark $\tilde{u}_R^\dagger$	0
$\bar{d}$	quark $d_R^\dagger$	$\frac{1}{2}$	squark $\tilde{d}_R^\dagger$	0
L	lepton (ν, e_L)	$\frac{1}{2}$	slepton $(\tilde{\nu}, \tilde{e_L})$	0
$\bar{e}$	lepton $e_R^\dagger$	$\frac{1}{2}$	slepton $\tilde{e}_R^\dagger$	0
H_u	higgs (H_u^+, H_u^0)	0	higgsino $\left(\tilde{H}_u^+, \tilde{H}_u^0\right)$	$\frac{1}{2}$
H_d	higgs $\left(H_d^0, H_d^-\right)$	0	higgsino $\left(\tilde{H}_d^0, \tilde{H}_d^-\right)$	$\frac{1}{2}$
	gluon g	1	gluino $\tilde{g}$	$\frac{1}{2}$
	W-bosons $W^\pm$, W^0	1	winos $\widetilde{W}^\pm$, $\widetilde{W}^0$	$\frac{1}{2}$
	B-boson B	1	bino $\widetilde{B}$	$\frac{1}{2}$

Table 3.1 The particle content of the MSSM. Each chiral supermultiplets containing quarks and leptons represents three family specific supermultiplets

superpotential of the MSSM can be parametrized as given in [20, page 30]. R-parity is a discrete symmetry related to a multiplicative conserved quantum number P_R which can be defined for each particle via

$$P_R = (-1)^{3(B-L)+2s} \quad , \tag{3.6}$$

where L and B label the lepton and baryon number and s is the spin. Postulating R-parity has several important consequences:

- R-parity eliminates terms in the MSSM superpotential that violate the conservation of lepton or baryon number. Such terms are strongly restricted by experimental bounds (e.g. the lower limit on the lifetime of the proton of $> 10^{29}$ years[5] [23]).
- The lightest supersymmetric particle (LSP) is stable. In the case of a neutral LSP this results in a potential dark matter candidate.
- Sparticles[6] decay in chains that result in a state with an odd number of LSPs.
- Only an even number of sparticles can be created in the interaction of Standard Model particles (e.g. in parton interactions during proton-proton collision at the LHC [7]).

Supersymmetry Breaking

Supersymmetry implies that the masses of the Standard Model particles are equal to those of their respective superpartners. Hence supersymmetry needs to be broken if we want to explain why supersymmetric particles have not yet been observed. In the process of breaking supersymmetry one would certainly like to keep the important asset of solving the hierarchy problem. This restricts the possible form of the SUSY breaking terms that can be added to the MSSM Langrangian. Terms that maintain the cancellation of quadratic divergencies are called "soft" (e.g. [22]). A general parametrization of MSSM soft SUSY breaking terms respecting R-parity is for example given in [22, eq. (61)].

After introducing supersymmetry in this way one remains with a high number of parameters. In order to build a viable theory the parameters need to follow several rules for which there seems to be no convincing justification. As this is clearly not a satisfactory situation one is looking

[5] mode dependent limits may be even more restrictive
[6] supersymmetric partners of the Standard Model particles
[7] Large Hadron Collider (see chapter 4)

for underlying principles presenting a convincing mechanism of "spontaneous superymmetry breaking". Spontaneous supersymmetry breaking implies that though the Lagrangian respects supersymmetry SUSY is broken in the ground state. This means that $Q_\alpha^\dagger |0\rangle \neq 0$ and $Q_{\dot{\alpha}}^\dagger |0\rangle \neq 0$ [20, page 47]. Several competing mechanisms of spontaneous superymmetry are under discussion. One of them is "gauge mediated supersymmetry breaking" (GMSB).

It should be mentioned that, after generating the full mass spectrum in the context of electroweak symmetry breaking through the MSSM version of the Higgs mechanism, the mass eigenstates do not correspond to the gauge eigenstates. For example the two neutral higginos, the bino $\widetilde{B}$ and the neutral wino $\widetilde{W}^0$ mix to form the four neutralino mass eigenstates χ_1^0, χ_2^0, χ_3^0 and χ_4^0 (in the order of increasing mass) [22].

GMSB

In the following paragraphs we consider a scenario with an additional supermultiplet containing the graviton and the gravitino and local supersymmetry.

It seems to turn out that for an appealing explanation of spontaneous SUSY breaking one needs to add a "secluded sector" with a number of additional superfields. The fields in the secluded sector have no renormalizable tree level couplings with the "visible sector" [24]. The visible sector contains all standard particles and their supersymmetric partners. In models with gauge mediated SUSY breaking (GMSB) the breaking of supersymmetry in the secluded sector is communicated to the visible sector by a number of chiral superfields that form the "messenger sector". These messenger fields couple to the visible sector through the usual Standard Model $\mathrm{U}(1) \times \mathrm{SU}(2) \times \mathrm{SU}(3)$ gauge interactions [20, page 53]. As the messenger fields also couple to the original SUSY breaking in the secluded sector they can relay the breaking to the visible sector. In this kind of setup the mass splittings between the particles and superpartners in the visible sector are the result of higher order corrections involving the gauge interactions between the visible sector and the messenger fields [24, page 10].

In GMSB scenarios the gravitino is the lightest supersymmetric particle (LSP). For the minimal R-parity conserving GMSB model (see [25]) that is used for benchmark scenarios in section 10 there are six remaining free parameters [26]:

$\mathbf{M_m}$ characterizes the mass scale of the messenger fields introduced by the SUSY breaking mechanism.

$\mathbf{\Lambda}$ $= F_m/M_m$ is related to the masses of the MSSM superpartners. Together with M_m the parameter F_m determines the vacuum expectation value of the superfield X in the hidden sector.

$\mathbf{N_m}$ fixes the number of generations for the messenger fields. For $N = 1$ the lightest neutralino $\tilde{\chi}_1^0$ tends to be the NLSP [8].

$\tan\beta = \langle H_u^0 \rangle / \langle H_d^0 \rangle$ is the ratio of the vacuum expectation values of the higgs fields H_u^0 and H_d^0.

sign μ labels the sign of the mass parameter of the higgs fields in the MSSM superpotential.

$\mathbf{C_{grav}}$ $= F_0/F_m$ in which F_0 is related to the scale of supersmmetry breaking is related to the mass of the gravitino and the lifetime of the NLSP [26, pages 23-24].

$$m_{\tilde{G}} \approx C_{grav} \cdot \left(\frac{\sqrt{M_m \Lambda}}{100\,\mathrm{TeV}} \right)^2 \cdot 2.4 \quad \mathrm{eV} \tag{3.7}$$

$$c\,\tau_{\mathrm{NLSP}} \approx C_{grav}^2 \cdot \left(\frac{100\,\mathrm{GeV}}{m_{\mathrm{NLSP}}} \right) \cdot \left(\frac{M_m}{\Lambda} \right)^2 \cdot 10^{-5} \quad \mathrm{m} \tag{3.8}$$

[8] next to lightest supersymmetric particle

3.2 Excited Muons

The Standard Model of particle physics does not give an explanation of the number of fermion generations. It does also not predict mass parameters or mixing angles. One can try to address these issues by postulating that there is a substructure for the Standard Model fermions. This idea usually runs under the name of "compositeness" (e.g. [27]) [9]. The particles building the substructure of the Standard Model fermions are called "preons". It might be argued that such an approach is in some way analogous to explaining the complicated mass patterns of mesons and baryons by describing these particles as bound states of quarks.

Models with composite leptons imply the possibility of excited lepton states. The respective Standard Model lepton would then correspond to the ground state of the bound preon system. Here it is assumed that the excited lepton state f^* has spin and weak isopsin of 1/2.

Without any detailed idea of the the underlying preon dynamics one tries to estimate observable effects with the help of effective interactions. At energies below the compositeness scale Λ the assumed new strong interactions between preons can be described with an effective Lagrangian. It describes the exchange of heavy resonances [28] with a four-fermion contact interaction of the form [29]

$$\mathscr{L}_{\text{contact}} = \frac{g_*^2}{2\Lambda^2} j_\mu j^\mu \quad . \tag{3.9}$$

Obviously the scale of compositeness Λ is an important free parameter of the model. We choose $g_*^2 = 4\pi$. It can be argued that $g_*^2 \approx 4\pi$ is necessary for a self-consistent model [28]. Neglecting right handed currents and setting specific factors for each term to one we can write the current j_μ of the contact interaction as [29]

$$j_\mu = \bar{f}_L \gamma_\mu f_L + \bar{f}_L^* \gamma_\mu f_L^* + \bar{f}_L^* \gamma_\mu f_L + + \bar{f}_L \gamma_\mu f_L^* + \text{h.c.} \quad . \tag{3.10}$$

Here $f_L = (\nu_f \, f^-)_L^T$ and $f_L^* = \left(\nu_f^* \, f^{*-}\right)_L^T$ are lefthanded isodoublets for the fermion states.

One can also model gauge interactions resulting in vertices which couple the excited fermion to a gauge boson and a ground state fermion. The respective effective Lagrangian for the case of excited leptons is [30]

$$\mathscr{L}_{\text{trans}} = \frac{1}{2\Lambda} \bar{f}_R^* \sigma^{\mu\nu} \left[g f \frac{\vec{\tau}}{2} \vec{W}_{\mu\nu} + g' f' \frac{Y}{2} B_{\mu\nu} \right] f_L + \text{h.c.} \tag{3.11}$$

$W_{\mu\nu}$ and $B_{\mu\nu}$ are the $SU(2)$ and $U(1)$ field strength tensors. Y is the weak hypercharge and $\vec{\tau}$ summarizes the $SU(2)$ structure constants [27]. g and g' are the electroweak gauge couplings. It has been argued that one might expect the parameters f_s, f and f' to be of order 1 [29].

If the description of processes via contact interactions as outlined in the context of equation 3.9 is valid it is possible to neglect the influence of gauge interactions on the production of excited leptons [27]. The cross section for the production of single excited muons at a center of mass energy $\sqrt{\hat{s}}$ in a $q\bar{q} \to l\bar{l}^*$, $\bar{l}l^*$ process via contact interaction is given by [29]

$$\hat{\sigma}\left(q\bar{q} \to l\bar{l}^* \, , \, \bar{l}l^*\right) = \frac{\pi}{6 \cdot \hat{s}} \left(\frac{\hat{s}}{\Lambda^2}\right)^2 \left(1 + \frac{\nu}{3}\right) \left(1 - \frac{m_{l^*}^2}{\hat{s}}\right)^2 \left(1 + \frac{m_{l^*}^2}{\hat{s}}\right) \tag{3.12}$$

with $\nu = \left(\hat{s} - m_{l^*}^2\right) / \left(\hat{s} + m_{l^*}^2\right)$.

In chapter 10 benchmark scenarios with $m_{\mu^*} = 0.4\,\text{TeV}$ and $m_{\mu^*} = 1\,\text{TeV}$ are evaluated at a compositeness scale of $\Lambda = 6.0\,\text{TeV}$. These scenarios are beyond the reach of current exclusion limits [31, 32]. In the following paragraph the decay mechanism of the excited muon is discussed in the context of the considered benchmark scenarios.

[9] One might also discuss a substructure for the Standard Model bosons, but such models have less appeal (see [28]).

For $\Lambda = 6\,\text{TeV}$ and the selected masses m_{μ^*} the most important contribution to the total decay width of the excited muon is due to gauge interactions (see [27, table 1]). The partial decay widths for $f^* \to f\gamma$ and $f^* \to fV$ with $V = W, Z$ are [29]

$$\Gamma\left(f^* \to f\gamma\right) = \frac{\alpha}{4} f_\gamma^2 \frac{m_{f^*}^3}{\Lambda^2} \tag{3.13}$$

$$\Gamma\left(f^* \to fV\right) = \frac{g_V^2}{32\pi} f_V^2 \frac{m_{f^*}^3}{\Lambda^2} \left(1 - \frac{m_V^2}{m_{f^*}^2}\right)\left(2 + \frac{m_V^2}{m_{f^*}^2}\right) \tag{3.14}$$

wherein T_3 denotes the third component of the weak isospin of f^* and

$$f_\gamma = fT_3 + f'\frac{Y}{2} \quad , \quad f_W = \frac{f}{\sqrt{2}} \quad ,$$
$$f_Z = fT_3 \cos^2\theta_W - f'\frac{Y}{2}\sin^2\theta_W \quad ,$$
$$g_W = \frac{\sqrt{4\pi\alpha}}{\sin\theta_W} \quad , \quad g_Z = \frac{g_W}{\cos\theta_W} \quad .$$

3.3 Large Extra Dimensions

Standard Model physics is based on the Lorentz spacetime with coordinates x^μ and metric $g_{\mu\nu}$, including 3 spacial dimensions. Extending the four-dimensional spacetime with n_{ED} additional spacial dimensions has been noticed as an interesting concept since Kaluza and Klein proposed to unify gravity and electromagnetism in a theory with one compact[10] toroidal extra dimension y. In the approach of Kaluza and Klein one starts from a five-dimensional generalization of the Einstein-Hilbert action of four-dimensional gravity

$$S = \frac{M_*^3}{2}\int d^4x dy \sqrt{g_5} R_5 \tag{3.15}$$

with the determinant of the metric $g_{n_{\text{ED}}} = \det\left(g_{MN}\right)$, $M, N = 1 \ldots 4 + n_{\text{ED}}$ and the five-dimensional Ricci scalar R_5 (see [33, page 12]). Expanding g_{MN} in harmonics with respect to the y coordinate with radius L and using a convenient parametrization for the zero mode g_{MN}^0, one can derive the four dimensional action of the zero mode fields [34]

$$S_{\text{zero mode,4d}} = M_*^3 \pi L \int d^4x \sqrt{g_4}\left(R_4\left(g_4\right) - \frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \frac{1}{4}e^{-\sqrt{3}\phi}F_{\mu\nu}F^{\mu\nu}\right) \quad . \tag{3.16}$$

In this equation, which has been derived from the single compact Lagrangian (3.15), the last term can be identified with the action of electromagnetism and the first term with the action of general relativity. Thus, it can be motivated by equation (3.16) that theories with extra dimensions can lead to interesting results and might be worth further investigation. Matching the definition of the four-dimensional Einstein-Hilbert action with (3.16) one arrives at

$$M_*^3 = \frac{M_{Pl}^2}{2\pi L} \tag{3.17}$$

The Planck mass M_{Pl} is defined as $M_{\text{Pl}} = 1/\sqrt{G_{\text{N}}} \approx 1.2 \cdot 10^{19}\,\text{GeV}$, where G_{N} is the gravitational constant. However, the outlined original approach by Kaluza and Klein introduces several problems including the emergence of the gravitationally coupled scalar field ϕ (see [35]).

For gravity with n_{ED} compact toroidal extra dimensions one can start from the $4 + n_{\text{ED}}$ generalization of (3.15)

$$S_{\text{grav}} = \frac{M_*^{n_{\text{ED}}+2}}{2}\int dx^{4+n_{\text{ED}}}\sqrt{g_{4+n_{\text{ED}}}} R_{4+n_{\text{ED}}} \quad . \tag{3.18}$$

[10] curled up and thus of finite size

Then integrating out the extra dimensions the generalized $4+n_{\mathrm{ED}}$-dimensional version of (3.17) reads as (e.g. [35])

$$M_*^{n_{\mathrm{ED}}+2} = \frac{M_{Pl}^2}{(2\pi L)^{n_{\mathrm{ED}}}} \quad . \tag{3.19}$$

Equation (3.19) implies that in a theory with extra dimensions the fundamental Planck scale M_* can be much lower than M_{Pl}. In theories with extra dimensions a reduced Planck scale implies that one would expect a change in the inverse square law of gravity near to the corresponding length scale L [35].

Consider a reduced Planck scale of $M_* \sim 1\mathrm{TeV}$. Then $L \sim 10^{11}\,\mathrm{m}$ for $n_{\mathrm{ED}} = 1$, which is clearly excluded. But already for $n_{\mathrm{ED}} = 2$ one arrives at values of $L \sim 10^{-4}\,\mathrm{m}$. Hence scenarios with $n_{\mathrm{ED}} = 2$ are close to the exclusion limits of current experiments which test the inverse square law of the gravitational force [36]. The case $n_{\mathrm{ED}} \geq 3$ is not restricted by such measurements. Since gravity in $4 + n_{\mathrm{ED}}$ dimensions with $M_* \sim 1\mathrm{TeV}$ leads to scenarios with $L \gg 1/M_{\mathrm{Pl}}$ such extra dimensions are usually called "large".

If one sets $M_* \sim 1\mathrm{TeV}$ then in some way this can be seen as a solution to the hierarchy problem (see section 3.1). But it is argued that the problem is only shifted towards the question of why the extra dimensions should have the appropriate size [33].

Assuming extra dimensions one can consider the possibility that the four-dimensional world, as it is described for sufficiently low energies by the Standard Model, is realized as a surface (a "brane") in a higher dimensional spacetime (the "bulk") [34]. The theoretical challenge in constructing such "braneworld" models is to localize the Standard Model gauge fields on the brane and to explain the four-dimensional gravity.

The "ADD" model [37] combines the idea of introducing n_{ED} compactified toroidal dimensions of size L with the concept of physics on branes [34, page 10]. It should be mentioned that the resulting model is not renormalizable and can therefore only be used as an effective low energy theory (e.g. [33]). Making some simplifying assumptions [35] one can write down the ADD action as

$$S_{\mathrm{ADD}} = S_{\mathrm{grav}} + S_{\mathrm{brane}} \quad , \tag{3.20}$$

where the Standard Model is confined to a three-brane in the $4 + n_{ED}$-dimensional bulk [35]. For S_{brane} one can write [33]

$$S = \int d^4x \sqrt{g} \mathscr{L}_{\mathrm{SM}} \left(\hat{g}_{\mu\nu}, \phi, A, \psi, \ldots\right) \tag{3.21}$$

with the induced metric [38]

$$\hat{g}_{\mu\nu} = g_{\mathrm{MN}} \left(X\left(x\right)\right) \partial_\mu X^M \partial\nu X^N \quad . \tag{3.22}$$

Here X^M are the coordinates of the $4 + n_{\mathrm{ED}}$-dimensional spacetime and x^μ the coordinates in which the brane is parametrized.

Following the calculations in [33] one can expand $\mathscr{L}_{\mathrm{SM}}$ around the flat induced metric

$$S_{\mathrm{SM}} = S_0 + S_{\mathrm{int}} = \int d^4x \left\{\mathscr{L}_{\mathrm{SM}} \left(g_{\mu\nu}, \phi, A, \psi, \ldots\right)\right\} + \int d^4x \left\{ \left.\frac{\delta S_{\mathrm{SM}}}{\delta \hat{g}_{\mu\nu}}\right|_{\hat{g}_{\mu\nu}=g_{\mu\nu}} \frac{1}{M_*^{\frac{n}{2}+1}} h_{\mu\nu}\left(x\right) \right\} \tag{3.23}$$

Interpreting $h_{\mu\nu}$ as the graviton field of superimposed Kaluza Klein modes [33] and expressing S_{int} in terms of the energy-momentum tensor of the Standard Model

$$T^{\mu\nu} = \frac{1}{\sqrt{g}} \cdot \left.\frac{\delta S_{\mathrm{SM}}}{\delta \hat{g}_{\mu\nu}}\right|_{\hat{g}_{\mu\nu}=g_{\mu\nu}} \quad , \tag{3.24}$$

one finds the linear coupling of the graviton to $T^{\mu\nu}$. This implies that the graviton couples to the energy content of the Standard Model. From equation (3.23) one can derive the Feynman rules for diagrams involving the graviton [39].

Apart from searching for direct graviton production another conceivable signature at collider experiments is given by virtual graviton exchange [35]. As the graviton couples to quarks and photons one possibility is to look for a diphoton signal. Such a kind of signature is used as a test scenario in chapter 10. However, because of the small energy spacing of adjacent Kaluza Klein modes of about $1\,\mathrm{meV} - 100\,\mathrm{MeV}$ [40] one does not expect to observe single Kaluza Klein resonances, but rather a continuous excess of diphoton events at high energies.

An interesting aspect of virtual graviton exchange is that cross sections and energy spectrum are not only related to the reduced Planck scale $M_{*,\mathrm{ADD}} \sim 1\,\mathrm{TeV}$ but also to the cutoff scale M_f of the ADD model. While it seems natural to assume that $M_{*,\mathrm{ADD}} \approx M_f$ [40], in principle those two scales can differ, since we have no knowledge of the nature of quantum gravity [35]. Thus, depending on M_f, virtual graviton exchange may provide the clearest signal of ADD extra dimensions.

4 The LHC Collider

The main design goal of the LHC machine is to provide two colliding beams with 7 TeV protons at an instantaneous luminosity $\dot{L}$ of up to $10^{-34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$. For an interaction process with given cross section σ the luminosity is linked to the rate of interactions by equation (4.1)

$$\dot{N} = \dot{L} \cdot \sigma \quad . \tag{4.1}$$

The LHC is designed as a synchrotron collider. After previous acceleration steps bunches of protons are injected into the main ring at already ultrarelativistic energies of 450 GeV. There are two beamlines in which protons circle in opposite directions. Due to the restricted space in the LHC tunnel an innovative layout has been chosen in which the beam lines lie close to each other and can thus share important parts of the infrastructure. The distance between bunches corresponds to a separation of 25 ns. Consequently the inverse of this time interval defines the frequency of bunch crossings at the interaction points of 40 MHz. As the overall length of the LHC ring is close to 26.7 km there are up to about 2800 proton bunches per beam direction. Each bunch contains about $1.1 \cdot 10^{11}$ protons.

The Lorentz force implies that the energy of protons with a velocity close to c contained in a ring of given circumference can be estimated by equation (4.2, in SI units) [41].

$$E = \frac{Q \cdot c}{2\pi} \oint \vec{B} \times \mathrm{d}\vec{l} \tag{4.2}$$

There are also straight beam segments within the tunnel and apart from the bending magnets additional components need to be included to keep the beam focused. Consequently, the effective radius of the ring is reduced. The global geometric setup dictates that magnetic fields of about $B = 8.4\,\mathrm{T}$ are needed to reach the design goal of $E_\mathrm{p} = 7\,\mathrm{TeV}$. The beam energy corresponds to a center of mass energy for the proton proton collisions of $\sqrt{s} = 14\,\mathrm{TeV}$.

Only by the use of superconducting magnets the necessary field strength can be realized while keeping the energy consumption under control. The LHC dipole magnets (see figure 4) which are used for the bending of the beam use NbTi as superconductive material. They are cooled down to an operating temperature of 1.9 K. This extremely low temperature is necessary to ensure superconductivity at the needed current densities but leads to a reduced heat capacity making the superconducting cable potentially vulnerable to quenching effects[1]. The cooling liquid is super fluid helium.

Without additional magnet components proton-proton interactions in each bunch would rapidly defocus the beam. To counteract this and other effects related to the complex field of LHC beam optics special quadrupole magnets have been designed. With respect to a plane perpendicular to the beam line such components can only focus in one direction while defocussing in the orthogonal direction. Thus magnets with a rotated arrangement of the poles need to be positioned at a sufficiently close distance to result in an overall focusing of the proton beam [41]. There are several additional magnet types including components which help to correct for higher multipole moments. A summarizing list of the superconducting magnet types and their respective numbers in the LHC setup can be found in [42].

[1] local loss of superconductivity for example caused by a material region with increased temperature due to non-compensated energy absorption. The increased resistance can then cause a global loss of superconductivity in the material.

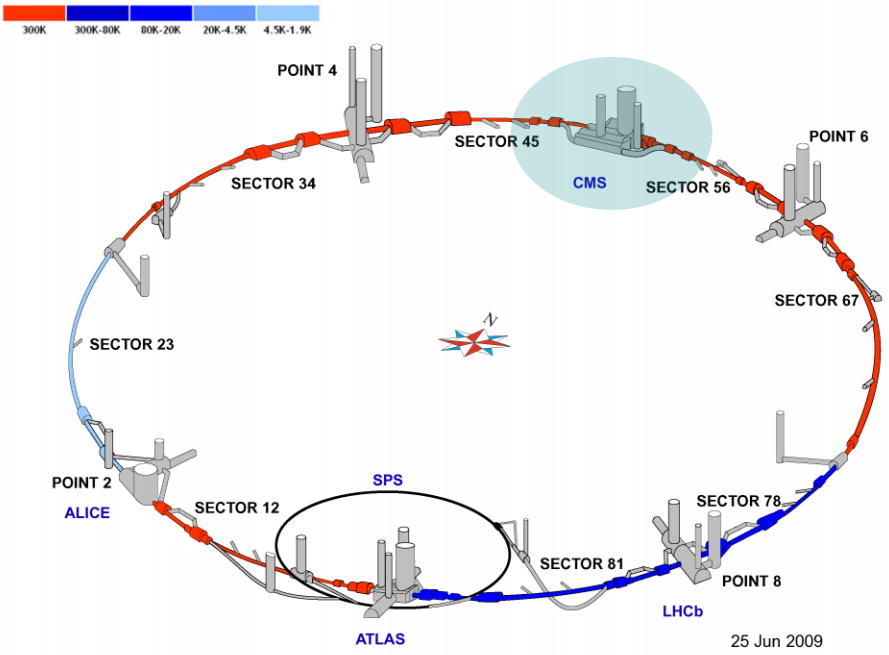

Figure 4.1 Snapshot showing the LHC detector sites and the current status of LHC cooling on June 25 2009 [43]. The shading of the CMS site has been added.

It will take about one hour to ramp the injected proton bunches up to the final energy level. Then the LHC can be kept in operation for roughly 10 hours. By then the beam quality has degraded to a level requiring a new injection cycle. Special magnets are used to focus the opposing beams at the interaction point. This causes a heightened luminosity because the effective area A, which characterizes the beam profile and is inversely proportional to the luminosity, is decreased. The dependency of $\dot{L}$ on the beam parameters is presented in equation (4.3) being stated similarly in [42].

$$\dot{L} \propto \frac{N_p^2 n_b \gamma f_r}{A} \tag{4.3}$$

Here N_p is the number of protons per bunch, n_b the number of bunches and f_r the revolution frequency. $\gamma = E/m$ is the relativistic factor and A characterizes the beam profiles and their intersection at the collision point. Incoming bunches from opposite directions interact even before reaching the nominal interaction region. This leads to destabilizing effects on the proton beams. To reduce the resulting problems the beams are brought to collision at a slight angle. This implies a reduction factor F for the luminosity which can be absorbed into the effective area A. By integrating $\dot{L}$ over a time interval one receives the integrated luminosity L in units of $1/\text{distance}^2$.

During a period after start up in 2009/2010 it is planned to run with a reduced luminosity and a center of mass energy of 10 TeV. All related results which are presented in this work are based on this scenario.

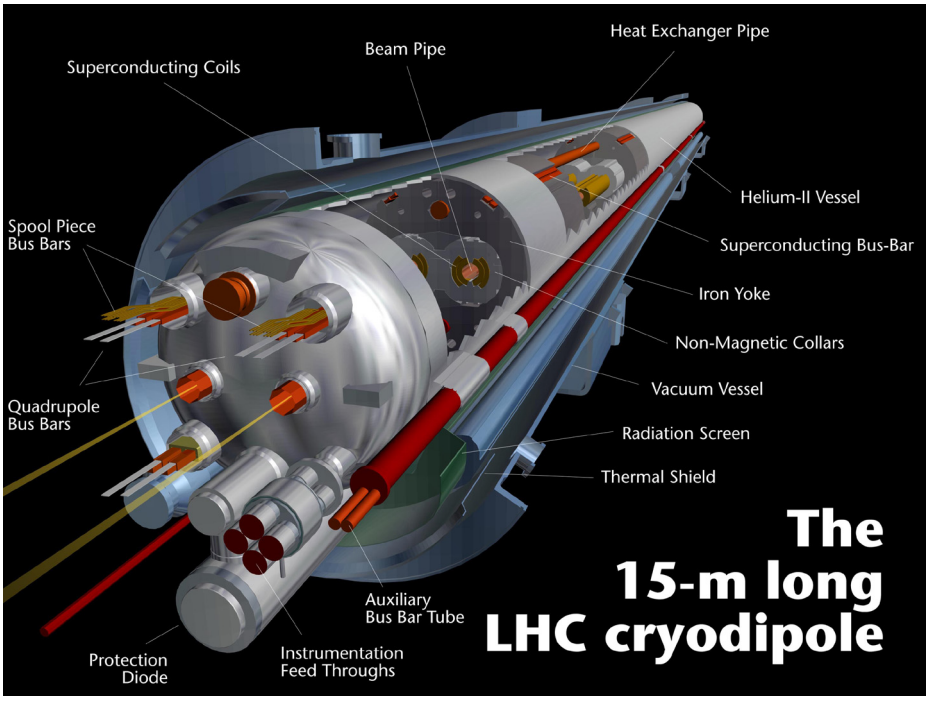

Figure 4.2 Schematic of an LHC superconducting dipole magnet [43]

5 The CMS Experiment

The CMS experiment has been designed as a multi purpose detector. The experimental setup is centered around one of the interaction points of the LHC proton-proton collider (section 4). The compact design of the CMS detector, in which a large fraction of the active detector components is contained within the 4 Tesla field of the CMS solenoidal magnet, provides high hermeticity close to the ideal angular coverage of 4π. Tracking detectors and calorimetry enable CMS to cover an extensive physics program ranging from precision measurements within the Standard Model to searches for physics beyond the Standard Model at the TeV scale. The overall shape of the detector is roughly cylindrical with a length of about 22 m and a diameter of about 14 m, weighing about 12500 t. In the order of increasing radial distance from the interaction point the main detector components comprise the following subdetectors (see also figure 5.2):

1. The **inner tracking system** is used for the identification and track measurement of charged particles.

2. The **electromagnetic calorimeter** provides energy measurements. Electrons and photons deposit almost all of their energy in this subdetector.

3. Typically the interactions of hadronic jets imply that most of their energy is deposited after transversing more material than given by the electromagnetic calorimeter. The main energy component of these objects is measured with the **hadronic calorimeter**.

4. The **muon system** identifies high energetic muons which can not be confined within the CMS experiment and measures their tracks and respective momenta.

Section (5.1) shortly addresses the influence of the detector on measured event rates and introduces the main coordinates which are used in the CMS experiment. Those coordinates are also referred to in the context of object reconstruction and data analysis. The following sections discuss the mentioned subdetectors and the CMS magnet. Additionally some basic performance parameters are covered. The underlying detection principles are mentioned but not explained in greater detail. The interested reader may be referred to [44] and [45] for detailed discussions concerning the physics of particle detection. The chapter concludes with a short section about the Level 1 trigger stage.

5.1 Measured Event Rates and CMS Coordinate System

Though the initial rate for an interaction process of interest is given by equation (4.1), the experimental setup introduces constraints on acceptance and reconstruction efficiencies of the produced particles. Here a "process" needs not to be directly linked to events from a closely restricted set of Feynman diagrams but can also refer to the collective set of events featuring objects fulfilling a certain set of criteria as is for example the case in the presented approach of a model unspecific search. There are also contributions from other processes falsely identified as being associated to the process of interest. In this case a more realistic description of the resulting event rates may be given by equation (5.1) in which i refers to the different contributing processes and j labels additional efficiency factors for each process.

$$\dot{N} \propto \sum_i L \cdot \sigma_i \prod_j \epsilon_{ij} \tag{5.1}$$

The coordinates which are mainly used for the global description of the detector components are z , r , ϕ and η (see figure 5.1).

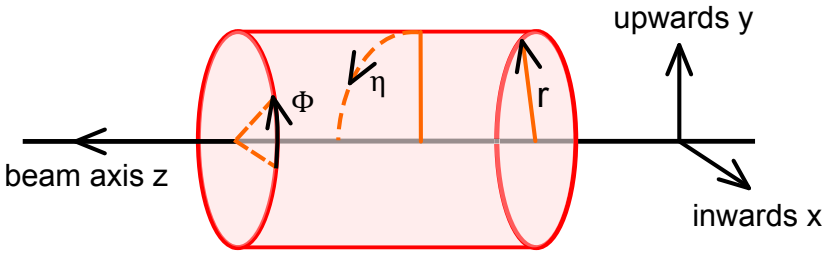

Figure 5.1 The CMS coordinate system

The z coordinate describes the distance from the nominal interaction point along the beam axis. The nominal interaction point is defined to be the origin of the CMS coordinate system. The positive range of the z axis points into the counter-clockwise direction if seen from the surface level. The radius r defines the distance between a coordinate point and the z-axis. ϕ and η fix the polar angles with respect to the nominal interaction point. Starting at zero for the direction pointing radially inwards with respect to the LHC collider ring the azimuthal coordinate ϕ defines the angular direction orthogonal to the beam pipe. The remaining polar angle θ is zero when pointing towards the surface with positive values towards +z and negative values towards -z. In most cases the pseudo rapidity η is stated instead of θ. The coordinate transformation between θ and η is given by equation (5.2).

$$\eta = -\ln\left[\tan\left(\frac{\theta}{2}\right)\right] \tag{5.2}$$

In the ultra relativistic limit (where particle masses can be neglected) the pseudo rapidity η can be used to define a quantity $\Delta\eta = \eta_2 - \eta_1$ for two objects with coordinates η_1 and η_2, which is invariant with respect to Lorentz boosts into the z direction. For ultra relativistic particles the pseudo rapidity is equal to the rapidity ψ. As defined with respect to the z direction ψ is given by equation (5.3) where v_z is the speed of the particle into the z direction and c the speed of light (see e.g. [46, page 12]).

$$\tanh\psi = \frac{v_z}{c} \tag{5.3}$$

The rapidity as defined in (5.3) is additive with respect to successive Lorentz boosts into the z direction. Hence, differences in rapidity ψ are invariant under such transformations. This property can be of interest for the analysis of events, as for parton-parton interactions within the proton the centre of mass frame is related to the the zero momentum frame of the detector by a boost into the z direction[1]. The graphical depiction of the coordinate directions is given in figure 5.1.

A frequently encountered measure of distance is given by

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} \quad . \tag{5.4}$$

It is often used in the the context of defining suitable isolation criteria for reconstructed objects.

[1] the small momentum imbalance in the plane orthogonal to z can be neglected in most situations

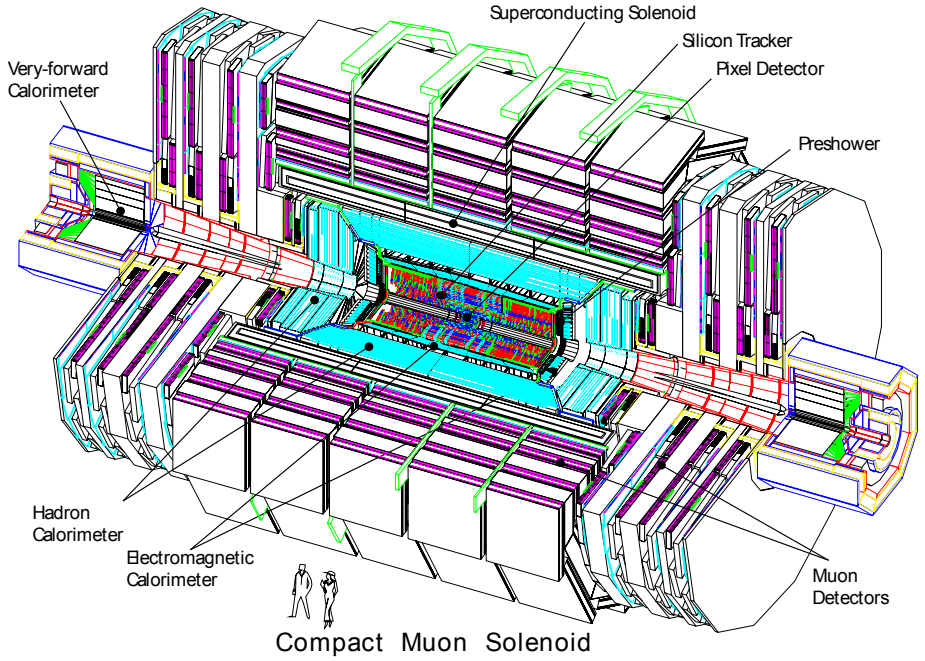

Figure 5.2 The CMS Detector [47]

5.2 Electromagnetic Calorimeter

The electromagnetic calorimeter (ECAL) is partitioned into a barrel section in the region $|\eta| < 1.48$ and two endcaps covering a range up to 3.0 in $|\eta|$. The main task of this detector component is to allow for a high quality measurement of the energy and direction of incoming photons and electrons. Figure 5.3 shows the main components including the "preshower detector" which is installed in front of the endcap regions.

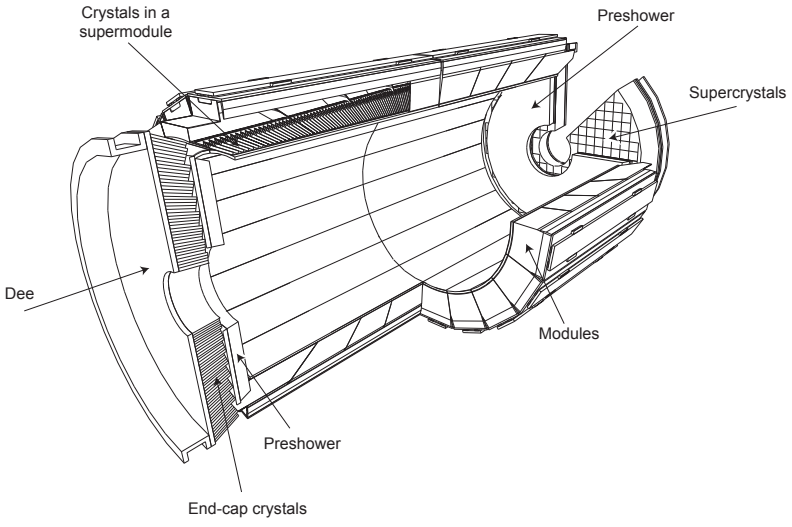

Figure 5.3 Main components of the electromagnetic calorimeter [47]

Both in barrel and endcaps scintillating $PbWO_4$ crystals are used as detection material. This type of material was chosen because of its short radiation length $X_0 = 0.89\,\text{cm}$ (For electrons X_0 is the length in which the energy of an ensemble of high energetic electrons is reduced to $1/e$ of the original value. For photons X_0 is the length in which the intensity of an ensemble of high energetic photons is reduced to $1/e^{7/9}$ of the original value. [48]) and small Molière radius

$R_M = 2.19\,\text{cm}$ (a measure of the width of an electromagnetic shower in the material). The crystals are sufficiently radiation-hard to ensure operation on a long timescale.

The barrel component consists of 61200 crystals granting a fine granularity that allows for a good separation between showers with closely aligned axes. With respect to the beam line the crystal front faces are positioned at a radius of 1.29 m. Each crystal has a depth of 230 mm corresponding to $25.8\,X_0$. Overall there are 17 geometrically distinct types of crystals covering angular intervals in $\eta \times \phi$ of 0.0174×0.0174 [49]. In order to prevent photons and electrons from having an angle of incidence corresponding to the line of intersection between adjacent crystals the components are aligned in a quasi-projective geometry tilting the crystals with respect to the nominal interaction point at an angle of about $3\,^\circ$ both in η and ϕ [47]. This reduces the probability of electromagnetic showers with substantial energy deposits that are not measured due to cracks between crystals. Groups of 5×5 crystals are combined into a supercrystal corresponding to a trigger tower for the L1 trigger stage. All barrel crystals are comprised into 36 supermodules. The supermodules are arranged into two cylindrical structures respectively covering the barrel range for $\eta > 0$ and $\eta < 0$. Thus each of the 36 supermodules covers an interval $\Delta\phi = 20\,^\circ$. Extending in η each supermodule has a substructure of four modules. This implies an overall number of eight modules in η for any fixed value of ϕ [50, page 26]. To detect the photons from the scintillating material two semiconductor based avalanche photodiodes (APDs) are used per barrel crystal.

The endcaps feature a number of 2×7324 crystals which are arranged in two D shaped structures (Dees) for each endcap. The detection material is the same as in the barrel section but there is just one position independent crystal shape with a depth of 220 mm corresponding to $24.7\,X_0$. To increase hermeticity the mounting features angles between $2\,^\circ$ and $8\,^\circ$ with respect to the nominal interaction point. The angular coverage of a crystal varies between 0.0175×0.0175 and 0.05×0.05 [50]. For the purpose of photodetection vacuum phototriodes (VPTs) are employed which are more radiation hard than the avalanche photodiodes and thus better suited for the higher doses in the endcap regions.

As the signal response of those components is very temperature sensitive it needs to be noted that the temperature of the crystals and photodetectors of the ECAL is kept constant within $\pm 0.05\,^\circ$ C [47, page 93]. The nominal value for operation of the ECAL is $18\,^\circ$C.

In the endcap region the "preshower detector" is installed as an additional detector component. It is placed in front of the ECAL crystals. The preshower detector consists of two lead layers with a material thickness of $2\,X_0$ for the first layer and $1\,X_0$ for the second. Both of them are covered by silicon strip detectors with an active area of 61 mm $\times$ 61 mm subdivided into 32 strips [47]. The orientation of strips is orthogonal for the two layers. Thus a two dimensional resolution can be achieved. The main goal of the preshower detector is an improved discrimination between π^0 mesons decaying into two photons and direct photon production by measuring differences in the respective typical shower shapes [51]. The preshower detector also allows for an improved position measurement of photons and electrons.

The expected energy resolution as a function of the incoming photon or electron can be parametrized as shown in equation (5.5) [47, page 116]. The equation can only be applied to particle energies lower than 500 GeV, since the influence of rear leakage included in parameter C is more prominent for higher energies.

$$\frac{\sigma_E}{E} = \sqrt{\left(\frac{S}{\sqrt{E/\text{GeV}}}\right)^2 + \left(\frac{N}{E/\text{GeV}}\right)^2 + C^2} \quad ; \quad S, N, C \text{ const.} \tag{5.5}$$

The "stochastic term" S summarizes contributions from lateral shower containment, photo-statistics and uncertainties on the energy absorption in the preshower detector for the endcap regions. The "noise term" N covers those uncertainties whose absolute values do not vary with

the energy comprising influences from electronics, digitization and pileup[2]. The "constant term" C includes the non uniformity of light collection within the crystal and rear leakage for energies below 500 GeV. Some of the mentioned contributions like the lateral shower containment will obviously show a dependency on the shower position in η and ϕ. A representative set of values as presented in [47, page 120] may be given by $S = 0.028$, $N = 0.12$ and $C = 0.003$. The resulting energy dependent resolution as calculated in [49] is displayed in figure 5.4.

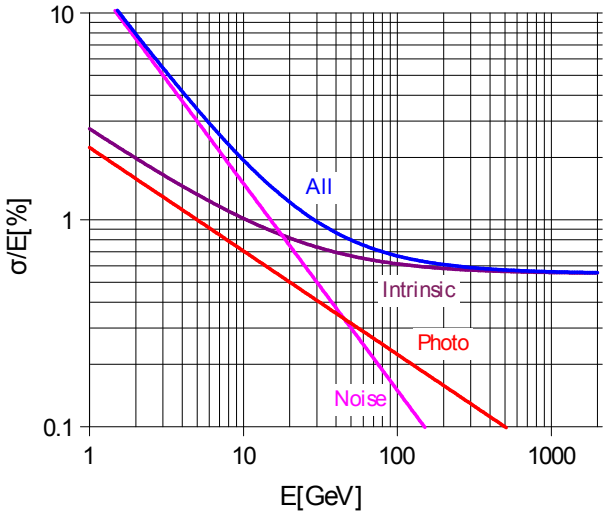

Figure 5.4 Expected energy resolution in the electromagnetic calorimeter [49]. The term labeled "intrinsic" includes shower containment and constant contributions. [50]

5.3 Hadronic Calorimeter

As for a given material the typical interaction length λ_I (a material dependent measure[3] of the distance in which a fraction $1/e$ of an ensemble of incoming hadronic particles features no inelastic interaction (e.g. [16, page 44]) of hadronic components is much higher than the radiation length X_0 an additional calorimeter is installed behind the ECAL detector and in the forward regions of the detector in order to measure the energy and position of hadronic jets. This "hadronic calorimeter" (HCAL) comprises several subcomponents being described in the following paragraphs. For the jet energies encountered at LHC energies a full shower containment within active material can not be achieved. Hence, the HCAL is built as a sampling calorimeter with alternating layers of scintillators and additional absorber material. The HCAL components within the CMS magnet extend from $R = 1.77\,\mathrm{m}$ to $R = 2.95\,\mathrm{m}$. The HCAL outer detector is the only HCAL component mounted outside of the CMS solenoid (see figure 5.5).

HCAL Barrel

The barrel section covering a range up to $|\eta| = 1.3$ is separated into two half barrels respectively covering the positive or the negative range in η. Most absorber layers are made of cartridge brass [47]. The scintillation material is given by plastic scintillator tiles. They are read out with the help of wavelength shifting fibers. The 16 absorption layers (including front- and backplate) are arranged in wedge-shaped components. Each wedge covers an angle of $20\,^\circ$ in ϕ . A staggered arrangement between layers ensures that there is no direction with only

[2] here: energy deposits from other overlapping proton-proton interactions

[3] usually stated in terms of g/cm^2 which of course can be transformed into a corresponding length by dividing by the respective material density

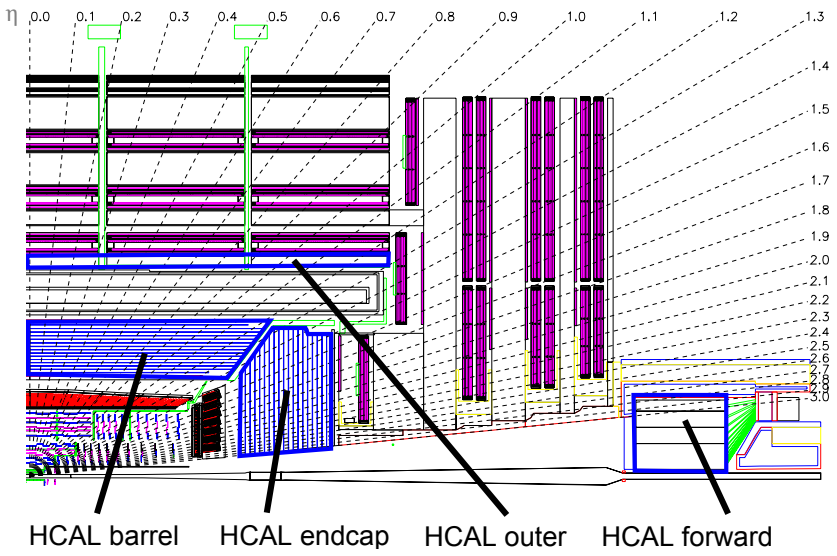

Figure 5.5 A cross section sketching one quadrant of the CMS detector in the η-z plane [47, page 123 (changed labels and added contours for the HCAL components)]. The η coordinate is indicated by the dashed lines.

dead material. Trays equipped with scintillator tiles are inserted between the absorption layers, starting from a scintillator layer directly behind the ECAL [52]. The optical readout is organized by integrating the signal from radially extending "towers".

Overall there are 36 (ϕ) times 16 (η) towers corresponding to a granularity of $(\Delta\eta, \Delta\phi) = (0.087, 0.087)$ [47, page 123]. Depending on η the material thickness of a tower can vary between five and eleven interaction lengths.

HCAL Outer

In order to improve the measurement of lately developing and high energetic jets in the central region, which would otherwise be the region of the calorimeter with lowest λ_I, there is an additional HCAL component in the region $|\eta| < 1.3$. It consists of five rings of scintillation detectors at a radius of about 4 meters using the CMS magnet as additional absorber material. For the central ring covering the range around $\eta = 0$ there are two scintillation layers mounted on front and back of a 20 cm thick plate of iron. The other rings feature just one layer of detection material. Except for the transition region between barrel and endcaps the minimum interaction length is thus increased to $\lambda_I = 11.8$ [47, page 138]. In order to build combined towers the scintillator tiles are arranged in a way roughly corresponding to the HCAL barrel design. Again the photodetection is improved by the use of wavelength shifting fibers. For the readout several fibers are combined via optical connectors.

HCAL Endcap

The general setup of the two HCAL endcap calorimeters is similar to the design of the barrel section. The hadronic endcap calorimeter covers the range of η between 1.3 and 3.0. Again a staggered geometry is used to arrange the subsequent layers of absorption material (cartridge brass) and scintillating plastic tiles within wedge-shaped structures.

There is a collective readout for those tiles that are pooled into an HCAL tower. The 14 (η) times 16 (ϕ) endcap towers correspond to a granularity of $(\Delta\eta, \Delta\phi) = (0.087, 0.087)$ for $|\eta| < 1.6$ and about $(\Delta\eta, \Delta\phi) = (0.17, 0.17)$ for $|\eta| > 1.6$ [47, page 133].

The most outwards towers of HCAL endcap and barrel regions are overlapping in order to improve hermeticity and to increase the effective radiation length in these transition regions.

HCAL Forward

The front faces of the HCAL forward detectors are situated on each side of the detector at a distance of 11.2 m from the nominal interaction point. They surround the beam pipe starting from a radius of 12.5 cm, featuring an overall cylindrical shape which is embedded in a multilayered radiation shielding. The range in η extends from about 3.0 to 5.2. As in this region the expected particle fluxes ($> 10^{10}$ Hz, [53]) and radiation doses (≈ 10 MGy, [47]) are much higher than the respective values that need to be expected for the other HCAL components, a more radiation hard technology had to be employed. The general layout is that of a sampling calorimeter. But instead of scintillating tiles polymer clad quartz fibers are used as active material. Because of the Cherenkov effect[4] spreading particle showers create light in the detection fibers. The fibers run parallel to the beam pipe. Half of the fibers have a 22 cm offset with respect to the plane where the other fibers start. This particular design allows for separation power between hadronically induced showers, and showers originating from photons and electrons, which develop earlier. The readout fibers are bundled into towers of about $(\Delta\eta, \Delta\phi) = (0.175, 0.175)$ [page 146][47].

Apart from measuring energy and position of jets the HCAL is also indispensable for the determination of missing transversal energy (MET) due to particles non- or only weakly interacting in the CMS detector. The specified requirement for the HCAL energy resolution can be roughly summarized by equation (5.6).

$$\frac{\sigma_E}{E} = \sqrt{\left(\frac{1.2}{\sqrt{E/\mathrm{GeV}}}\right)^2 + 0.05^2} \tag{5.6}$$

This design goal can indeed be achieved within a high magnetic field, as has been demonstrated with the help of test beam setups [54].

5.4 Superconducting Magnet

High magnetic fields are needed within the CMS experimental setup to allow for a momentum measurement of charged particles in the tracker and muon system (see sections 5.5 and 5.6). The CMS superconducting magnet is designed as a solenoid with four layers of superconducting (NbTi) cable with 542 turns per layer. [55]. The cable is embedded into a reinforcing structure of pure aluminium and aluminium alloy. This makes the conductor coil self sustained against the pressure and deformation exerted by the magnetomotive force during operation. In order to reach the critical temperature for superconductivity the “cold mass” is cooled with liquid helium. The length of the coil is 12.5 m with an inner radius of $r_i \approx 3.2$m and an outer radius of $r_o \approx 3.5$m. This makes it possible to install not only the tracker but also the main parts of the CMS calorimeters within the coil. Otherwise the performance of the calorimeters would suffer from the material budget of the magnet.

Rod ties support the magnet coil from different directions. The coil is positioned within a vacuum tank to achieve the necessary thermal insulation. The CMS magnet can induce a homogeneous magnetic field of up to 4 T in the region of the interaction point. The operation value is set to 3.8 T which is applied to all event simulations used in this work.

Outside the coil the magnetic flux is returned through a 10000 t iron yoke. Hosting large fractions of the CMS muon system the yoke features a complex geometry. As can be seen in figure 5.2 the yoke is segmented into five barrel wheels and two endcap discs.

[4] Cherenkov light is emitted when a charged particle crosses a material in which the effective velocity of light is lower than the velocity of the transversing particle (see e.g. [44, chapter 3]).

5.5 Inner Tracking System

The task of the CMS tracker is to reconstruct the trajectories of passing charged particles with transverse momentum $p_T > 1\,\text{GeV}$ in a range of approximately up to $|\eta| = 2.5$. The CMS tracker allows for a precise measurement of the transverse momentum as the trajectories of the charged particles are bent in the $\approx 3.8\,\text{T}$ magnetic field of the CMS magnet. All tracking components rely on silicon based semiconductor technology [56]. They provide fast readout and a radiation hardness up to the challenge of an operation period of up to ten years with an acceptable loss in efficiency. To reduce the radiation damage and to prevent problems due to "thermal runaway" (see [57, page 40] for a description of this effect) the tracker is cooled down to a temperature of about $-10\,^\circ\text{C}$. During the lifetime of the detector this temperature needs to be lowered down to $-27\,^\circ\text{C}$. The cooling system uses liquid C_4F_{16}. An exact alignment of the subcomponents of the tracking system is indispensable for an exact reconstruction of particle trajectories resulting in a high resolution for the track parameters. A precise mounting of detector components and a laser alignment system are important inputs for the alignment, which can then be complemented by an analysis of reference data [58].

Due to converting photons, multiple scattering and other material interactions, including the important effect of bremsstrahlung for electrons, the particle identification quality of the track reconstruction is reduced with an increasing amount of material that is encountered by a particle transversing the CMS tracker. Additionally, these effects hamper a clean energy measurement

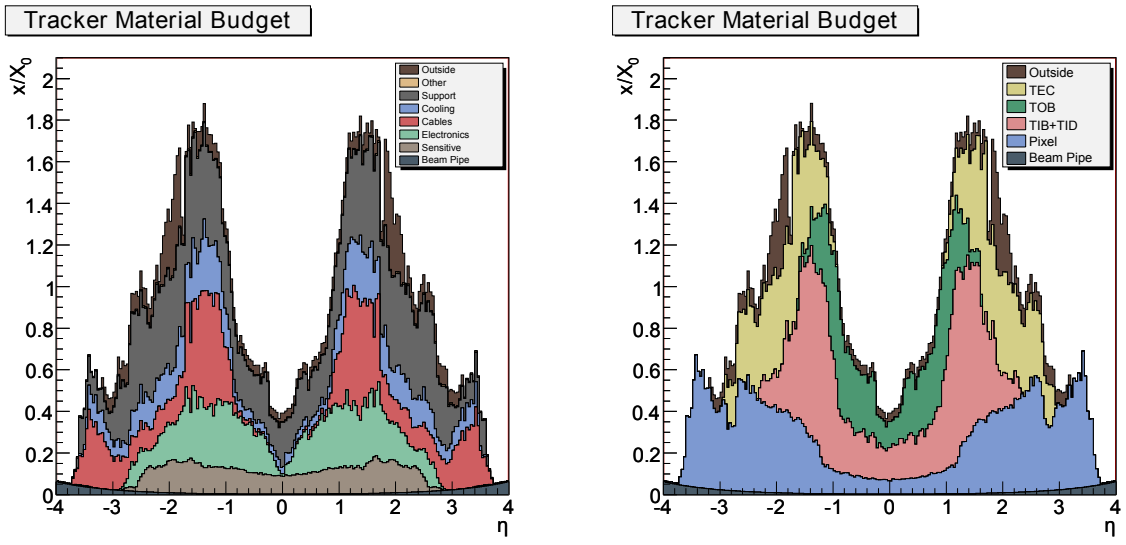

Figure 5.6 Expected material budget of the CMS tracker in units of X_0 broken down to different contributions as a function of η [47]

in the ECAL system for both photons and electrons. Thus the requirement of a highly granular multi layered tracking system with active cooling needs to be brought in agreement with the requirement of an acceptable material budget of the track detector. In figure 5.6 the material budget is presented in terms of the radiation length[5] X_0 as a function of η. The transverse momentum resolution of the tracker can roughly be parameterized as given by equation (5.7) where the linear term can be explained by the decreasing bending within the magnetic field for higher p_T. The constant term is caused by multiple scattering (e.g. [59]). In addition to these effects the resolution decreases further if the object looses a significant amount of energy via

[5] see section 5.2 for a definition of X_0

bremsstrahlung. For values of $p_T < 10\,\text{GeV}$ the resolution is expected to be better than 1 % while the decrease in relative resolution with p_T is approximately $0.2\,\text{TeV}^{-1} \cdot p_T$ [60].

$$\frac{\sigma\left(p_T\right)}{p_T} = \sqrt{\left(\left(a \cdot p_T\right)^2 + b^2\right)} \tag{5.7}$$

The tracking system is subdivided into a silicon strip detector and an embedded silicon pixel detector. The pixel detector is the detector component that is closest to the nominal collision point and encircles the beam pipe in the region of the expected primary vertex. A cross-section of the tracker layout is presented in figure 5.7.

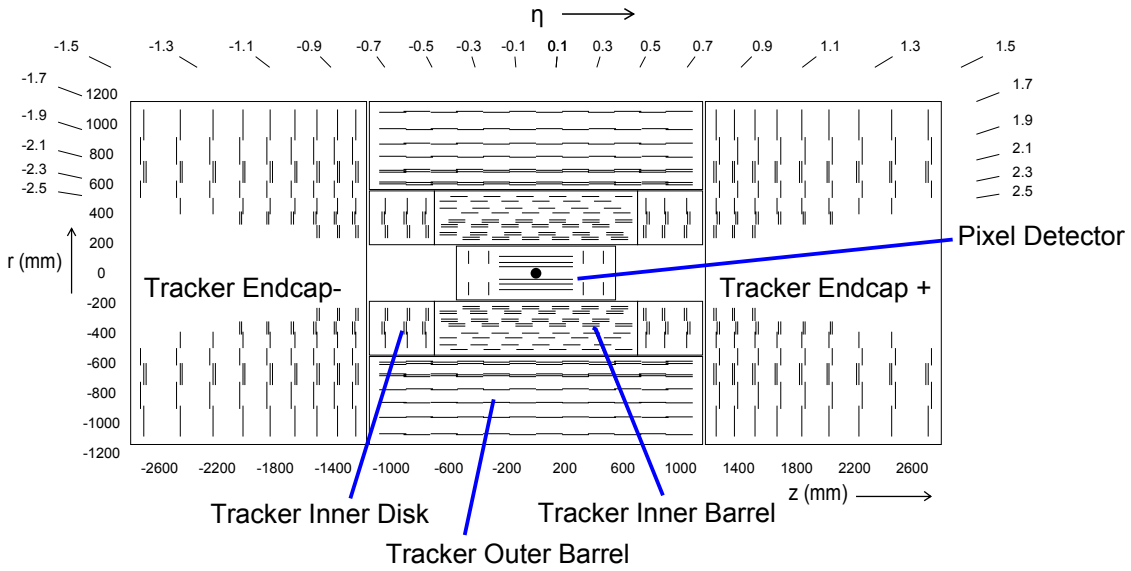

Figure 5.7 Schematic cross-section through the CMS tracker. Inner lines indicate tracker modules. Double lines present two sided modules with stereo layers. In the sections with strip detectors the lines also correctly indicate the main strip direction [47, page 31 (changed labels of components and resized labeling of coordinates)].

5.5.1 Pixel Detector

The CMS pixel detector has a cylindrical shape with three pixel layers at radii r of 4.4 cm , 7.3 cm and 10.2 cm and two disk segments at $z = \pm 34.5\,\text{cm}$ and $z = \pm 46.5\,\text{cm}$. In this way a range in η of up to 2.5 is covered by the pixel detector though the pixel coverage is reduced for higher values of η as can be seen in figure 5.8. Each pixel covers an area of $100 \times 150\,\mu\text{m}^2$ ensuring a good resolution both in the r-ϕ plane and the z direction. Altogether there are about 66 million pixels and corresponding readout channels. Hence the active area is roughly $1\,\text{m}^2$. The Lorentz drift spreads the created charges perpendicular to the magnetic field of the CMS solenoid. This results in a distribution of the overall charge over a range of several pixel elements. Accordingly, the detector resolution can be improved by interpolating with respect to the "charge sharing" between pixels. As a side effect this method has a positive influence on the noise rejection [47]. In order to utilize charge sharing also in the disk segments of the pixel detector the sensor elements are tilted by 20° with respect to the nominal axis of incidence, resulting in a turbine-like arrangement [61]. The achieved spatial resolution in ϕ and z for the barrel and respectively ϕ and r for the disks is about 15 to $20\,\mu\text{m}$. In the barrel section the pixels are arranged on modules with 8 or 16 read out chips per module. Each of these read out chips corresponds to an array of 52×80 pixels. In the forward pixel detector read out chips are arranged on "plaquettes" of differing size. The plaquettes are combined to twelve panels for

each of the two pixel layers [47]. The fine granularity of the CMS silicon based pixel detector is important for a successful track seeding and a reliable reconstruction of secondary vertices. Furthermore the granularity is crucial to keep the expected pixel occupancy per bunch crossing at the LHC design luminosity of $10^{-34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ below 1 %. Figure 5.8 presents a measure of the pixel coverage as a function of η.

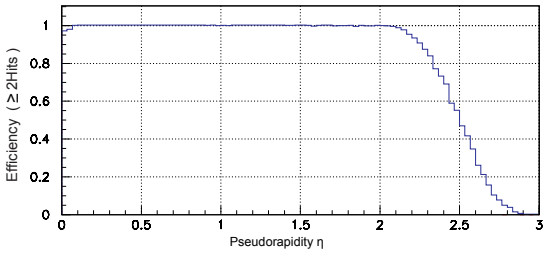

Figure 5.8 Efficiency of receiving at least two hits in the pixel detector as a function of the pseudorapidity η [47]

5.5.2 Silicon Strip Tracker

The strip detector can be subdivided into several components [47]. Overall there are 15 differing sensor geometries. Each sensor comprises 512 or 768 detection strips arranged in 15148 detector modules supporting one or two sensor elements. The active area of the silicon strip detector sums up to $198\,\mathrm{m}^2$ containing about 9.3 million detection strips. The strip pitch varies from $80\,\mu$m in the inner layers of the inner barrels to $184\,\mu$m in parts of the tracker endcaps.

The strip direction in the tracker inner barrel and the tracker outer barrel provides r$-\phi$ measurements (see figure 5.7). Extending from $z = -700\,\mathrm{mm}$ to $z = 700\,\mathrm{mm}$ the tracker inner barrel hosts four roughly equidistant detection layers of cylindrical shape. They surround the beam pipe at radii ranging from $r = 255\,\mathrm{mm}$ to $r = 498\,\mathrm{mm}$. The tracker outer barrel covers the range $|z| < 1090\,\mathrm{mm}$ with six detection layers between $r = 608\,\mathrm{mm}$ and $r = 1080\,\mathrm{mm}$. The tracking system is completed by the tracker inner disks and the tracker endcaps with strips providing measurements in z-ϕ. The inner disks contain three detection layers which are mounted on each side of the inner barrel. They are contained in the coordinate range of $200\,\mathrm{mm} < r < 500\,\mathrm{mm}$ and $800\,\mathrm{mm} < |z| < 900\,\mathrm{mm}$. Each of the tracker endcaps on each side of the tracker outer barrel contains nine disks equipped with sensors positioned in the range of $220\,\mathrm{mm} < r < 1135\,\mathrm{mm}$ and $1240\,\mathrm{mm} < |z| < 2800\,\mathrm{mm}$.

Both for barrel and disk sensors the position of the track along the direction of the strips needs to be measured to allow for a good resolution of all track parameters. Thus some of the layers are equipped with a double sided sensor arrangement (see figure 5.7). The strips in the two sublayers of these “stereo layers” have a stereo angle of 100 mrad. By matching between the corresponding strips the missing coordinate can be recovered if a transversing particle causes hits in both sublayers.

5.6 Muon System

The CMS muon system is designed to identify muons and measure their respective momenta. Due to its time resolution in the order of several nanoseconds it also plays an important role in the CMS trigger scheme (see section 5.7). The system has three basic types of detector components, which are all based on the effect of charge amplification in a gas of suitable composition filling a volume with an electric field of sufficient strength. The charge amplification is triggered by ionized molecules which are caused by particles transversing the detection material (e.g. [62]). Due to the design of the CMS magnet yoke, muon detectors can be mounted not only on the

front and back but also in recessed spaces within the yoke structure. Drift Tube (DT) and Resistive Plate Chamber (RPC) detectors are employed in the five barrel wheels. The detection components in the four roughly disk shaped endcap detection layers are Cathode Strip Chambers (CSC) and again RPCs. The basic arrangement of the muon system components is presented in figure 5.9.

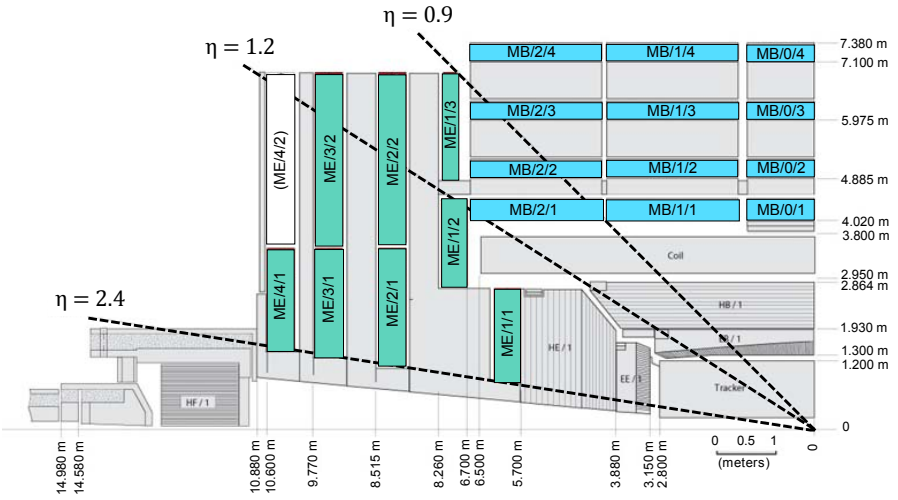

Figure 5.9 Cross section showing one quarter of the CMS detector with highlighted positions of muon system components. The DT stations (blue) can be seen in the barrel region. The CSC components are indicated in the endcaps (green). Attached RPCs are not shown separately. ([47], modified)

If the muon momentum is lower than 300 GeV the momentum resolution without the inclusion of additional information from the inner tracker is expected to be better than 10 % [47]. An exception to this rule of thumb are intersection regions that are not fully covered by the muon detector. The reachable performance in this energy range is mainly limited by multiple scattering. For 1 TeV muons the resolution is reduced to 15 % − 40 % [63]. This is due to the influence of bremsstrahlung which increasingly degrades the detector peformance for energies higher than 100 GeV and also because of the limited resolution of measured track coordinates. Combining track information from the inner tracker and the muon system is expected to decrease this uncertainty to values below 5 % in a momentum range of up to 1 TeV. An optical alignment system using laser beams and LEDs controls the position of the muon detector components with respect to each other and as related to the position of the inner tracker. In this way all positions can be measured with a presicision of less than several hundred micrometers.

5.6.1 Drift Tubes

The drift tube (DT) detectors are mounted in the barrel section of the CMS detector. There are four detector stations. Each of them corresponds to a detection layer of roughly cylindrical shape. In the z direction the barrel is intersected into five wheels which are subdivided into twelve sectors in ϕ. Most sectors comprise one DT chamber for each of the four stations[6]. The Drift Tube systems covers the range $|\eta| < 1.2$. The range $|\eta| < 0.8$ is covered by all four drift tube stations.

The basic detection unit in a Drift Tube chamber is the 42 mm × 13 mm drift cell with a length of about 2.5 m [64]. A gold plated stainless steel wire is stretched in its center . The wire

[6]Due to technical reasons, a few of the sectors in the outermost stations feature two DT chambers which are mounted side by side [47].

provides the anode which is operated at a voltage of about 3.6 kV. The cathode strips are made of 11.5 mm wide aluminium. It is set to a voltage of about −1.2 kV. Two field shaping electrodes (field strips) of 16 mm wide aluminium tape operating at about 1.8 kV are inserted. This allows for a linear relation between the signal time and the distance between the transversing muon and the anode wire. The geometric setup of the mentioned components and the resulting field configuration is shown in figure 5.10. The drift cells are filled with an Ar/CO_2 gas mixture.

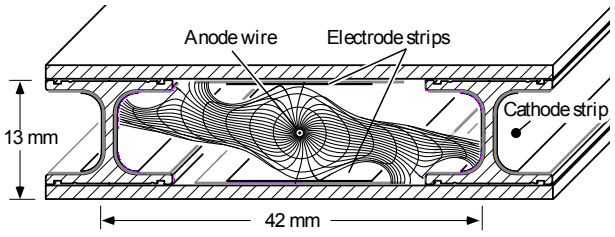

Figure 5.10 Field shaping components and field configuration in a DT drift cell for an external field of $B_{\text{radial}} = 0.8\,\text{T}$ [65].

The basic substructure of each chamber is the DT superlayer. Each superlayer contains four layers of drift tubes in parallel arrangement. Layers are staggered by half a cell to avoid directions of continuous intersection material. Chambers in the outermost station are equipped with two superlayers with cells running in the z direction. Accordingly, both superlayers provide position measurements in r$-\phi$. The chambers in the other stations feature an additional third superlayer with drift cells which are orthogonal to those in the other superlayers. This allows for a dedicated measurement of the z position.

5.6.2 Cathode Strip Chambers

The Cathode Strip Chambers (CSCs) are designed as multiwire proportional chambers of trapezoidal shape. Each chamber covers an angle of about $\Delta\phi = 10°$ to $20°$. In $|\eta|$ a range of up to 2.4 is covered with at least three chambers. In the region $0.9 < |\eta| < 1.2$ muons can be measured with both the CSC and the DT systems. In order to ensure a continuous coverage in ϕ most chambers have some overlap with neighboring chambers in the respective layer [47]. Each chamber features several subsequent panels. In the panel intersections wires are stretched perpendicular to the r direction. An exception is the innermost disk of chambers. In these chambers the wires are tilted at an angle of 29° in ϕ. This compensates for the effects due to the high magnetic field in this region which would otherwise cause a complicated charge spread that is not diagonal to the cathode strips.

Cathode strips running at a constant angle ϕ are milled on one of the two panels opposing each wire layer. The strips have a pitch varying between 8.4 mm and 16 mm. Neighbouring strips have a distance of about 0.5 mm. The CSCs are filled with a mixture of Ar, CO_2 and CF_4. They are operated at a gas gain of about $7 \cdot 10^4$ corresponding to a voltage of 3.6 kV. A typical charge avalanche caused by a transversing muon is spread over several strips which allows for a high position resolution based on the method of charge interpolation. The combination of wires and cathode strips allows for measurements in the $r - \phi$ plane with a precision in the order of 100μm.

5.6.3 Resistive Plate Chambers

Resistive Plate Chamber (RPC) based muon detectors are employed as a complementary detection system both in barrel and endcaps. The CMS RPC chambers are built as double gap chambers operating in avalanche mode. In the barrel region there are two layers of planar rectangular RPC chambers for each of the inner stations. They are fixed to the front and back plane

of the Drift Tube chambers. There is one additional layer for each of the outer two stations which is attached to the front of the respective DT chambers. Each chamber covers a position dependent area of roughly $3 - 4\,\mathrm{m}^2$. The endcap RPCs are of planar trapezoidal shape. The inner RPCs in the endcap regions are mounted in front of the magnet yoke disk which is closest to the interaction point. Further layers are attached to the back of this inner disk and to the front of the third disk. The endcap RPCs are overlapping in ϕ to avoid unequipped material intersections between neighboring chambers. In $|\eta|$ a range of up to 1.6 is covered with RPCs.

Within an RPC chamber a transversing muon may produce charge avalanches in two gas layers (96.2% R134a ($C_2H_2F_4$), 3.5% iC_4H_{10}, 0.3% SF_6 plus water vapour [47]) of about 2 mm depth. The gas layers are confined between highly resistive graphite coated plates. The faster electron component of the avalanches from both gaps is collected on the middle plane which is separating the gaps. Aluminum strips within this middle plane sum the signal from both gaps. They are read out separately. In order to perform p_T measurements strips in the barrel RPCs run in the z direction, while RPC strips in the endcaps run into the r direction at constant ϕ.

The Resistive Plate Chambers have a time resolution which is considerably shorter than the nominal 25 ns bunch spacing [66]. Thus they can be used to achieve a redundant bunch crossing identification and provide an additional fast and independent trigger. Though the precision of a complementary p_T measurement by the RPCs is not as high as for the DT or CSC system the RPCs can contribute to a global measurement and help to resolve ambiguities between tracks.

5.7 Trigger and Data Acquisition

For a 25 ns bunch crossing frequency at the full LHC luminosity the amount of processing power and storage capabilities that would be needed to store the full amount of data taken by the CMS detector is far beyond what is possible to realize. The task of the Level 1 (L1) trigger [67] is to reduce the rate of events to below 100 kHz. This frequency corresponds to the capacity of the next steps of data processing. The trigger criteria are designed to filter out the most interesting events for further analysis. To avoid dead time the Level-1 trigger needs to be capable of performing a trigger decision at the rate of the bunch crossing. The readout of the detector response and the transmission and trigger evaluation of the signal can not be conducted within 25 ns. Therefore the information corresponding to up to 128 bunch crossings needs to be stored locally. This implies a maximum time frame of 3.2 μs until the Level-1 decision triggers the full readout or the rejection of the event data. Thus only fast and efficient algorithms can be used for the Level-1 trigger. Physics requirements include triggers for single objects and multi object triggers with programmable algorithmic procedures and adjustable thresholds for parameters like transverse momenta or isolation requirements. This is necessary in order to balance the relatively large uncertainties on the cross sections and branching ratios of proton-proton interactions at the collision energies provided by the LHC. To meet all of these requirements a complex and custom made system of electronics has been implemented. It combines information from the CMS calorimetry and the muon system to achieve the necessary amount of data reduction while keeping the physically most promising events. It should be emphasized that the inner tracking system is currently not part of the Level-1 trigger scheme[7].

The L1 Muon Trigger combines the information about hits in the Drift Tubes and Cathode Strip Chambers (see section 5.6) into a collection of track candidates. The information is complemented by the evaluation of the strip signal being provided by the Resitive Plate Chambers (see section 5.6.3). The best track candidates are forwarded to the Global Trigger stage. The Calorimeter Trigger provides information about those local energy deposits which are most likely corresponding to interesting physics objects. A sketch of the complete Level-1 trigger scheme is

[7] this might change during future updates of the CMS experiment [68]

shown in figure 5.11. The final L1 trigger decision is determined at the Global trigger stage. It is based on the combined input from calorimetry and muon system.

The data is then transfered to a computer farm at which High Level Triggers (HLTs) reduce the rate to about 100 Hz [69]. All events passing at least one of the HLT criteria are stored permanently for the purpose of further analysis.

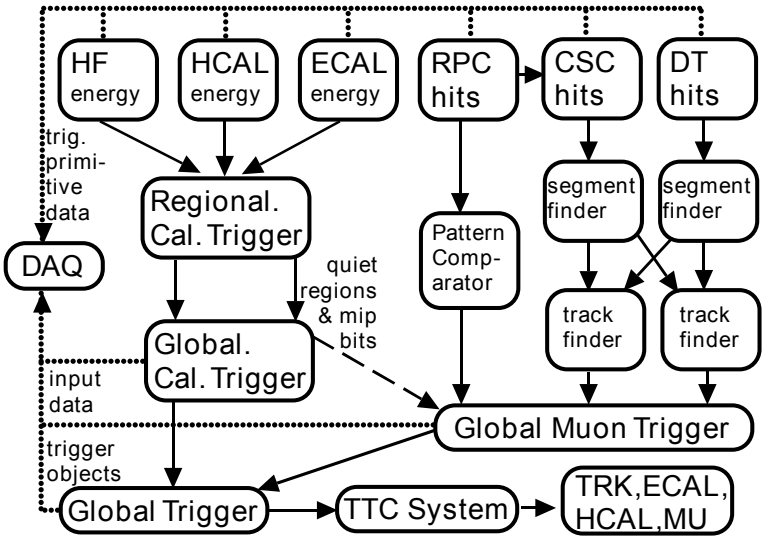

Figure 5.11 The global concept of the CMS Level-1 trigger [67]

6 Software and Analysis Framework

Most modern analyses in experimental particle physics utilize an extended framework of software tools. Only those tools which are of the most direct importance in order to understand the content of this work are covered in the following sections. For example, any details concerning the various generator tools that are used to picture the physics of the proton-proton collision at the LHC had to be omitted for brevity's sake. The following sections introduce MUSiC (**M**odel **U**nspecific **S**earch for New Physics **i**n **C**MS) and discuss the application of software packages and Grid computing within the analysis framework

6.1 CMSSW

CMSSW is the umbrella term for the simulation and reconstruction software used in the CMS experiment. The software package also includes the necessary interfacing with related software components like for example GEANT [70] or event generators like Sherpa [71]. The main parts of the source code are written in C++. Code modules and input/output are channeled and controlled via python based configuration files which are executed by the cmsRun binary.

During simulation and reconstruction the events pass several formats which comprise a selection of the data collections that have been produced up to the respective step. At some points within this work it is referred to these different formats. The corresponding basic functional chain both for data taken by the detector and event simulation is summarized in figure 6.1. The main parts of the CMSSW related results were implemented in the 2.2.X branch. The used event samples from the official production were largely produced with the CMSSW components that were the standard during Summer08 and Fall08 production.

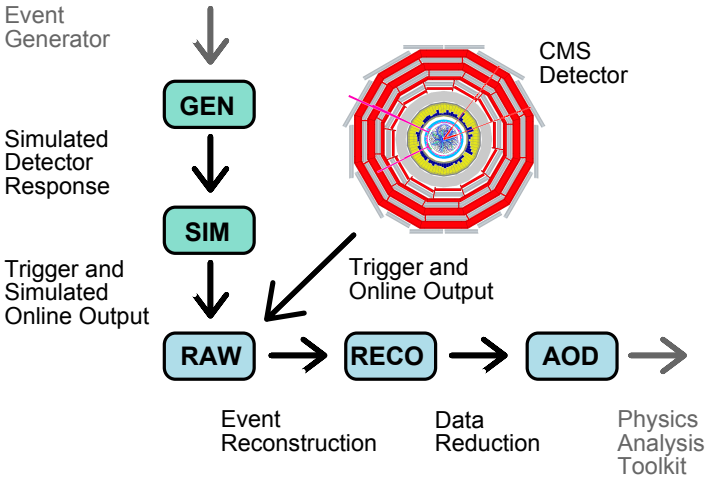

Figure 6.1 Main data formats for data from the CMS detector and event simulation (depiction of the detector from [72])

6.1.1 PAT

Since the introduction of the 2.X version CMSSW contains a newly added component called **P**hysics **A**nalysis **T**oolkit (PAT) [73]. A main feature of the PAT layer is to provide data

collections for important analysis objects, which are clean and easy to use. The idea is to bridge the existing gaps between the needs of experts for the reconstruction of the various physics objects, and the interests of those users who are working on "high level" analyses [74]. Hence the basic output from various algorithms is reprocessed in order to achieve a higher level of transparency and accessibility for the non-specialist. The steps which are performed at PAT layer 0 include matching between different objects, and a reorganization of information and references. In a next step the layer 0 output is stored in the data collections of PAT layer 1. It is important to note that the PAT cleaning steps can be customized according to the needs of a specific analysis. In a rather loose sense can PAT layer 1 be thought of as a standard data format.

In the course of this work the analysis was migrated to CMSSW 2.X. The Physics Analysis Toolkit is now fully embedded into the MUSiC workflow. Concerning the matching algorithms this allowed for some valuable crosschecks between private code and PAT Tools.

6.2 Application of Grid Computing

The CMS experiment will store data in the order of Pbytes per year. Additionally the large scale production of simulated events will easily add hundreds of Tbyte to this number. A completely central storage of all datasets at a single site would be difficult to achieve and also not sensible with respect to data safety. Apart from the requirements in terms of data storage a high amount of computing power is needed for event reconstruction and data analysis.

In order to meet these computational challenges the computing resources associated to the CMS experiment are incorporated into a Grid structure. Important characteristics of a Grid are a decentralized resource coordination, standardized open source protocols and the provision of quality of service [75, page 8]. The infrastructure of the CMS computing grid is based on an association of resource providing institutions with so-called "tiers". Tasks like data storage are coordinated between the different CMS tiers. Details about this structure can be found in [76].

Within the CMS collaboration Grid computing allows for the sharing of resources between users. The shared resources can encompass computing power, storage and services (e.g. software). One effect of this is a more efficient usage of computing elements. As the output of typical analysis steps is by orders of magnitude smaller than the datasets on which one needs to run, it would be extremely inefficient to transfer complete data sets to every user. With the help of the CMS grid jobs can be executed at those sites where the necessary data is already available. If not much external data is required, the distribution of analysis steps to a number of parallelized grid jobs can still be useful to accelerate the analysis.

In the MUSiC analysis (see section 6.6) the CMS grid infrastructure is used to access remote data (see section 6.3) and to achieve a parallelized execution of the event selection and the search algorithm.

6.3 CRAB

The CMS Remote Analysis Builder (CRAB) [77] is a Python based tool which has been developed within the CMS collaboration to provide a user friendly framework for the submission of CMSSW based grid jobs. CRAB automatizes the process of data localization and handles the preparation and submission of jobs. The user defines tasks in the form of CMSSW configuration files and can choose from several options.

The use of the CRAB analysis server provides extended features. For the user an improved job management and monitoring is one of the advantages [78].

In the context of this work CRAB has mainly been used to access sets of simulated events at various remote sites. This is necessary for the collection of data that is needed as an input for

the MUSiC analysis. The job output is given by files in the pxlio[1] format, which are sufficiently small for local storage.

6.4 PXL

The C++ based PXL (Physics eXtension Library) classes [79] provide a data container which has been tailor made for the needs of collider experiments in particle physics. The pxl::Event and pxl::Particle classes [80] come up as natural objects to which most of the relevant data in the context of a high level analysis at CMS can be associated.

One PXL event can contain several EventView objects. Different objects of this type can be used for a separate storage of data that is related to generation, detector simulation and reconstruction of a simulated LHC event. The syntax which is used for recurring tasks like adding variables to an event or retrieving information like the 4-momentum of a single particle is straightforward. Demonstrating code examples can be found in [80].

This work uses PXL 2.0. The version 2.0 includes the concept of "soft links" that can establish a customized net of relations between pxl::Relative objects (from which e.g. pxl::Particle objects are derived). This kind of relation management has been used to control the information that is necessary to establish the various flavors of reconstruction efficiencies and fake rates (discussed in chapter 7).

6.5 ROOT

ROOT [81] provides a framework for scientific data analysis. It has been developed with respect to the specific requirements in modern experimental particle physics like the need to process large quantities of data in an efficient way. ROOT provides powerful libraries including the graphical display of graphs and histograms and a multitude of tools that support the statistical evaluation of data [82]. ROOT is developed in object oriented C++. Hence a C++ based analysis environment as it is is used in this work can easily include and utilize needed classes and libraries. For example, ROOT classes are used for the creation of plots which monitor various aspect of the data and for the graphical presentation of the results from analyzing event classes (see section 6.6) within the search algorithm. The ROOT framework comes along with the scriptable CINT command line interpreter.

6.6 The MUSiC Framework

The MUSiC (**M**odel **U**nspecific **S**earch for New Physics **i**n **C**MS) framework [2] comprises the efforts of building and improving the tools needed for a model unspecific analysis. MUSiC is a constantly evolving project with several active contributors. Some results presented in this work have thus been derived extending previous studies (e.g. [64]) that have been conducted within this framework. Without going into any detail it should be mentioned that parts of the efforts related to the MUSiC project may of course also be considered in terms of extending, updating and maintaining the respective source code.

The following paragraphs outline the main analysis steps while including references to those chapters where the respective steps are discussed in greater detail.

The MUSiC code can be thought of as taking over after the processing and coordination of information within the PAT (see section 6.1.1) layers. Thus the most crucial pieces of information processed in MUSiC are taken from the PAT object collections with only supplementary inputs from the AOD, RECO or SIM levels (see section 6.1). All the data that is of interest for the following analysis steps is then extracted and efficiently stored in the pxlio-format provided by

[1] the format that corresponds to the PXl data container (see section 6.4)

PXL (see section 6.4). In a next step further selection criteria can be applied to the stored objects, like for example additional track quality requirements for muons. For a chosen set of selection criteria up to several hundred customizable control plots can be evaluated and stored for further analysis. The respective code is summarized in the so-called "control plot factory".

For a chosen final set of cuts the events are then distributed to "event classes" [3]. Details about the current status of selection requirements are given in section 7.2. In the event classes the event data is sorted according to the object content. There are two basic kinds of event classes. Exclusive event classes restrict the multiplicity of physics objects to fixed integer values (e.g. exactly one photon and two jets but nothing else). In contrast to that inclusive event classes place only minimal requirements on the object content (e.g. at least one muon and three jets). While this implies that each event can be accepted for insertion into several inclusive event classes an event can only be part of one exclusive class. The event content of exclusive classes is thus disjunct. For each class different distributions of event based variables like the scalar sum of transverse momenta of all accepted objects within an event can be considered for further investigation. A schematic depiction of the MUSiC workflow up to this step is presented in figure 6.2.

In the following steps dedicated algorithms can be applied to search for deviations between the Standard Model expectation and the data. The corresponding algorithmic procedures and the underlying statistical principles are explained in section 8. Simulated examples evaluating event selection and search algorithms in the context of several test scenarios are presented in section 10.

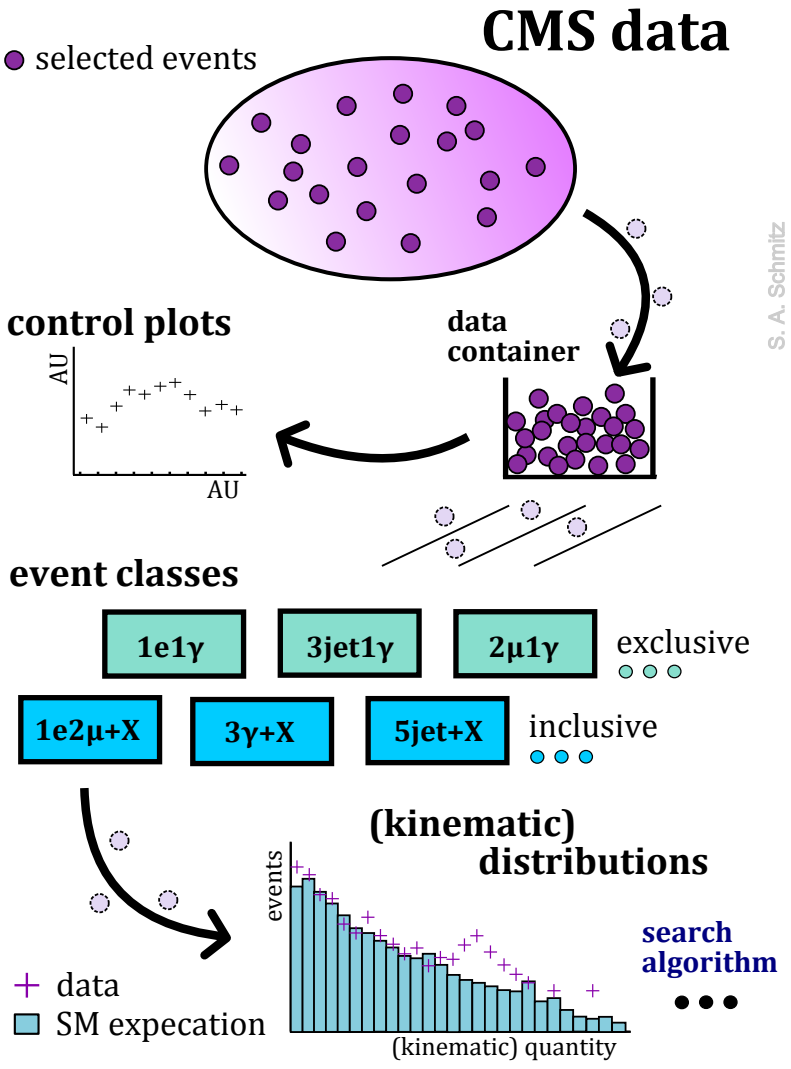

Figure 6.2 Schematic overview of the MUSiC framework excluding the steps within the actual search algorithm

7 Reconstruction and Object Selection

Though the presented test scenarios in section 10 focus on signatures with photons, all objects need to be well understood in order to perform a model unspecific search. This is because most classes combine information from several different physics objects with what is given by the final photon collection.

Obviously, the applied reconstruction algorithms and the subsequent selection steps are subject to an ongoing chain of improvements, and details within the reconstruction may be subject to turbulent developments during the analysis of first data. Thus what is presented has to be understood as a snapshot of current knowledge and methodology.

7.1 Basic Reconstruction Steps

The following paragraphs cover aspects concerning the main algorithms that are needed for the offline reconstruction of physics objects. Only those events are reconstructed which pass the initial selection provided by the L1 Trigger and the High Level Trigger. Special reconstruction methods used in the faster algorithms that are needed for decisions at the trigger stage are not discussed. Respective details can be found in [69].

7.1.1 Finding and Reconstructing Tracks

Finding tracks and measuring their respective parameters is important for the identification and measurement of charged particles. Additionally, tracks contribute to the definition of isolated particles or may be used to distinguish electron positron pairs from converting photons. Track algorithms for a detector characterized by a large number of channels with high occupancy are demanding in computing power. Based on the input from different subdetectors, like the calorimeter, sometimes only parts of the tracking information may be fully reconstructed.

Many of the methods used for track reconstruction within the CMS experiment are based upon the idea of applying concepts from the field of dynamical systems to the problem of evaluating a combined algorithm for track finding and track fitting.

The recursive method of the Kalman filter (see e.g. [83]) is based on a system described by the equations

$$x_k = Ax_{k-1} + w_{k-1} \quad \text{and} \quad z_k = Hx_k + v_k \quad , \tag{7.1}$$

where x_k may be thought of as a vector with five components [84, page 226] describing the state of the transversing particle at detection layer k and z_k is the actual measurement at layer k. The matrices A and H represent the linearized relations between subsequent states and between a given state and the related measurement. While the measurement can be made linear by the choice of the parametrization of the vector x_k, the linear propagation by matrix A is inherently approximative in the presence of a magnetic field permeating the tracking system [85]. This concerns both the inner tracker and the muon system in CMS. The multivariate random variables w_k and v_k represent independent disturbances in the measurement of z_k and the evolution of the process towards x_k. The Kalman filter provides equations that can be used for

Prediction Estimation of x_k based upon all earlier measurements z_{k-i} where $0 < i \leq k$, which can be used to predict the relevant tracker region in the next layer.

Filtering Estimation of the state x_k based upon all measurements z_{k-i} where $0 \leq i \leq k$. This updates the track by including the input from layer k.

Smoothing Estimation of x_k based upon a number of later measurements in layers z_i where $k < i < k + j$.

The Kalman filter provides the optimal solution[1] if w and v are normal random variables while for other distributions it can only be shown to be the optimal linear solution [85]. In [85] also a summarizing set of equations describing the steps of the Kalman filter is given. Electrons can loose a substantial amount of their energy via bremsstrahlung while passing the tracker material of the CMS inner tracking system. Thus a Gaussian Sum Filter (GSF) is applied which provides a non-linear extension of the principles presented for the Kalman Filter [86]. The application of this method results in an improvement if compared to the performance of the Kalman filter, because the influence of bremsstrahlung on the random variable is not well represented by a normal distribution [87]. The main idea in the Gaussian Sum Filter is to approximate the energy loss via bremsstrahlung by a weighted sum of normal distributions (see [88]).

The next paragraphs summarize the main steps within the track reconstruction for electrons and muons.

Electron Track Reconstruction

Electron track reconstruction starts from seeds given by hits in the silicon pixel detector. The evaluated range in the pixel detector is restricted to those regions which are compatible with a reconstructed supercluster in the ECAL (see section 7.1.2). Pairs of two hits fulfilling certain quality criteria are considered as seeds for the formation process of electron tracks [89]. In the next step the GSF algorithm is used to extrapolate to outer layers and build track candidates from compatible hits. Only those tracks having a transverse momentum $p_T > 3\,\text{GeV}$ and meeting certain criteria like comparatively low χ^2-values[2] and a sufficient number of hits are kept (see [91] for actual numbers and additional details). The parameters of remaining track candidates are then fitted by the application of GSF smoothing steps.

Muon Track Reconstruction

Standard muon track reconstruction starts in the muon system. Both inhomogeneities in the magnetic field and multiple scattering need to be taken into account in the track algorithm. Additionally, the influence of bremsstrahlung is growing in importance for energies above 100 GeV. This implies that high energy muons will often be accompanied by photon induced showers which can complicate the track measurement by causing additional hits. Hits in a DT or CSC chamber are combined into local track segments. The track information is supplemented by the RPC measurements. Seeding is provided by hit combinations that are compatible with the beam spot [92]. Starting from the seeds, an algorithm based on the Kalman filter is applied for track finding and fitting. Track candidates are accepted or rejected according to a χ^2 criterion. The "stepping helix" propagator including both field inhomogeneities and material effects [92] is applied for track propagation in the muon system. Using the parameters from the fitting procedure for comparison muon system tracks are matched to the best fitting tracks from the inner tracking system which are built by application of the Kalman filter. In the following step a global fit is conducted combining the matched hits from the inner tracker with the information from the muon system. Each accepted global track results in a global muon candidate.

[1] in terms of minimizing the mean square error of the estimation

[2] χ^2 is an estimator for the goodness of fit (see e. g. [90] for details).

7.1.2 ECAL Clustering

For the reconstruction of electrons and photons the energy deposits in the single ECAL crystals need to be be combined into clusters. This is done in order to recover the full energy of an object. While typically about 97% of the energy will be contained in a 5x5 array of crystals [52], the algorithms have to extend the energy collection in the ϕ direction. The reason is given by the effects of photon conversions and Bremsstrahlung. The two basic concepts being applied at CMS are so-called "hybrid" and "island" clustering algorithms, with the previous being employed in the barrel section and the latter for the endcap regions.

The hybrid algorithm starts from a collection of seeds corresponding to energy deposits above a certain threshold in a crystal. In the next step the energy contribution from the two adjacent crystals in η is added to a seed. The selected crystals are extended to a 1×5 bar[3] if the initial 1×3 bar contains energy deposits above a fixed threshold. This procedure is then iteratively repeated for crystals lying adjacent in ϕ. In this way the algorithm collects 1x3 and 1x5 bars which feature energy deposits above a fixed threshold while extending the supercluster in both directions in ϕ. Iteration in a given direction of ϕ stops if the predefined limit of 17 $1 \times x$ bars has been collected in each direction of ϕ.

In the ECAL endcaps an island algorithm is used to build a set of superclusters. All basic 5x5 clusters built from a subset of local maxima of energy deposition meeting certain criteria (e.g. an assigned transverse energy of more than 400 MeV) are taken as possible seeds for superclusters. Starting from the seed with highest energy a window in η and ϕ is considered for each seed in order to build the supercluster collection. While the interval in η is the same for all seeds, the interval in ϕ can be chosen dynamically according to the energy of the seed. This feature is currently not used during reconstruction (see [91]). If the position of a different cluster is within this search window it is added to the supercluster. Subsequently it is removed from the list of clusters to ensure that a cluster can not be part of multiple superclusters. The energy of the resulting supercluster is given by the sum of energies from its constituents. Its position is the energy weighted sum of the positions from all contributing clusters.

In the endcap regions the superclusters need to be corrected for energy depositions in the preshower detector. In order to improve the achievable resolutions further position dependent correction steps are applied to the superclusters both in the barrel and endcap regions.

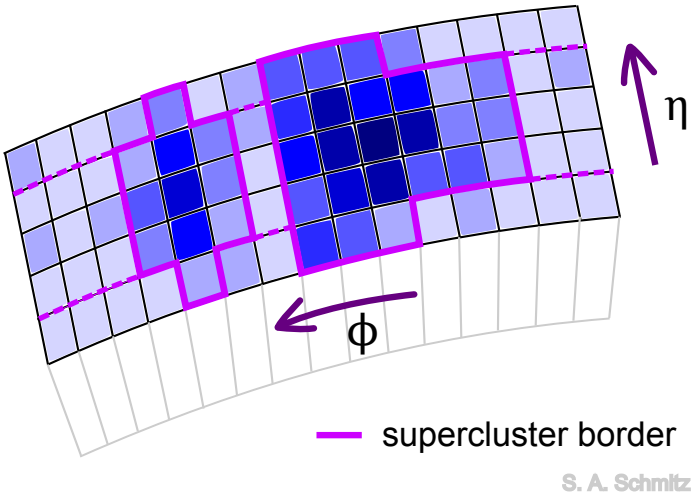

Figure 7.1 Illustration of a barrel supercluster as reconstructed by the hybrid cluster algorithm (not to scale).

[3] number of crystals in ϕ direction $\times$ number of crystals in η direction

7.1.3 Jet Algorithm

A standard approach to jet reconstruction is to use an iterative cone algorithm. Starting from a collection of seeding calorimeter towers the cone algorithm finds a collection of stable jets. In order to test stability a first proto-jet is defined by the energy deposits within a fixed radius R (usually defined in coordinates ϕ and rapidity ψ) around the seed. A distance r between the four-vector of the seed and the sum of all four vectors within the proto-jet is calculated. If this distance exceeds zero, a new proto-jet is built by collecting entries within a cone R around the axis of the first jet. Iteratively, the distance between consequently built jets is calculated, and a new jet is constructed. The jet is stable if r converges to zero within this procedure. Stable jets are removed from the collection of calo towers, and the procedure is repeated until no further stable jets are found.

In practice, a lower limit needs to be set on the energy entries that are used as seeds, because the iterative cone algorithms are infrared unsafe. This means that adding a small energy deposit can lead to new stable jets. Such an additional deposit might for example be caused by pile-up. A frequently encountered example illustrating the problem is illustrated in figure 7.2. Infrared unsafety leads to substantial inconsistencies when comparing results to theoretical predictions from perturbative calculations [93]. While improved iterative cone algorithms adding additional seeds between stable jets have been proposed (e.g. the midpoint algorithm as explained in [94]), those can only reduce the issue of infrared unsafety, but not completely solve it.

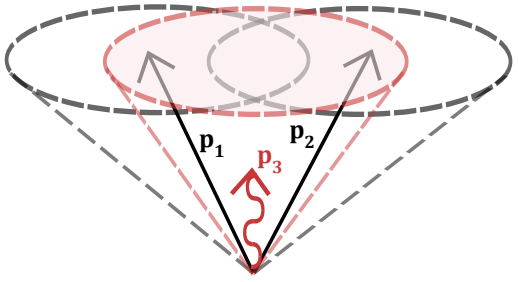

Figure 7.2 Simple example demonstrating infrared unsafety within iterative cone algorithms: Addition of a new seeding soft particle p_3 leading to a new stable cone configuration

Thus several alternative algorithms being inherently infrared safe have been proposed. One of them is the SISCone algorithm [93] whose implementation in the CMS software framework [95] has been employed for jet reconstruction in this work. In this seedless approach the basic idea is to consider pairs (i,j) of particles (i.e. calorimeter towers). For each pair of particles two cones of size R in the (ϕ, ψ) plane can be constructed, for which the particles lie on the edge of the cone. If the corresponding jets have not yet been found, each of the four jets, that can be formed for each of the two cones by including or excluding the entries i and j on the edge, are evaluated [95]. The jet is unstable if the cone around its combined axis does not include the same particles as the original cone [93, page 13]. For the remaining subset of jets stability can be explicitly tested as explained above. This removes the small fraction of remaining unstable jets (which in practice may be negligible). As the particle content of tested cones will be frequently recurring, the bookkeeping about which combinations of particles have already been considered proves to be crucial for a reasonable performance. For a detailed discussion see [93].

Unfortunately none of the mentioned methods ensures that the found set of stable jets is mutually disjunct. Thus a split and merge procedure needs to be appended in order to prevent particles from being part of more than one set. In this step jets sharing more than a certain fraction of energy are merged into a single jet. If this criterion is not met, the calo towers shared by a pair of jets are assigned to the jet with the nearest central axis. After each splitting step

the jet positions are updated. A more detailed discussion including arguments addressing the order of considered jets is presented in [96]. For an analysis of the performance of the SISCone algorithm within the CMSSW framework see [95].

7.1.4 Reconstruction of the Missing Transverse Energy

Though decay products with large longitudinal momentum can escape detection, the hermeticity of the CMS detector ensures that objects that are measurable with the calorimeter can be detected up to $|\eta| = 5$. This implies that all particles with $p_T > 0.013 \cdot E$ are covered [97]. Thus the missing transverse energy $\not{E}_T$ estimating the residual transverse component of the sum of the momenta of all objects that are detectable in CMS can be measured as a meaningful quantity. Aside from potential mismeasurements, weakly interacting particles with a sufficient lifetime are the source of high $\not{E}_T$ events. Within the Standard Model the neutrinos provide such a type of object. Several models beyond the Standard Model predict particles like for example the graviton, which could also result in observed events with significant $\not{E}_T$ contributions.

The starting point of $\not{E}_T$ reconstruction is the missing transverse energy at calorimeter level. It is defined via the n energy deposits E_n in the calorimeter towers as given in equation (7.2).

$$\not{E}_{T,\text{calo}} = \sqrt{\vec{\not{E}}_{T,\text{calo}}} = \sqrt{\left[\sum_n \left(E_n \sin\theta_n \cos\phi_n \hat{\vec{e}}_x + E_n \sin\theta_n \sin\phi_n \hat{\vec{e}}_y\right)\right]^2} \tag{7.2}$$

Further reconstruction steps are applied to correct for several important effects. As the relative calorimeter response is varying with quantities like η or the total amount of energy deposited in a component of the calorimeter, such dependencies need to be corrected for on the basis of a jet collection like the previously described SISC5 jets. After applying these "jet energy scale" corrections $\vec{\not{E}}_{T,\text{calo}}$ is updated via equation (7.3). Here i loops over a subset of the jet collection including jets with a maximum ECAL energy fraction of 0.9 which are above an energy threshold of 10 GeV (see [97] for a more detailed description).

$$\vec{\not{E}}_{T,\text{corrected}} = \vec{\not{E}}_{T,\text{uncorrected}} - \sum_i \left(\vec{p}^{\,i}_{T,\text{corrected}} - \vec{p}^{\,i}_{T,\text{uncorrected}}\right) \tag{7.3}$$

It is clear that $\vec{\not{E}}_{T,\text{calo}}$ also needs to be corrected for potential muons in the event whose momenta can not be determined in the calorimeter, but in the inner tracker and the muon system. Thus the difference between the transverse momentum of each muon and the respective transverse energy deposit of the muon as a minimum ionizing particle must be added to $\not{E}_{T,\text{uncorrected}}$ via (7.3). Now i runs over the global muons with $|\eta| < 5$, $p_T > 10$ GeV and a sufficiently good track measurement [98].

Further, more sophisticated ways of improving the $\vec{\not{E}}_T$ measurement comprise a special correction for jets from decaying τ leptons, additional track based corrections and a dedicated treatment of unclustered calorimeter entries and out of cone effects [98].

7.2 Event and Object Selection

Events are chosen by defining a suitable set of Level 1 and HLT triggers. All events firing at least one of the specified triggers are considered for further evaluation. In the next step, starting from object collections provided by the PAT format, further acceptance and quality cuts are implemented. This is necessary to suppress backgrounds from misidentified objects while balancing fake probabilities with the achievable efficiencies. A main goal is to achieve a selection that is sufficiently clean but still efficient over a wide range of energies and angles. If possible, cut sets which are suggested by expert groups within the CMS collaboration are

adopted. Where such cuts are not found, or considered to be incomplete, or not meeting the needs of a model unspecific search they are changed or replaced by a set of cuts specifically designed for the purpose of the MUSiC analysis.

To limit the amount of data that needs to be handled, and to restrict the algorithm to a controllable range of event topologies which are deemed to be favorable to the aim of finding new physics, there is a further basic restriction. At least one electron, muon or photon needs to be accepted by the quality cuts (implying for example a certain degree of isolation). In previous work conducted within the context of model unspecific search in CMS (see [64] and [2]), only events with at least one isolated electron or muon have been considered. This requirement has now been extended to isolated photons. Events not meeting the criterion of containing at least one isolated photon, electron or muon are discarded. It should be emphasized that, if at least one of the required objects is present, also all other selected objects in the event are stored for further evaluation.

It needs to be checked if the considered objects are not ambiguous reconstructions of the same particle. To avoid such "ghost objects" a simple cleaning procedure checks for a set of selected objects if the candidates do not share substantial parts of their respective characteristics. The applied criteria are summarized in table 7.1.

criterion	**selected object**
muon candidates with $\Delta R < 0.2$	candidate featuring the track with lowest value of χ^2
electrons with $\Delta R < 0.2$ sharing either track or supercluster	candidate featuring the highest estimated energy
photons with $\Delta R < 0.2$ sharing the supercluster	candidate featuring the highest estimated energy

Table 7.1 Applied cleaning steps

After this step possibly remaining ambiguities between objects of different types are resolved by keeping only one of the possible candidates.

While in this work the performance of the MUSiC analysis is evaluated with respect to signatures with high p_T photons, the model unspecific approach forbids an a priori restriction of accompanying objects. Thus it would be rather artificial not to describe at least to some extent all additional physics objects which are by now included in the analysis, though not all of them have an equal share in the specific scenarios discussed in chapter 10. Of course, the photon selection is discussed in somewhat greater detail.

The reconstruction efficiency at simulation level, which is considered to be an important selection criterion, is here defined as

$$\epsilon_{reco}(\alpha) = Pr\left(\text{object}_{\text{gen}(\alpha)} \overset{matches}{\rightarrow} \text{object}_{\text{rec}}\right) \quad . \tag{7.4}$$

"object$_{\text{gen}}$" refers to objects of a certain type (e.g. electron) at generator level and "object$_{\text{rec}}$" to objects of corresponding type at reconstruction level. α is typically chosen from general kinematic quantities like transversal momentum p_T or pseudo rapidity η. Matching is achieved by evaluating a cone of sensibly chosen size Δ_R around the generated object. In order to ensure the functionality of the matching procedure, a private matching algorithm was compared to results from matching steps performed at the PAT (see subsection 6.1.1) stage. For equally chosen matching criteria no differences between the two implementations were found.

7.2.1 Trigger Selection

The trigger selection for a model unspecific search needs to be based on a set of general purpose triggers. Additionally, there is a strong influence on what can be considered a sensible choice

of high level triggers because of the restriction to an evaluation of events with at least one high quality photon, muon or electron with high transverse momentum. For each of these objects dedicated trigger collections exist within the CMS trigger menu. Hence the MUSiC analysis can resort to these standard objects. Care has been taken to choose triggers whose respective thresholds are sufficiently below the cuts applied during the final object selection. This prevents the analysis from being subject to additional difficulties caused by the uncertain behavior of the trigger efficiency in the vicinity of its threshold. An event is selected if it has passed one or more of the specified trigger requirements. Thus the set of selected events $D_{selected}$ corresponds to the union of the sets T_i including all events passing trigger i.

$$D_{selected} = \bigcup_i T_i \quad , \quad i = 1 \ldots \#_{\text{selected triggers}} \tag{7.5}$$

Once actual data is recorded at CMS, one needs to make sure that there is no accidental double counting of events. As datasets might be recorded according to a specific set of fired triggers, single events can turn up within several samples. Of course, each event has a specific unique event ID. Thus one could always solve the problem by some additional bookkeeping. However, it is also possible to define a scheme realizing specific trigger requirements for each type of dataset in order to make sure that events with a certain trigger fired pass the trigger selection in exactly one of the datasets.

Dedicated single and double object triggers are included for each of those object types being sufficient for the event selection. The selected photon specific triggers are presented in table 7.2. The full trigger set for the current MUSiC analysis can be found in table B.7.

name	L1 prerequisite	basic requirement
IsoPhoton20 L1R	e/γ candidate with $p_T > 15\,\text{GeV}$	$p_T > 20\,\text{GeV}$ photon candidate, track isolation criterion
Photon25 L1R	e/γ candidate with $p_T > 15\,\text{GeV}$	photon candidate with $p_T > 25\,\text{GeV}$
DoubleIsoPhoton20 L1R	two e/γ candidates with $p_T > 8\,\text{GeV}$	two $E_T > 20\,\text{GeV}$ photon candidates, track isolation criterion
DoubleIsoPhoton20 L1I	two e/γ candidates with $p_T > 10\,\text{GeV}$	two $E_T > 20\,\text{GeV}$ photon candidates, track isolation criterion

Table 7.2 Photon specific triggers

7.2.2 Photons

Photon candidates are selected from corrected superclusters reconstructed in barrel and endcap sections of the electromagnetic calorimeter. As mentioned above the main purpose of the photon selection is to remain with a set of clean and well understood objects. During selection a reliable identification implying low fake probabilities from hadronic jets and electrons is considered to be more important than a maximized efficiency. It was decided to include converted photons into the selection though the respective efficiency for those objects is allowed to be lower than for non converting photons. As the selection of photons converting before reaching the ECAL by a dedicated algorithm was not yet at a sufficiently mature stage within CMSSW 2.2.X, which has been used for the present study[4], all photon candidates are subject to the same set of cuts. In the future it might be beneficial to introduce a special set of cuts for those objects which are likely to be photons which converted before reaching the ECAL. Especially for possible signatures of new physics with several photons setting converted photons aside would be problematic. It would

[4] there have been a lot of recent developments which are included in the 3.X version

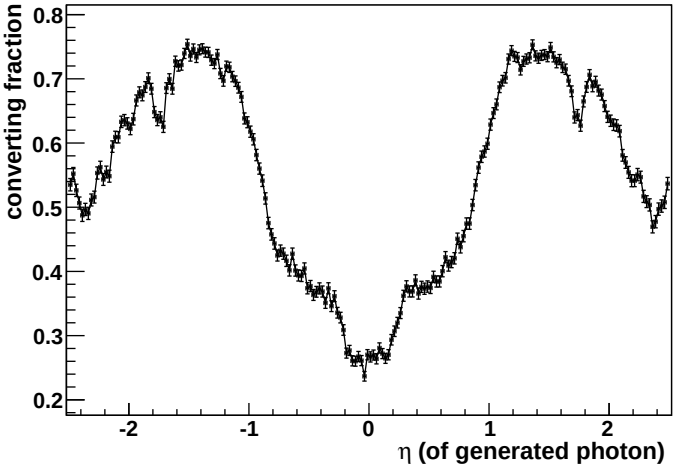

Figure 7.3 Fraction of photons converting before reaching the ECAL as a function of η (linear extrapolation between entries; used photon sample: Photon+Jet p_T80, a sample comprising direct QCD photon production with hard scatterings in the range of $80\,\text{GeV} < \hat{p}_T < 170\,\text{GeV}$ (see table C.2))

mean to accept a highly reduced efficiency and hamper a correct interpretation of results. This becomes obvious when analyzing the fraction of photons converting during the simulation of events before reaching the tracker as shown in figure 7.3. Within the simulation the notion of a converted photon always refers to a generated photon that can be matched to a conversion vertex (photon to electron plus positron) before reaching the ECAL. The necessary information was directly retrieved from the results stored in the SIM output[5] of the CMS detector simulation. The result presented in figure 7.3 is found to be fully consistent with the expected amount of tracker material as shown in figure 5.6. The adopted photon quality and acceptance cuts are presented in table 7.3. The photon selection is restricted to a range in pseudorapidity with $|\eta| < 2.5$ corresponding to the region which is covered both by tracker and ECAL. Only photons with a measured transversal momentum $p_T > 30\,\text{GeV}$ are kept to ensure a good energy measurement. This choice also reflects the assumption that observable signatures of new physics might feature an excess of objects at high transverse momenta.

The following cuts are designed specifically for the MUSiC analysis by studying the spread of the respective variables as a function of variables like energy and transverse momentum. In this way reasonable cut values resulting in an acceptable loss of efficiency could be found. At a first stage several isolation requirements are applied to improve the quality of the measurement and to separate the photons from hadronic jets. A hollow cone track isolation with $0.04 < \Delta R < 0.4$ is adopted to discard objects showing tracks in the vicinity of the direction suggested by the supercluster. The choice of a hollow cone is necessary if photons converting early enough to allow for a track reconstruction are to be kept within the selection. The tracker isolation is chosen to be dependent on the p_T variable, which is well measured in the tracker. In contrast to this the calorimeter cuts are chosen to be linear in the candidate object's energy. The "Jurassic ECAL isolation" suppresses photon candidates with high electromagnetic activity in neighboring crystals. In addition to the $0.06 < \Delta R < 0.4$ hollow cone requirement a strip of width $\Delta\eta = 0.08$ in ϕ direction is excluded from the cone criterion. This compensates for the discrimination against bremsstrahlung from converted photons that would otherwise significantly reduce the photon efficiency. The cut on HCAL isolation in a hollow cone of size $0.1 < \Delta R < 0.4$ reduces the number of objects which originate from hadronic activity. The

[5] see figure 6.1

variable	value	motivation
kinematic cuts		
p_T	$> 30\,\text{GeV}$	reliable energy measurement
$\lvert\eta\rvert$	< 2.5	possible separation from electrons via tracker
isolation cuts		
"Hollow Cone" track isolation	< 5	suppression of background from hadronic jets
"Jurassic" ECAL isolation	$< 8 + 0.0073 \cdot E/\text{GeV}$	suppression of background from hadronic jets while keeping converted photons emitting bremsstrahlung
HCAL isolation	$< 5 + 0.012 \cdot E/\text{GeV}$	suppression of background from hadronic jets
shower shape		
R9	< 0.8	discrimination against strongly radiating electrons and rejection of hadronic jets
others		
pixel seed	false	rejection of electrons and disjunction of photon and electron collection
HoEm	< 0.05	rejection of background from hadronic jets

Table 7.3 Summarized acceptance and quality cuts for photons

rejection of fake objects is enforced by requiring the ratio of energy measurement in HCAL and ECAL, HoEm $= E_{HCAL}/E_{ECAL}$, to be smaller than 0.05.

The shower shape variable R9 is defined by the ratio of the energy deposit in the central 3×3 cluster and the energy deposit in the entire supercluster, $R9 = E_{3\times3}/E_{SC}$. Requiring $R9 > 0.8$ discriminates against hadronic showers which have on average a broader shower development. As an undesirable side effect this cut also rejects converted photons with a significant amount of bremsstrahlung. The effect of the R9 cut as simulated for a sample of QCD photon plus jet events is depicted in figure 7.4. The R9 cut has the advantage of showing sharp distributions with a similar shape in both barrel and endcap regions of the ECAL, and is thus simple and straightforward in its application. Thus the variable has been selected for this study. As a possible alternative one might also consider different shower shape variables which rely mainly on the η shape of the cluster. Those variables are supposed to allow for a better estimation of the photon efficiencies by application of data driven methods using electrons. The pixel seed veto provides important rejection power against electrons. At the same time it rejects a fraction of those photons converting in the pixel seed detector. However, the pixel seed detector has a comparatively low material budget with respect to the complete tracker system (see figure 5.6). Further requirements on the track reconstruction have not been imposed, as electrons failing to produce a seed in the pixel detector can hardly be distinguished from converting photons with high p_T, for which electron and positron feature highly collinear tracks.

Resulting reconstruction efficiencies after application of all cuts as defined in equation (7.4) are presented in figure 7.5. At simulation level a reconstruction efficiency of close to 80 % is achieved for the combined Photon + Jet samples over a wide range in p_T. From studying the plots showing the efficiency as a function of η it is clear that the efficiency as a function of p_T is not necessarily sample independent. Thus the respective efficiency plots for p_T can show

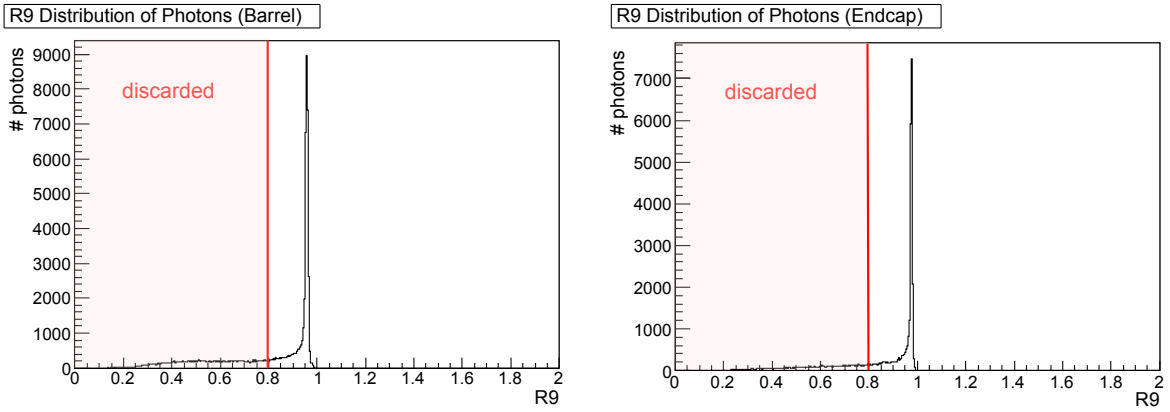

Figure 7.4 Left: R9 distribution for photons from the "Photon+Jets p_T80" sample in the ECAL barrel section. **Right:** the same but for the ECAL endcap section

some dependency on the dominating sample. This results in slight changes in efficiency between those regions where the efficiency is dominated by events from different subsets of the overall Photon + Jets sample. For high values of p_T these observed effects may be increasing due to a feature during reconstruction resulting in a physically meaningless treatment of saturated ECAL crystals. For values of p_T below 200 GeV a reduced efficiency is observed for the converting photons (see label no. 6 in figure 7.5). This is expected, because the tracks from converting photons tend to be more separated for low p_T, resulting in a behavior more distinguished from non converting photons. Additionally, for a given energy low p_T objects imply higher values of η, causing an increased rejection from the pixel seed veto. This may at first seem counterintuitive since the pixel detector performance decreases with increasing η as shown in figure 5.8. But this effect is overcompensated by the influence of the material budget of the pixel detector which increases with η and results in a higher rejection rate for converted photons. This explanation is supported by the distribution of the η dependent efficiencies showing a reduced efficiency for increasing η (see label no. 5 in figure 7.5). Testing the same set of other cuts while eliminating the pixel seed cut results in a significantly increased efficiency in the respective regions. The damping at the edges of the barrel (label no. 4 in figure 7.5) is due to the R9 shower shape cut. With increasing overal tracker material budget the photons convert with an on average higher material budget ahead of them. Hence it is expected that they should have superclusters with shapes diverging more significantly from the typical central cluster of unconverted photons.

Apart from evaluating efficiencies the performance of a chosen set of cuts can be monitored with respect to fake probabilities. During event simulations the probability that an object of type i will incorrectly be identified as an object of type j can be defined as given by equation (7.6).

$$f_{i\mapsto j}\left(\alpha\right) = Pr\left(\text{object}_{\text{gen,i}(\alpha)} \overset{matches}{\rightarrow} \text{object}_{\text{rec,j}}\right) \tag{7.6}$$

The matching is performed by cones of extension ΔR in the same way as explained for the calculation of reconstruction efficiencies. In figure 7.6 the fake probability jet $\rightarrow \gamma$ is shown as a function of η and p_T. The jets used for this evaluation are taken from QCD samples (see table C.2). For example it is estimated that $f_{\text{jet}\mapsto\gamma} \approx 0.0001$ for photons in the central region of the detector with $p_T = 80\,\text{GeV}$. Though the isolation cuts are less strict for higher p_T one can see that with rising p_T the fake probability decreases while the efficiency stays practically constant. This implies that jets of higher transverse momentum are less likely of producing isolated objects. Since the photons from reconstructed jets tend towards an underestimation of the object's p_T on generator level, the fake probability is increasing in the region near to the photon p_T threshold. For the chosen set of cuts and higher values of η the fake probability increases due to the degrading performance of the tracker and calorimeter components. The

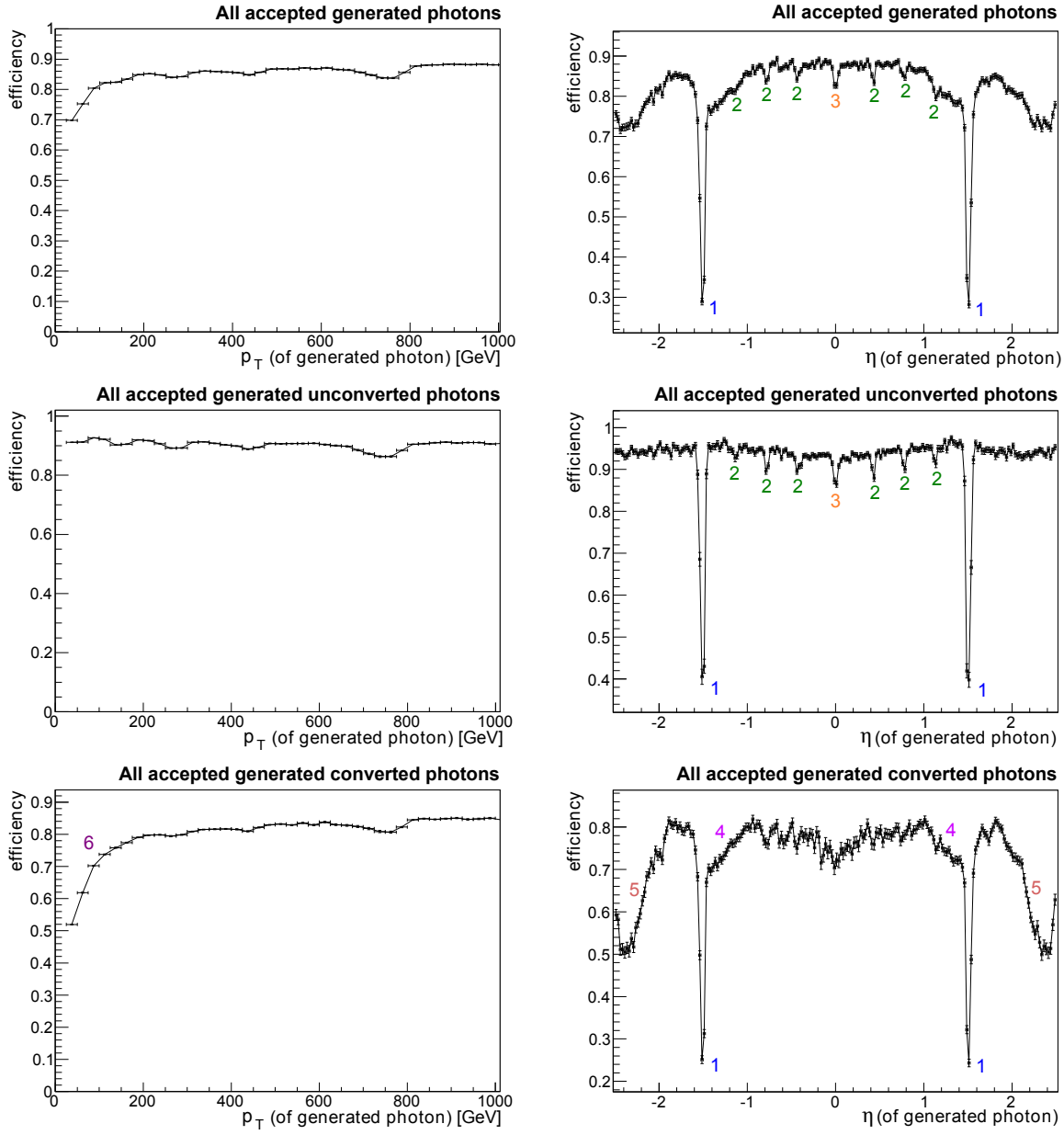

Figure 7.5 Left: reconstruction efficiency as function of p_T for generated photons from combined Photon + Jets samples. **Right:** reconstruction efficiency as function of η for generated photons from the "Photon+Jets p_T80" sample (see figure 7.3)

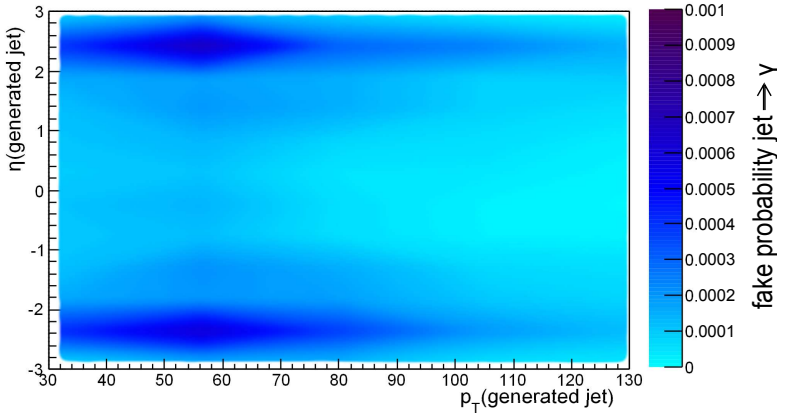

Figure 7.6 Fake probability jet $\mapsto \gamma$ for generated jets as function of p_T and η of the generated jet. The jets are taken from the complete QCD sample.

average fake probability as a function of p_T is shown in figure B.9.

If an electron does not produce any hit in the pixel detector, it may easily be misidentified as a photon, especially if unconverted photons are not discarded. The resulting fake probability for electrons from the W+Jets sample (see table C.2) is shown in figure 7.7. The fake probability is found to be approximately flat in p_T which is not unexpected if the cuts are chosen as shown above and feature an efficiency close to constant in p_T. Evaluating the fake probability of simulated events as a function of η a significant increase of the fake probability is found above $|\eta| \approx 2.3$. Thus, especially during early data taking, a further restriction of the acceptance cut for $|\eta|$ to values smaller than 2.3 should be considered. For the given set of selection variables this would be beneficial for the reduction of fake contributions in the collection of photon candidates. The p_T resolution of the photons selected from a photon+Jets sample is presented in figure 7.8.

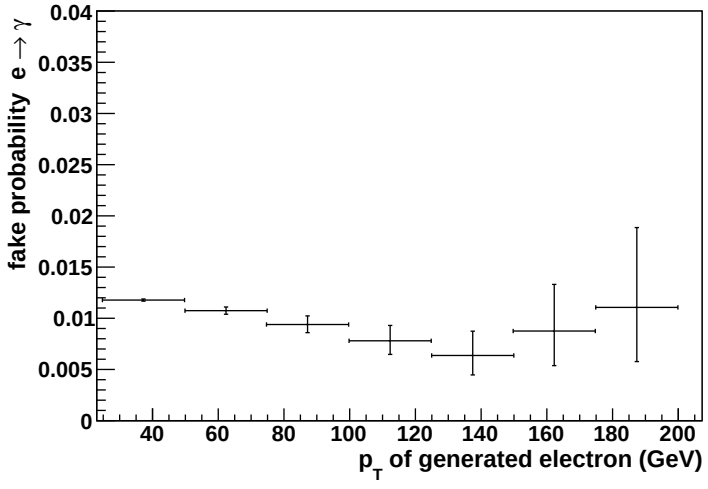

Figure 7.7 Fake probability e $\rightarrow \gamma$ for generated jets as function of p_T of the generated electron. The electrons are taken from the complete W + Jets sample.

As expected from the underlying detection principle resulting in equation (5.5), the simulated relative resolution improves for higher momenta. The tail of $\Delta p_T/p_T$ into the direction of negative values can be explained by energy deposits in non-active material. The additional tail into the direction of positive values for lower values of p_T is due to noise effects and overlapping energy deposits from other particles. Here it has been chosen to show the p_T resolution instead

of the energy resolution as the scalar sum of transverse momenta of the reconstructed objects within a collision event is an observable directly used within the search algorithm presented in section 8.

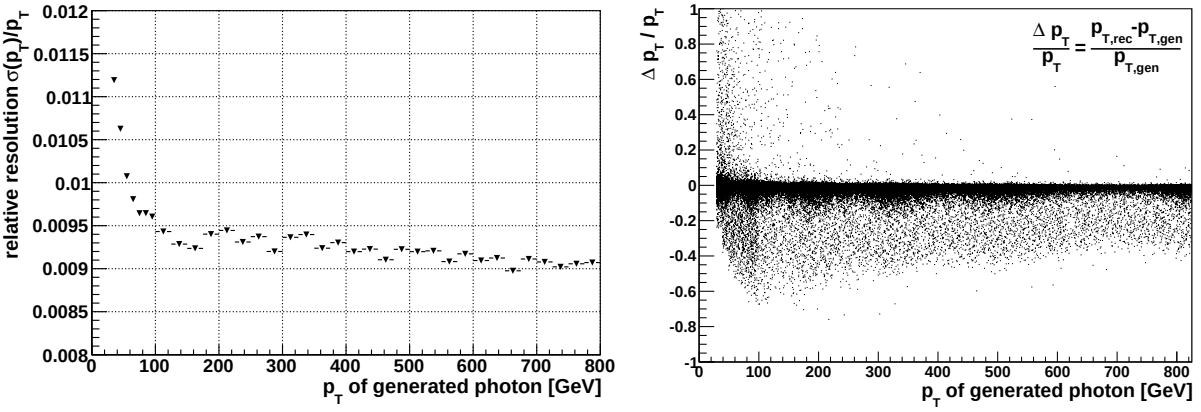

Figure 7.8 Left: relative resolution as function of p_T of the generated photons from the combined Photon + Jets samples. **Right:** scatter plot showing the relative difference between generated and reconstructed transverse momentum p_T as function of p_T of the generated photons from the Photon+Jets p_T80 sample.

7.2.3 Electrons

Electron candidates are built by combining the information from the supercluster with a matched GSF track (see section 7.1.1) featuring a seed in the pixel detector. A robust cut based selection needs to take into account the combined effects of an high magnetic field and the considerable amount of bremsstrahlung due to the tracker material.

Only electron candidates with $|\eta < 2.5|$ are selected. This corresponds to a region in which both calorimetry and tracker show reliable performance. For lower momenta the expected fake contributions are much harder to control. Consequently a p_T cut of 30 GeV is applied in order to reduce fake contributions and to select well measured high energy objects. As we are not interested in electrons produced during the formation of hadronic jets, a relative track isolation is included for the reduction of the corresponding fake objects. Those should often be accompanied by reconstructed tracks in the vicinity of the electron candidate. A maximum distance of 2 mm

variable	value	motivation
kinematic cuts		
p_T	> 30 GeV	object well measured in tracker and ECAL
$\|\eta\|$	< 2.5	tracker and ECAL acceptance
isolation		
track isolation $\Delta R = 0.3$	$\sum p_T/p_T < 0.1$	rejection of non isolated objects (e.g. from hadronic jets)
others		
d_0	< 2mm	discrimination against objects from pile-up interactions
Electron ID	Tight	high quality measurement and background rejection

Table 7.4 Summarized acceptance and quality cuts for electrons

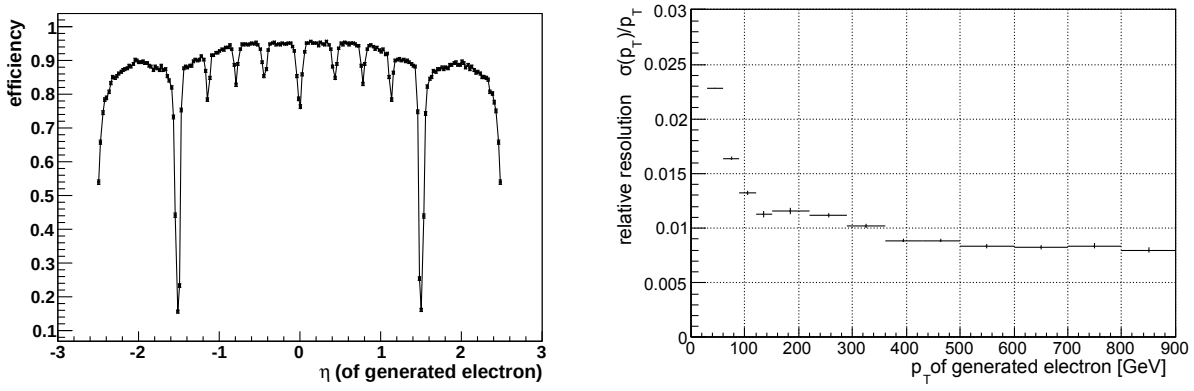

Figure 7.9 Left: electron reconstruction efficiency as a function of η for reconstructed electrons from the W+Jets sample **Right:** relative electron p_T resolution as a function of p_T for reconstructed electrons from W events

between the vertex associated with the electron candidate and the reconstructed primary vertex of the event is requested in the plane orthogonal to the beam direction. During data taking this will support the suppression of objects from pile-up interactions[6]. The cuts for the selection of electrons are summarized in table 7.4.

In order to increase the quality of the MUSiC electron sample a set of recommended cuts is applied which defines a "tight" electron identification [99]. This set of cuts has been designed to feature a high background rejection and a robust performance with respect to the possible influences of alignment during early data taking. The used variables are

1. H/E which is the ratio of the estimated energy fraction deposited in the HCAL and the respective value for the ECAL. Electron fakes from jets are more likely of having a higher value for H/E.

2. the shower shape variable $\sigma_{\eta\eta}$, being defined as

$$\sigma_{\eta\eta} = \frac{\sum_i^{5\times5} w_i \left(\eta_i - \bar{\eta}_{5\times5}\right)^2}{\sum_i^{5\times5} w_i} \quad . \tag{7.7}$$

   Here i runs over all crystals within the central 5×5 cluster and w_i is a tuned weight based on the relative energy of the crystal i [100]. There is an additional correction due to the endcap specific geometry. While the extension of the supercluster in ϕ is only loosely defined, electrons feature a characteristic shower shape in η.

3. $\Delta\eta_{\text{in}}$ which is the difference between the η positions of supercluster and the track at the estimated vertex. These should correspond as tracks are not bent in η.

4. $\Delta\phi_{\text{in}}$ which is the difference between the ϕ positions of the supercluster and the track at the estimated vertex. These should be close to each other if bremsstrahlung has been correctly recovered during the reconstruction of track and supercluster.

5. $E_{\text{seed}}/p_{\text{in}}$ defining the ratio of the energy associated with the seed of the supercluster and the estimated track momentum p_{in} at the vertex.

There are different sets of cuts for the barrel and the endcap regions of the electromagnetic calorimeter. Additionally, the electrons are subdivided into categories according to some basic

[6] Additional proton-proton interactions during bunch crossing

criteria which characterize the electron candidate. The exact definitions of these categories and the respective cut values can be found in the appendix (tables B.5 and B.6).

Figure 7.9 shows two selected control plots demonstrating the performance of the chosen set of cuts during simulation for a selected event sample. The reconstruction efficiency in η is found to be close to 90 %, except for detector intersections, which have already been discussed in the context of photon selection, and small intervals close to the borders of the acceptance region. The second graph depicts the relative p_T resolution.

7.2.4 Muons

The muon selection starts from reconstructed global track muons whose tracks have been refitted according to the combined measurements of the muon system and the inner tracker. As we are interested in objects with high energies, and in order to select muons comfortably above the muon trigger thresholds, a cut in p_T of 30 GeV is applied. The region of acceptance is restricted to $|\eta| < 2.1$. A relative p_T isolation in the inner tracker is used to suppress a significant fraction of muons originating within hadronic jets. The characteristics of the global track fit are natural candidates for controlling the quality of the final muon selection. Hence the χ^2 value per degree of freedom of the global track fit $(\chi^2/\text{dof}) < 10$ and the number of contributing hits N_{hits} are used as cut criteria. The compatibility C_{comp} is a high level quality criterion that utilizes the properties of muons as minimum ionizing particles. It is calculated via

$$C_{\text{comp}} = 0.8 \cdot C_{\text{calo}} + 1.2 \cdot C_{\text{segment}} \quad . \tag{7.8}$$

C_{calo} defines a likelihood variable defined in the range $[0, 1]$. The variable tests the compatibility of the muon hypothesis based on the three measurements of energy deposits in the electromagnetic calorimeter, the hadronic calorimeter and the hadronic outer calorimeter. The variable C_{segment}, which is also defined between 0 and 1, checks to which extent the muon hypothesis is

variable	value	motivation
kinematic cuts		
p_T	$> 30\,\text{Gev}$	focus on high energy objects, trigger threshold
$\|\eta\|$	< 2.1	trigger acceptance and detector coverage
isolation		
track isolation $\Delta R = 0.3$	$\sum p_T/p_T < 0.1$	rejection of non isolated objects from hadronic jets
track quality cuts		
$\frac{\chi^2}{\text{dof}}$	< 10	rejection of fake tracks and mismeasurements
N_{hits}	> 11	rejection of fake tracks and mismeasurements
others		
compatibility C_{comp}	> 1.0	calorimeter and segment compatibility with respect to muon hypothesis
d_0	$< 2\text{mm}$	discrimination against cosmic muons and objects from pile-up interactions

Table 7.5 Summarized acceptance and quality cuts for muons

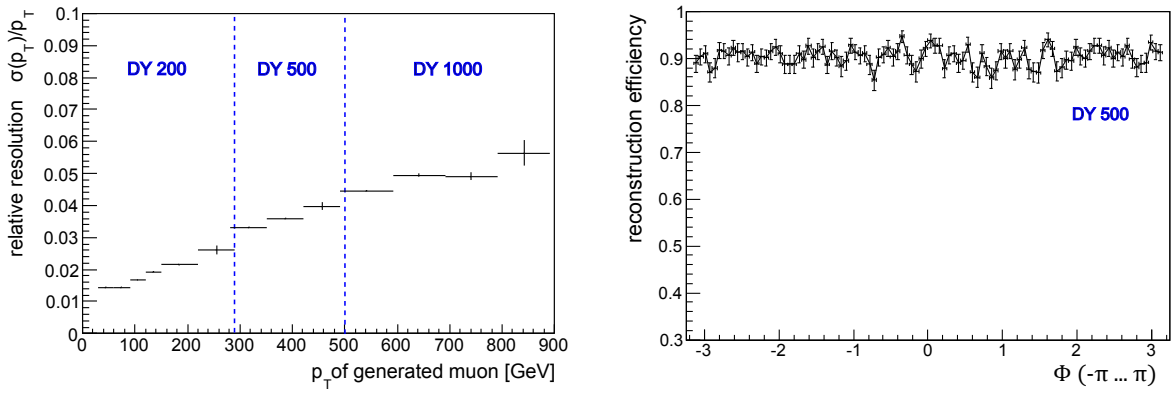

Figure 7.10 Left: relative muon p_T resolution as a function of p_T for reconstructed muons from Drell-Yan offshell samples (see table C.2). **Right:** reconstruction efficiency as a function of ϕ for a selected Drell-Yan offshell sample

in agreement with the found segments in the crossed muon stations. For example, not finding compatible segments in stations which should have been crossed by the muon casts doubt on the nature of the muon candidate. A detailed discussion of the compatibility variables can be found in [101]. Similar to what has been recommended for a tight selection, muons with $C_{\text{comp}} < 1$ are discarded. A cut in the impact parameter d_0 of $d_0 < 2\,\text{mm}$ is applied in order to increase the rejection power against pile-up and cosmic muons. This also partially suppresses decays-in-flight from pions and kaons [101]. All muon acceptance and quality cuts are summarized in table 7.5.

Figure 7.10 shows the muon p_T resolution as a function of the transverse momentum of the generated muon and the reconstruction efficiency as a function of the angle ϕ.

The efficiency as a function of the quantities η or p_T as depicted for photons in figure 7.5 is clearly more interesting in terms of parton interactions because the interactions are expected to be symmetric in ϕ. But exactly this property makes the ϕ dependent efficiency a simple crosscheck of a working detector simulation and event reconstruction. Obviously most of the observed effects in ϕ dependent distributions need to have their origin in aspects of the detector setup. Thus, in order to spot software bugs or other potential problems, it can be valuable to crosscheck these distributions because for them one usually has a clear expectation which is largely decoupled from the used event sample. For the shown Drell-Yan offshell sample (see table C.2) the efficiency is found to be above 90%. This holds true for a large range of possible muon energies. Only for energies above several hundred GeV a non-negligible loss in efficiency was observed independently of the used simulated event sample. This is due to the current performance of the standard global muon reconstruction. Being still under discussion, this issue does not yet seem to be fully understood. The muon p_T resolution is decreasing at higher transverse momenta. That is expected because of the decreasing curvature of the track. Additionally increasing effects from Bremstrahlung may effect the resolution for $p_T \gtrsim 100\,\text{GeV}$.

7.2.5 Jets

The selection of jets starts from the SISC5 jet collection with a cone radius of $\Delta R = 0.5$. The underlying jet algorithm has been outlined in section 7.1.3. Several correction factors are applied to improve the estimation of the 4-momentum of the initial objects at parton level. These "L1-L3" jet corrections comprise the treatment of pile-up and noise levels and address momentum dependent effects due to the non linear response of the detector [102]. Additionally they correct for the position dependent response of the calorimeter. Once CMS is taking data optimized correction factors can be estimated with the help of data driven methods.

Additional criteria are applied to improve the quality and robustness of the selected objects. In order to remain with well measured high p_T objects only jets with a minimum transverse momentum of 60 GeV are selected. A threshold in the pseudo rapidity of $|\eta| < 2.5$ is applied to ensure that the full shower is contained in the acceptance region of the CMS detector. Only jets with a hadronic energy fraction of at least 5% are selected to achieve separation power against photons and electrons. Figure 7.11 presents the jet reconstructon efficiency for simulated QCD multijet events as a function of p_T. The efficiency approaches 100% for jets with very high transverse momenta. It is found to be above 90% over the full spectrum of selected transverse momenta.

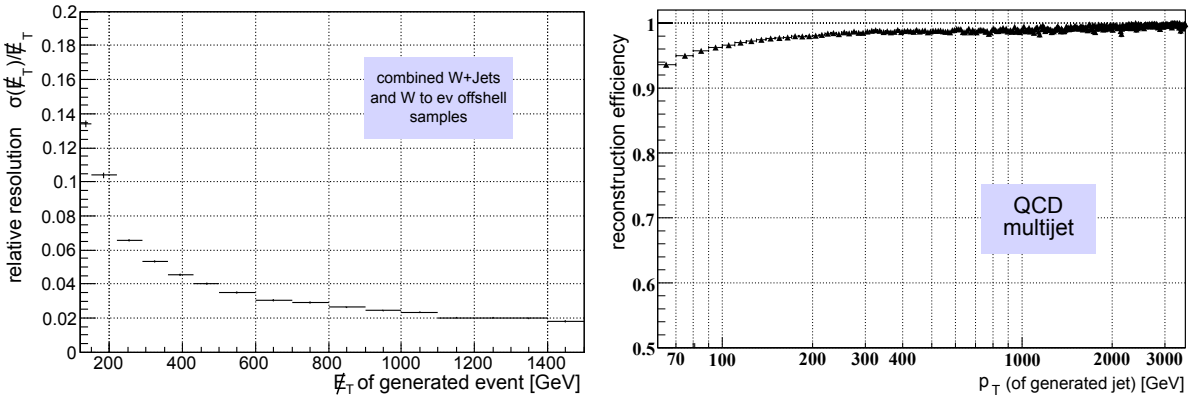

Figure 7.11 Two exemplary performance plots concerning the reconstruction and selection of jets and missing transverse energy. **Left:** relative $\not{E}_T$ resolution as a function of $\not{E}_T$ for events from the W+Jets and the $W \rightarrow e\nu$ offshell samples (see table C.2). **Right:** reconstruction efficiency as a function of p_T for QCD multijet events between 15 GeV and 3000 GeV

7.2.6 Missing Transverse Energy

The $\not{E}_T$ selection starts from a data collection in which the missing transverse energy has already been corrected for jet energy scale effects and muons as outlined in section 7.1.4. During the period of early data taking the $\not{E}_T$ reconstruction will be less predictable than most other reconstructed objects as basically all detector components influence the $\not{E}_T$ performance and therefore need to be reasonably well understood to reduce the involved uncertainties. Hence for now only very basic selection criteria are applied to $\not{E}_T$. In order to ensure a robust identification and a reasonable $\not{E}_T$ resolution the selection restricts the $\not{E}_T$ range of events for which $\not{E}_T$ is evaluated to those featuring a minimum of

$$\not{E}_{T,\mathrm{min}} = 100\,\mathrm{GeV} \quad .$$

Figure 7.11 shows the relative $\not{E}_T$ resolution as a function of $\not{E}_T$ for a combined set of samples with W+Jets and $W \rightarrow e\nu$ offshell events. However, it should be noted that the $\not{E}_T$ resolution is highly dependent on the evaluated sample. For example it clearly makes a difference which fraction of the contributions to the sum of transverse energy in the event is associated with jets and muons. Also the respective object multiplicities have an important influence. For events within the context of supersymmetric scenarios which often tend towards high jet multiplicities a lower $\not{E}_T$ resolution needs to be expected.

8 Search Strategy

After defining a class of recorded events (e.g. two selected photons and three selected jets, see section 6.6) and choosing an appropriate binning[1] for a kinematic variable the data can be represented in a histogram. In each histogram the selected CMS events are compared to the stacked contributions from the selected simulated events that are part of the included generated background samples. In order to be interpretable, the expected contribution of the simulated samples to each class needs to be scaled according to luminosity and cross section. A depiction using a logarithmic scale for the number of expected events is usually suitable for variables like the scalar sum transverse momenta of the selected objects in an event. A generic example showing this kind of graphical representation for a notional class is presented in figure 8.1 for the special case of an equidistant bin size.

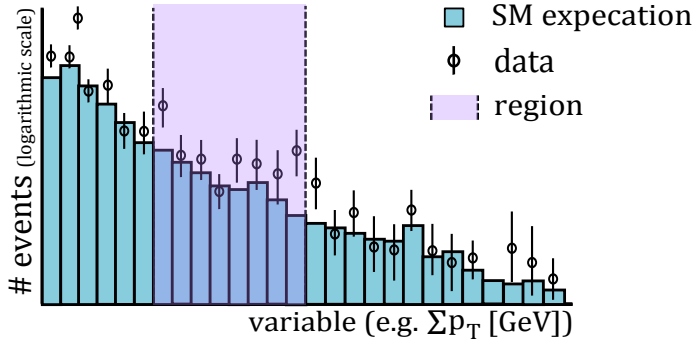

Figure 8.1 Schematic of a generic distribution for a given event class. One of the regions which would be evaluated by the search algorithm is tagged.

While only one-dimensional distributions have been considered in this work, the presented statistical evaluation could also be applied to multidimensional distributions if the potential lack of Monte Carlo expectation either due to the limited number of events or the limited modeling of Standard Model physics can be handled.

Given a single class and distribution one is looking for a quantified assessment of the discrepancy between data and Standard Model expectation. Like in other model independent searches (see e.g. [103] or [104], [105]) a combined set of data bins is referred to as a "region". Identifiable deviations show up as inconsistencies between the recorded and the expected number of events in one or more bins. Without assuming additional information about the location and the shape of possible deviations any combination of bins might be evaluated. In the present algorithm this is accomplished through comparison between the summed number of selected CMS events inside the region and the Standard Model expectation. In order to identify significant deviations the uncertainties on the expected numbers of events in each bin need to be taken into account.

It has been proposed to define regions in a way also transferable to multidimensional distributions, which depends on the recorded events by the use of Voronoi diagrams after transforming to a uniform background expectation [106]. This approach is not applied as the statistical implications of defining regions based on the data seem not to be fully clear. Here all regions are formed by connecting predefined bins. Previously this approach has for example been applied in the evaluation of H1 data [4]. The width of all bins is fixed before looking at the data.

[1] Here binning refers to the definition of a set of intervals in the considered variable (the bins) for which the information is condensed into the number of contained events.

In practice only "connected regions", which exclusively combine adjacent bins, are taken into account by the algorithm. This restriction is useful in order to limit the computational efforts of analyzing each event class and to restrict the statistical effects associated with an increasing number of evaluated regions. The latter aspect is discussed in section 8.2. After defining the set of considered regions one has to evaluate their statistical significance. Accordingly, the discrepancy between data and Standard Model expectation needs to be quantified for every considered region (see sections 8.1 and 8.3). For each distribution evaluated for a given event class the region which is assessed to be most inconsistent with the given expectation is referred to as a "region of interest". The following sections clarify the steps applied in the algorithm and discuss aspects concerning the statistical interpretation of results.

8.1 p-Values as a Measure of Surprise

It has been outlined that for any evaluated region a quantified assessment of the amount of discrepancy between the background expectation and the respective number of events observed in the actual data needs to be defined. This is achieved by the use of "p-values", which are discussed in this section.

Due to the vast amount of regions considered during a model unspecific scan of all considered event classes, it is necessary to condense the measure of surprise into a single number. This number is required to be strictly decreasing for an increase of the deviation and strictly increasing in case of an increasing uncertainty of the background expectation[2]. We are looking for a probabilistic interpretation of the measure of surprise and are interested both in deficits and excesses of observed events. Thus it is sensible to define the discrepancy in each region on the interval $(0, 0.5)$ [3].

p-values are a well-known and widely spread statistical tool for conveying evidence against a hypothesis H_0. Based on the school of statistical reasoning developed by Fisher (see e.g. [107]) they provide a mean of inductively inferring information on the likelihood of H_0 without stating a specific alternative. Therefore their application arises quite naturally in the context of a model unspecific search.

The notion of a p-value [108] can be defined by the equation

$$p \equiv Pr\left(T\left(X\right) \geq T\left(x_{obs}\right) | H_0\right) \quad , \tag{8.1}$$

where $T\left(X\right)$ is a statistic based on the observation X. Hence the p-value is the probability of finding a deviation that in terms of $T\left(X\right)$ is "more extreme" than the value $T\left(x_{obs}\right)$ found in the data if the null hypothesis holds true. Thus observing a low p-value reduces the degree of belief in H_0.

In High Energy Physics it is a common convention to state p-values in terms of tail probabilities Z of the standard normal distribution. They are expressed by the number of standard deviations one has to digress from the mean so that the one-sided or two-sided tail probability corresponds to the numerical value p. The one-sided tail probability relating p and Z is given by

$$\begin{aligned} p &= \frac{1}{\sqrt{2\pi}} \int_Z^\infty \mathrm{d}x \exp\left(-\frac{x^2}{2}\right) = 1 - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^Z \mathrm{d}x \exp\left(-\frac{x^2}{2}\right) \\ &= 1 - \Phi\left(Z\right) = 1 - \frac{1 + \operatorname{erf} Z/\sqrt{2}}{2} \\ \Leftrightarrow \quad Z &= \sqrt{2}\, \operatorname{erf}^{-1}\left(1 - 2\,p\right) \quad , \end{aligned} \tag{8.2}$$

[2] The regions together with the $\leq$ relation for the number signaling the amount of discrepancy form a partially ordered set.

[3] in other situations $(0, 1)$ might be a more intuitive choice

where $\Phi(x)$ is the cdf[4] of the normal distribution and $\mathrm{erf}(x)$ the error function (definition and relations e.g. in [109]). A graph showing the relation between p and Z is displayed in figure 8.2.

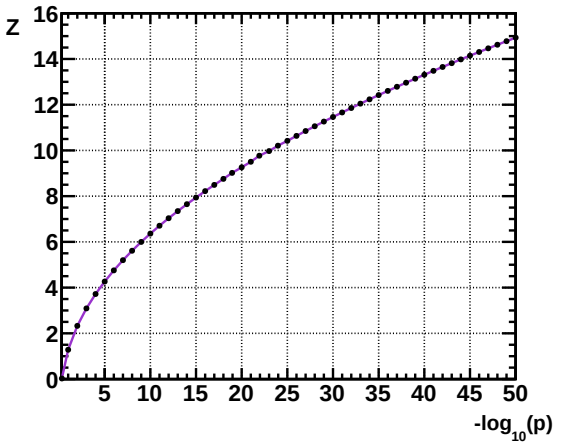

Figure 8.2 Z (one sided) as a function of $-\log_{10}(p)$

In an experiment where one is interested in counting the observed number of some kind of rare event the probability of finding a number N of such events is given by the Poisson probability (e.g. [90])

$$Pr(N) = \frac{\mu^N e^{-\mu}}{N!} \quad , \tag{8.3}$$

where μ is the mean of the distribution. In the given context equation 8.3 describes the probability of observing a number of events N_{obs} in an evaluated region (see above). Of course, the correct value of μ is not known a priori but one can evaluate the significance of the observed number of events under the assumption that μ is given by the Standard Model expectation. The corresponding p-value based upon an observed value $N_{obs} > \mu$ is then defined by

$$p = \sum_{N=N_{obs}}^{\infty} \frac{\mu^N e^{-\mu}}{N!} \quad . \tag{8.4}$$

Difficulties arise when one tries to incorporate an uncertainty on the mean into the definition of the p-value. However, this is of crucial importance as in this work the mean μ is always some kind of estimation of the expected Standard Model background in a specified region. In general this estimation is subject to non-negligible uncertainties (see section 8.4.1). Clearly an increasing uncertainty on the mean should be reflected by an increasing p-value as the vagueness of the expectation for the mean makes an observed deviation more credible.

In practice modeling the uncertainty on μ becomes increasingly difficult if it is necessary to combine the effects from multiple systematic uncertainties into an assessment of μ. The input for the parameters influencing the uncertainty of μ is often addressed in terms of so-called "nuisance parameters" [110]. Unfortunately those can in practice be quite vague in their statistical properties. This implies that formulating a prior function describing the given uncertainty on the background will usually involve some kind of educated guess.

It is possible to include the uncertainty on the background mean into the definition of the p-value by defining

$$p = \int_{R(b)} p(N > N_{data}|b)\, \pi(b)\, db \quad . \tag{8.5}$$

[4] cumulative distribution function

Here $\pi(b)$ is an appropriate prior function for the parameter b and $R(b)$ its range. In the literature on the application and the interpretation of p-values this method is sometimes called the "prior-predictive method" (e.g. [110]). Some authors prefer the notion of a posterior pdf[5] for $\pi(b)$ as it may be derived from one or more auxiliary measurements. For a counting experiment observing an excess of events these ideas translate into a p-value of the form

$$p = \int_0^\infty \sum_{N=N_{obs}}^{\infty} \frac{\mu^N e^{-\mu}}{N!} \pi(\mu)\, d\mu \quad . \tag{8.6}$$

In this work normal and lognormal probability density functions are considered as appropriate choices describing a prior for the uncertainty of the background mean in a counting experiment. A detailed discussion of the resulting p-values is given in section 8.3.

In order to evaluate the performance of a special choice of p-value one may for example investigate the following properties (see [111] for a more detailed discussion of some related arguments, including examples):

- the p-value should increase monotonously with increasing uncertainties[6]
- the possibility to interpret the p-value as a Type-I error rate[7] (coverage properties, see section 9)
- the discrimination power against alternative models
- the adaptability of the chosen p-value to possible changes in the requirements (e.g. inclusion of additional systematic uncertainties)

Those criteria are certainly not exhaustive and may vary in their relative importance based upon the context of the test.

It should be mentioned that there has been concern about the strict statistical meaningfulness of p-values. Problems arise both in the context of Bayesian and Frequentist reasoning. For a short introduction to those two statistical frameworks see e.g. [112].

From the perspective of a Bayesian it is reasonable to compare p-values with posterior probabilities of the kind $Pr(H_0|X)$. Berger and Sellke [113] evaluate the example of testing the mean of a set of random variables X_i that are independent and identically distributed (iid) according to $N(\mu, \sigma)$ against a null hypothesis H_0 for the mean θ, choosing the test statistic

$$T(X) = \sqrt{n}\, \frac{|\overline{X} - \theta_0|}{\sigma} \quad .$$

They find non-negligible discrepancies between the p-values and $Pr(H_0|X)$. Hence there exist cases in which the value of p is misleading as a Bayesian posterior probability. It should be noted that for certain classes of scenarios there have been related efforts to find a remedy by applying certain recalibration techniques to the considered p-value [114], but those could not be translated into an advisable strategy in the context of this work.

Introduced as a means of significance testing by Fisher the methodology is in conflict with the ideas of Neyman and Pearson centering around the notion of Type I and Type II[8] errors (often labeled α and β). It has been argued that confusion between those two fundamental concepts is tending to be ubiquitous in modern scientific work and that it is deeply entangled in many standard textbooks introducing significance testing [115]. An example related to particle

[5] probability density function
[6] for the p-values used in this work this is exemplified in figure B.5
[7] the probability of falsely rejecting H_0 under the assumption that H_0 is true
[8] falsely accepting H_0 under the assumption that H_0 is wrong

physics showing that p-values can disagree with Type I error rates can be found in [111, page 15 et seq.].

A subtlety arises when interpreting p in terms of the number of standard deviations Z as given by equation (8.2). Using one-sided or two-sided tails is merely a convention when strictly following the definition of a p-value. But it is not trivial to make the "right" choice if the p-value is meant to be interpreted as a Type I error rate. Here the following criterion is considered as a rule of thumb for counting experiments. If the p-value is defined in a way suitable for detecting both excess and deficit and one would in principle be willing to refute the hypothesis H_0 based upon a significant deficit the two-sided tail probability is applied. This holds true for the presented model unspecific approach, and hence the two-sided tail probability is used for the evaluation of test scenarios in chapter 10. Otherwise the one-sided alternative is chosen.

Though the concepts of p-values (p) and error rates (α) differ it is a reasonable question if the often found practice of reinterpreting p-values as Type I error rates may be feasible under controlled conditions. For example the question applies to a Frequentist interpretation as it is sometimes found within the context of using the p-value CL_b within in the CL_s method. This approach is frequently used and has been advocated in the CMS community (e.g. [116]). In order to give an answer to this problem a well defined framework is needed. Subsequently, one can examine the Frequentist properties of a given p-value when being interpreted as a Type I error rate. In this context the "exactness" of the p-value [110] plays a crucial role by fixing a terminology for the relationship between error rate α and p.

$$p \quad \text{is} \quad \begin{cases} \text{exact, if } Pr\left(p < \alpha \,|H_0\right) = \alpha \\ \text{conservative, if } Pr\left(p < \alpha \,|H_0\right) < \alpha \\ \text{liberal, if } Pr\left(p < \alpha \,|H_0\right) > \alpha \end{cases} \tag{8.7}$$

Consequently a p-value whose numerical value overstates α is called "conservative" and a p-value underestimating α is referred to as "liberal". Some tests evaluating the exactness of selected p-values are presented in [117]. In [117] the results are presented in terms of "coverage" which is linked to the notion of exactness by the relation

$$\begin{aligned} \text{conservative} \quad &\longleftrightarrow \quad p\text{-value has overcoverage} \\ \text{liberal} \quad &\longleftrightarrow \quad p\text{-value has undercoverage} \quad . \end{aligned} \tag{8.8}$$

The notion of "coverage" is inspired by the use of this term as an important characteristic of ensembles of confidence intervals (see e.g. [118] for a concise definition). Some of the tests that are performed to evaluate the p-values used in this work check and extend upon results that have been presented in [117]. Thus in order to simplify comparisons a largely congruent terminology is used in the following sections.

8.2 The "Look-Elsewhere Effect"

A large number of regions is considered for each distribution in a given event class. Hence the p-value suggested by the single region of interest[9] can not be equal to the significance of the overall distribution. Let p_{min} be the p-value associated with the region of interest of a distribution when comparing data and background expectation. Intuitively it is clear that the more regions one takes into account the more likely it is to find a p-value of a given value p_{min} or lower.

If the effect is not properly taken into account the evaluation of many possibly interesting scenarios might lead to a misinterpretation of significances. This statistical issue is called the "look-elsewhere effect" [119]. In order to make a valid statistical statement about the deviations

[9] meaning the lowest p-value of all regions evaluated for the distribution

within a complete distribution one is interested in the actual probability of finding a p-value smaller than p_{min} for an arbitrary region within the complete distribution. In general correlations between different regions, either due to correlated errors or because of overlap between regions, effect the correct result. This makes it virtually impossible to find a general analytical formula for this problem except for highly simplified scenarios. To address this difficulty one can approximate the result by dicing a large number of "pseudo experiments". For each pseudo experiment the entries of all bins describing the background expectation are fluctuated with respect to all included statistical and systemic uncertainties. Correlations between bins are taken into account. A detailed evaluation of the practical implementation of this procedure is given in section 8.4. Then one states the fraction $\widetilde{P}$ of hypothetical experiments in which a region with $p < p_{min}$ is found because of the expected statistical and systematic uncertainties on the background observation.

$$\widetilde{P} = \frac{\text{number of } H_0 \text{ experiments with a region featuring } p < p_{min}}{\text{total number of } H_0 \text{ experiments}} \tag{8.9}$$

Thus $\widetilde{P}$ can be interpreted as the significance associated with p_{min} after taking the "look-elsewhere-effect" into account.

Of course, one would like to avoid cases in which $\widetilde{P}$ overestimates the significance of an evaluated distribution. Strong overestimations of the significance of $\widetilde{P}$ could result in cases of "false alarm", which would draw undeserved attention to the affected regions and hamper the analysis of deviations which are indeed interesting. Hence $\widetilde{P}$ should be an exact or conservative estimate of the significance in as many cases as possible.

In this context one may be tempted to ask if the property that the p-value is exact, conservative or liberal as defined in equation (8.7) is playing any role for the evaluation of $\widetilde{P}$ which is what is finally stated as the significance of a specific distribution. That in practice the exactness of the found value p_{min} will indeed have an important influence on $\widetilde{P}$ can be understood by evaluating a simplified example.

Consider an experimental setup performing counting experiments in three disjunct regions R1-R3 in which fluctuations are assumed to be independent. Then without loss of generality one may assume that the algorithm finds p_{min} in region R1. The respective value is labeled $p_{1,min}$. Consider the following two cases:

I. The p-value of choice is exact in all of the three considered regions.

II. The p-value is exact in R2 and R3 (i.e. $Pr\left(p_{2,3} < \alpha\right) = \alpha$), but liberal in R1 (i.e. $Pr\left(p_1 < \alpha\right) > \alpha$).

The situation is summarized in table 8.1. In case I the p-value is exact in R1-R3 which implies $p_{1,min} = p_{1,true}$, where $p_{1,true}$ is the actual probability under repetition that $p_1 < p_{min,1}$. Thus $\widetilde{P}$ is given by

$$\begin{aligned}
\widetilde{P}_{\mathrm{I}} &= Pr\left(p < p_{1,min}\right) \\
&= 1 - Pr\left(\left(p_1 > p_{1,min}\right) \wedge \left(p_2 > p_{1,min}\right) \wedge \left(p_3 > p_{1,min}\right)\right) \\
&\overset{\text{independent}}{=} 1 - Pr\left(p_1 > p_{1,min}\right) \cdot Pr\left(p_2 > p_{1,min}\right) \cdot Pr\left(p_3 > p_{1,min}\right) \\
&\overset{\text{p exact in R2,R3}}{=} 1 - \left(1 - p_{1,true}\right)\left(1 - p_{1,min}\right)\left(1 - p_{1,min}\right) \\
&\overset{\text{p exact in R1}}{=} 1 - \left(1 - p_{1,true}\right)^3 \quad .
\end{aligned} \tag{8.10}$$

But when evaluating case II one finds

	case I	case II
R1	$Pr\,(p_1 < p_{1,min}) = p_{1,min}$	$Pr\,(p_1 < p_{1,min}) > p_{1,min}$
R2	$Pr\,(p_2 < p_{1,min}) = p_{1,min}$	$Pr\,(p_2 < p_{1,min}) = p_{1,min}$
R3	$Pr\,(p_3 < p_{1,min}) = p_{1,min}$	$Pr\,(p_3 < p_{1,min}) = p_{1,min}$

Table 8.1 Scenario with three regions illustrating the importance of the exactness of the p-value

$$\begin{aligned}
\widetilde{P}_{\mathrm{II}} &= 1 - Pr\,(p_1 > p_{1,min}) \cdot Pr\,(p_2 > p_{1,min}) \cdot Pr\,(p_3 > p_{1,min}) \\
&\overset{\text{p exact in R2,R3}}{=} 1 - (1 - p_{1,true})\,(1 - p_{1,min})\,(1 - p_{1,min}) \\
&\overset{(p_{1,min} < p_{1,true})}{>} 1 - (1 - p_{1,true})^3 \\
&= \widetilde{P}_{\mathrm{I}} \quad . \qquad (8.11)
\end{aligned}$$

One sees that a p-value liberal in the region of interest can cause an underestimation of $\widetilde{P}$. This would mean assigning too much significance to the region of interest. Hence it is reasonable to state that only the application of a p-value which is close to exact or at least conservative over most of the occurring parameter space can ensure a reliable interpretation of $\widetilde{P}$ as an error rate α.

Generalizing the presented example it is straightforward to show that assuming a globally exact p-value and independently fluctuating background means μ_i , $i = 1\,..\,n$ for n regions $\widetilde{P}$ is given by

$$\widetilde{P} = 1 - (1 - p_{min})^n \quad . \qquad (8.12)$$

It is an interesting question under which conditions this remains a feasible approximation. As dicing a high number of complete pseudo experiments is the most computing intensive step in the application of the algorithm it would be very useful to have an analytical estimation of the expected outcome for $\widetilde{P}$ just from combinatorial arguments. Of course in practice the regions do not fluctuate independently during the process of dicing as systematic uncertainties in different bins may be highly correlated[10], and because most regions can be considered as subsets or combinations of others.

To get a better understanding of the underlying behavior one can again consider the simple example of three regions R1-R3. This time p is assumed to be exact in all cases. Here the evaluated scenarios are that

1. the fluctuations in different regions are statistically independent.
2. the fluctuations are not independent. In this case we will consider the special cases that
   a) the fluctuations are fully correlated.
   b) R1 is independent with respect to R2 and R3, but R2 and R3 are correlated or anticorrelated.
   c) the regions are mutually correlated or mutually anticorrelated.

[10] the uncertainty on the luminosity is an obvious example of an uncertainty that is fully correlated for all bins

$\widetilde{P}_1$ for case 1 can be identified with the special case $n = 3$ for equation (8.12). If the regions are no longer defined to be independent one can derive the result from basic statistical relations (e.g. equations (2.3) to (2.5) in [90]).

$$\begin{aligned}
\widetilde{P}_2 &= 1 - Pr\left((p_1 > p_{min}) \wedge (p_2 > p_{min}) \wedge (p_3 > p_{min})\right) \\
&= 1 - Pr\left((p_1 > p_{min}) \wedge (p_2 > p_{min})\right) \cdot Pr\left(p_3 > p_{min} | \, (p_1 > p_{min}) \wedge (p_2 > p_{min})\right) \\
&= 1 - Pr\left(p_1 > p_{min}\right) \cdot Pr\left(p_2 > p_{min} | p_1 > p_{min}\right) \\
&\quad \cdot Pr\left(p_3 > p_{min} | \, (p_1 > p_{min}) \wedge (p_2 > p_{min})\right) \quad . \qquad (8.13)
\end{aligned}$$

For the case 2a) one finds from equation (8.13) that

$$\begin{aligned}
\widetilde{P}_{2a)} &= 1 - Pr\left(p_1 > p_{min}\right) \cdot 1 \cdot 1 \\
&\overset{\text{p exact}}{=} p_{min} \quad . \qquad (8.14)
\end{aligned}$$

Not surprisingly one obtains $\widetilde{P} = p_{min}$ in this limit. More interesting is case 2b).

$$\begin{aligned}
\widetilde{P}_{2b)} &= 1 - Pr\left(p_1 > p_{min}\right) \cdot Pr\left(p_1 > p_{min}\right) \cdot Pr\left(p_3 > p_{min} | p_2 > p_{min}\right) \\
&= 1 - Pr\left(p_1 > p_{min}\right) \cdot Pr\left(p_1 > p_{min}\right) \cdot \kappa \cdot Pr\left(p_3 > p_{min}\right) \\
&= 1 - \left(1 - p_{min}\right)^3 \cdot \kappa \quad ; \kappa > 0 \quad . \qquad (8.15)
\end{aligned}$$

If R2 and R3 are correlated then $\kappa > 1$. In the case of anticorrelation one finds $\kappa < 1$. Hence one sees that $\widetilde{P}_{2b)} < \widetilde{P}_1$ for correlation and $\widetilde{P}_{2b)} > \widetilde{P}_1$ for anticorrelation.

Similarly one can find that $\widetilde{P}_{2c)} < \widetilde{P}_1$ when R1-R3 are mutually correlated and $\widetilde{P}_{2c)} < \widetilde{P}_1$ when they are mutually anticorrelated. Those results can be generalized to n regions with a given set of correlations or anticorrelations. If correlations and anticorrelations are not mixed existing correlations lead to an overestimation of $\widetilde{P}$ (conservative). Anticorrelations cause an underestimation of $\widetilde{P}$ (liberal). As many of the considered statistical dependencies between regions disfavor anticorrelations one might conjecture that using (8.12) leads to conservative results in many but not all practical cases.

In summary it is possible conclude that in order to avoid misinterpretations both the application of a predominantly exact or conservative p-value and an appropriate treatment of correlations between regions is of high importance for the MUSiC analysis.

8.3 p_N and p_{LN} as p-Values for a Model Unspecific Search

In the search algorithm a p-value is used to evaluate the deviation between data and Standard Model expectation in every considered region (see section 8.1). Currently two p-values of the kind presented in equation (8.5) are fully implemented in the analysis framework. One of them has already been used for previous results [2] and also in the context of studies within different experiments like the H1 detector at HERA [103]. The second p-value (p_{LN}) is applied to a model unspecific search of the presented scope for the first time. As defined in equation (8.16) p_N uses a normal prior function to model the uncertainty on the background mean.

$$p_{\mathrm{N}} = \begin{cases} \displaystyle\sum_{i=n_{on}}^{\infty} C \cdot \int_0^{\infty} \mathrm{d}b \, \exp\left(\frac{-\left(b - b_0\right)^2}{2\,\sigma_b^2}\right) \cdot \frac{e^{-b}\, b^i}{i!} & \text{if } n_{on} > b_0 \\ \displaystyle\sum_{i=0}^{n_{on}} C \cdot \int_0^{\infty} \mathrm{d}b \, \exp\left(\frac{-\left(b - b_0\right)^2}{2\,\sigma_b^2}\right) \cdot \frac{e^{-b}\, b^i}{i!} & \text{if } n_{on} < b_0 \end{cases} \qquad (8.16)$$

Here n_{on} is the number of observed events, b_0 the mean of the background expectation and σ^2 the variance of the background prior. The constant factor C ensures the normalization of the normal pdf, which is used as a prior function, on the interval $[0, \infty)$. p_N can only be expected to be fully appropriate if the combined errors contributing to the background prior can be modeled with a normal distribution. An additive contribution of random variables favors a normal model because of the central limit theorem. A straightforward example is the summation of different background contributions with independent uncertainties on their respective cross sections. If independent uncertainties can be treated with a normal pdf then their sum is also given by a normal distribution (e.g. [120, page 49]). This property can be used to combine uncertainties. It is also important for the design of a consistent strategy for the dicing of pseudo experiments.

During a global scan of all considered event classes there are regions where the standard deviation σ_b of the normal background prior is not small if compared to its mean b_0. In those cases the restriction of b to values with physical meaning above 0 by normalizing with the factor C has a considerable effect on the shape of the distribution. This truncation of the normal pdf at $b_0 = 0$ leads to a " (singly) truncated normal pdf" [120, page 80]. Such a model may be considered to be somehow arbitrary. For example one could argue that a pdf like the truncated normal pdf not fulfilling $\lim_{b_0 \to 0} = 0$ is unintuitive as this implies that for a segmentation into sufficiently small intervals of size ϵ the probability is practically constant in the vicinity of $b_0 = 0$.

A different argument for the consideration of alternatives to a normal prior is that one can argue that the estimation of the combined prior for the background mean is often described by a product and not by a sum of several prior functions. For a single background process this argument is supported by equation (5.1).

For such a scenario one can separately model the uncertainty on each variable contributing to the number of events and estimate the resulting background prior by the product of the priors of each single uncertainty. Then it might be more appropriate to apply a p-value using a lognormal prior. If the n factors building the product are each following normal distributions then in general their product follows a complicated distribution for which there seems to be no closed analytic expression. For the case of $n = 2$ some results evaluating special cases can be found in [121]. In this context the lognormal pdf turns out to have interesting properties. Some of the important properties of the lognormal pdf are summarized in appendix A. One advantageous feature of the lognormal pdf is that a product of lognormal random variables again follows a lognormal distribution for which the parameters are easy to calculate[11]. Thus the lognormal distribution is a convincing choice for the prior if the background can be described by a product of lognormal distributed variables.

The mathematical expression for a p-value with a lognormal prior is presented in equation (8.17).

$$p_{\mathrm{LN}} = \begin{cases} \displaystyle\sum_{i=n_{on}}^{\infty} \int_0^{\infty} \mathrm{d}b \, \frac{1}{\sqrt{2\pi}\ln k} \cdot \frac{1}{b} \cdot \exp\left(\frac{-\ln^2(b/b_0)}{2\,\ln^2 k}\right) \cdot \frac{e^{-b}\, b^i}{i!} & \text{if } n_{on} \geq b_0 \\ \displaystyle\sum_{i=0}^{n_{on}} \int_0^{\infty} \mathrm{d}b \, \frac{1}{\sqrt{2\pi}\ln k} \cdot \frac{1}{b} \cdot \exp\left(\frac{-\ln^2(b/b_0)}{2\,\ln^2 k}\right) \cdot \frac{e^{-b}\, b^i}{i!} & \text{if } n_{on} < b_0 \end{cases} \tag{8.17}$$

This notation is based on the parametrization proposed in [122] which is the only reference known to the author where this type of p-value has already been considered for a practical application in an analysis. Here b_0 is the median of the background distribution and k is chosen to be $1 + \sigma_{rel}$, where σ_{rel} is the estimated relative uncertainty on b_0. This evaluation of a "relative" uncertainty is a shift with respect to the normal prior for p_N which is mainly concerned with

[11] for details see the paragraph leading to equation (A.14) in appendix A

"absolute" uncertainties on b_0. That this difference implies more than just a terminological subtlety is shown in section 9.

Some examples of lognormal probability densities are presented in figure 8.3 and in figures B.1 to B.3.

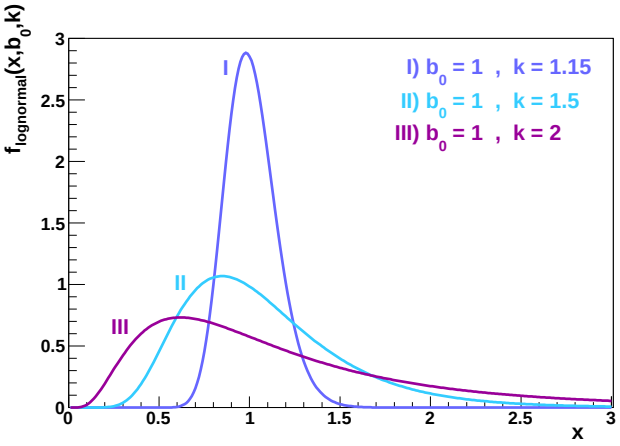

Figure 8.3 Lognormal pdf with parametrization (A.7) for different values of k if b_0 is fixed to 1

Consider the case where the combination of all errors can be modeled by a lognormal distribution. Then for k close to one the shape of the lognormal pdf approaches a normal distribution $N(\mu, \sigma)$ with mean $\mu = b_0$. Some motivation for this behavior is given in appendix A. Making use of this property one can ensure that the predictions of p_N and p_{LN} are similar for sufficiently small errors if the median b_0 of the lognormal distribution is chosen to be equal to μ and $k = 1 + \sigma_{rel} = 1 + \sigma/b_0$ as suggested in [122]. An example for the case of $k = 1.1$ is presented in figure 8.4.

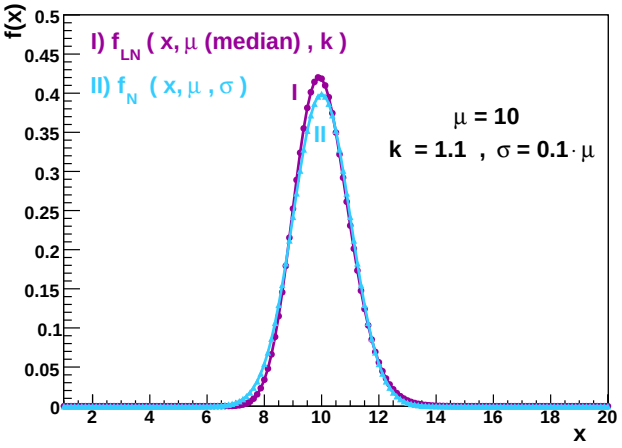

Figure 8.4 Lognormal pdf approaching a normal distribution for k close to 1 (linear extrapolation between evaluated points)

Additionally the proposed choice of parameters has the nice property that the background interval $(b_0 \cdot k\,,\, b_0/k)$ always corresponds to 68% of integrated probability (the special case of $c = 1$ for equation (A.12)). If σ_{rel} is interpreted as a relative uncertainty this is an intuitively reasonable behavior for large errors, i.e. k not close to 1.

The shape of the lognormal distribution differs increasingly from the normal distribution if k deviates from 1. For $k = 1.6$ this is exemplified in figure 8.5 in which the lognormal distribution

is compared to a normal distribution that has a mean equal to the median of the lognormal pdf and a standard deviation of $\sigma = 0.6 \cdot \mu$.

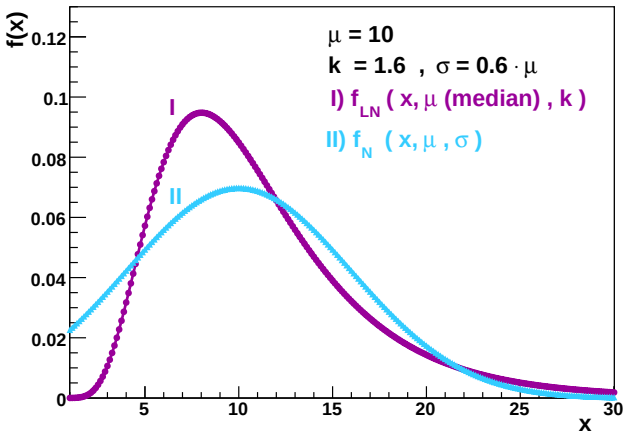

Figure 8.5 Lognormal distribution for $k = 1.6$ comparing to a normal distribution with a relative uncertainty of 0.6

It can be shown that for every normal distribution $N(x, \mu, \sigma)$ one can find a limit x_0 so that compared to a lognormal distribution $\Lambda(x, b_0, k)$ it holds true that $\Lambda(x, b_0, k) > N(x, \mu, \sigma)$ for $x > x_0$. Expressed less formally it means that the lognormal pdf has a larger upper tail than the normal distribution. Considering that the lognormal pdf is a heavy-tailed distribution this is certainly not surprising. As for the given applications one is trying to estimate background means this implies that for a lognormal model more probability is assigned to extreme upward deviations from the assumed median. This directly translates into higher p-values for cases in which a large excess with respect to the expected number of events is observed. As only backgrounds with $b > n_{on}$ can make considerable contributions to (8.6), lower expected backgrounds are effectively suppressed by the Poisson distribution in (8.6). If the used p-values are transformed into normal standard deviations via equation (8.2) they are labeled according to those subscripts that are used for the respective p-values leading to the notation Z_N and Z_{LN}.

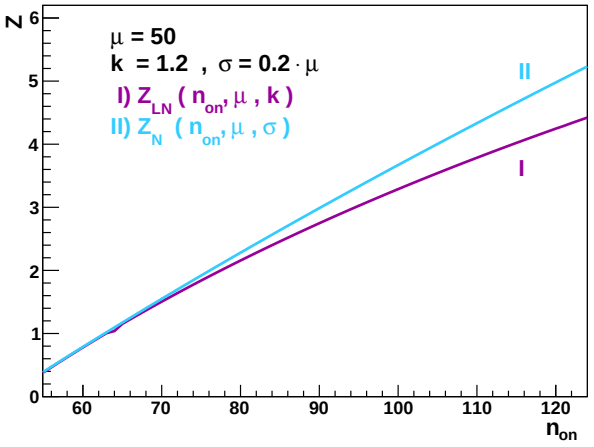

Figure 8.6 Z_{LN} and Z_N as a function of the observed number of events n_{on} (other parameters fixed)

In case of an observed excess and choosing $k = 1 + \sigma/\mu$ and $b_0 = \mu$ the resulting p-value p_{LN} will be more conservative than a p-value p_N with parameters μ and σ. This is shown for two examples in figures 8.6 and B.4. However, it should be noted that for cases in which less events than expected are observed the resulting p-value for the lognormal prior can be considerably smaller than p_N as it assigns less probability to backgrounds close to zero (see figure 8.5, exemplifying p-values are given in table B.9). One should also keep in mind that both p-values are functions of three parameters. For example one may compare p_{LN} and p_N for a varying background median μ while keeping k and σ_b fixed at corresponding values and choosing n_{on} constant. In this case there are regions in which p_{LN} predicts lower values for Z even if considering an excess of events. This is exemplified in figure 8.7.

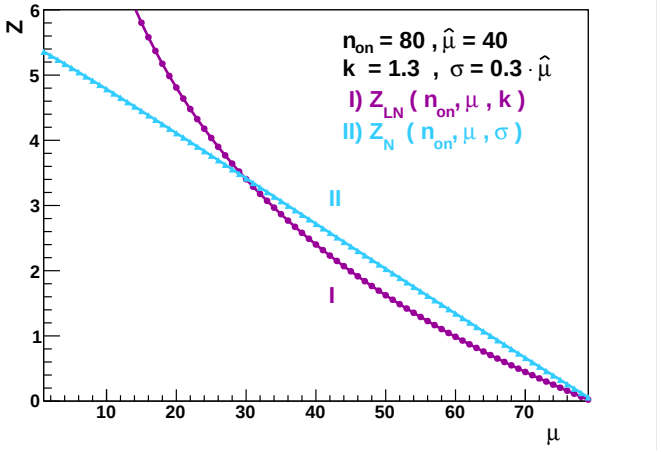

Figure 8.7 Z_{LN} and Z_N as a function of the background median μ (other parameters fixed)

Table 8.2 compares results for Z_N and Z_{LN} for several numbers n_{on} of observed events, if the suggested choices of parameters are applied to a background expectation of 10 and a relative error of 20 %. Results for an extended number of parameters can be found in appendix B. Both Z_N and Z_{LN} were found to be in agreement with an independent implementation of the p-values [123].

The question remains how one should combine the information about the different sources of statistical and systematic uncertainties into a choice for the parameters of p_N and p_{LN}. One also has to implement a consistent dicing of H_0 pseudo experiments. These problems are addressed in the following sections.

$\mu = b_0 = 10$, $k = 1.2$, $\sigma_{\mathrm{rel}} = 0.2$		
n_{on}	Z_{LN}	Z_N
15	1.15	1.18
20	2.22	2.27
25	3.15	3.25
30	3.98	4.16
35	4.73	5.01
50	6.61	7.73

Table 8.2 Evaluation of Z_{LN} and Z_N assuming a background expectation of 10 and a relative uncertainty of 20 %

8.4 Dicing of Pseudo Experiments

It has been explained in section 8.2 that in order to calculate the significance of the value p_{min} by taking into account the number of regions and their correlations it is necessary to dice large numbers of pseudo-experiments according to the background only hypothesis. It is shown in chapter 9 that reasonable coverage properties for $\tilde{P}$ can be expected for both p_N and p_{LN} if the resulting background distribution for a given region under repetition of pseudo experiments corresponds to the background prior of the p-value. In practice this consistency is difficult to achieve using p_{LN}. In this case a pragmatic procedure needs to be applied as an approximate solution. For both of the considered p-values motivating arguments are at hand. One of them is the possible underestimation of tails in the normal distribution in the case of multiplicative systematic influences.

Two dicing strategies are considered. One leading to the assumption of a normal distribution and one assuming lognormal errors for each background bin. In order to offer a concise description the first dicing strategy is labeled D_N, and the alternative one which has been customized in the context of p_{LN} is referred to as D_{LN}.

D_N Dicing strategy with the aim of combining systematic errors into a normal distribution for each bin

D_{LN} Dicing strategy with the aim of combining systematic errors into a lognormal distribution for each bin

For a single error it is often difficult to decide if it is modeled better by a normal or a lognormal distribution. Thus for a given relative error a normal prior $N(\mu, \sigma_{rel})$ is assumed for D_N and a lognormal prior $\Lambda(\mu, 1+\sigma_{rel})$ for D_{LN}. As shown in section 8.3 this implies that there are no noteworthy differences between the priors if the relative error is small. For D_N the diced value for a bin with n_0 originally expected events is then calculated by

$$n_i = n_{i-1} + n_0 \cdot X_{N\left(x,0,\sigma_{rel,i}\right)} \quad , i{>}0 \tag{8.18}$$

where n_{i-1} is the value before considering uncertainty i and X a random value drawn for this pseudo experiment according to the respective normal pdf. Successive application of (8.18) corresponds to adding normal distributed quantities and leads to a normal distribution. As equation (5.1) suggests, some combinations of uncertainties may be modeled better by products of prior functions. This approach is applied when working within the framework of D_{LN}. In this case the uncertainties are treated via

$$N' = n_0 \cdot X_{\Lambda\left(x,b_0=1,k=1+\sigma_{rel,i}\right)} \tag{8.19}$$

In practice one can get lognormal random variables from standard normal random variables by the application of

$$X_{\Lambda(b_0,k)} = b_0 \cdot e^{\ln k \cdot X_{N(0,1)}} \quad . \tag{8.20}$$

This can be derived from the relation between the involved distributions (see e.g. [124, page 28]). The iterative application of (8.19) to different uncertainties guarantees a resulting lognormal distribution because of equation (A.14). After dicing the background mean n_{sys} according to the modeled systematic uncertainties the final bin entry for the pseudo experiments is drawn from a Poisson distribution with mean n_{sys}.

Unfortunately, there is one important subtlety arising in the dicing procedure for p_{LN} when combining different bins into a region. While the outlined procedure works fine for each single bin the distribution for the observed sum of events in a region combining bins is no longer of strictly lognormal shape. Thus, even if for each region the errors are combined into a lognormal distribution which would be a reasonable choice by itself, the assumptions can not be in

full concordance for regions with more than one bin. As outlined in section 8.2 this leads to inconsistencies when dicing hypothetical experiments by dicing each bin according to Poisson experiments drawing their mean from the background prior. For p_{LN} the background prior is assumed to have a lognormal shape. Several conceivable direct solutions to this problem are found to be either impractical or cannot be implemented within the scope of this work:

- One could dice each region with respect to the lognormal function characterizing the background prior for p_{LN}. Though straightforward in implementation this would severely reduce the speed of dicing pseudo experiments and slow down the algorithm by a factor that is not acceptable.

- Literature suggests [125] that a sum of correlated random variables with lognormal distribution may under certain conditions again be approximated by a lognormal pdf. In this way one could achieve an approximate consistency by estimating the lognormal prior for a combined region from the lognormal priors of the contributing bins. A general approach using a recursive algorithm to estimate the needed parameters is presented in [126]. Unfortunately such an algorithm would make the search strategy less straightforward and be rather computing intensive. Additionally it is not fully clear how well the approximation would perform in practice.

- Obviously summing the observed event in a number of bins results in an observable which is much in favor of being treated under the assumption of normal distributions, because if the distribution of a linear function $Y = \sum_i a_i \cdot X_i$ is supposed to have the same type of distribution as each of the random variables X_i this ensures normality [120, page 51]. Thus it might be possible to look for other observables, like for example the product of events in each bin, which would be suited better under lognormal conditions. This idea might be worth further investigation but it would make the interpretation of results more complicated and less intuitive.

- If the p-value has good coverage one might consider to calculate the product of the p-values for each bin. This would lead to a new combined estimation of the significance. This would be similar to likelihood based approaches and imply a strategy being somewhat different from the original MUSiC algorithm.

- Finally, one could try to find a general numerical solution for approximating the sum of correlated lognormal random variables. Probably such a solution, if existing, would be either highly complicated or very computing intensive.

The mentioned drawbacks show that all of these options are not fully satisfactory. A further possible solution is given by the following "trick". Assuming actual data it is certainly possible to evaluate every region using p_{LN}. Under the hypothesis that the influence of the Look-Elsewhere-Effect is mainly caused by the multiplicity of the regions and their correlations, but largely independent from the difference in the applied background prior, one can correct $p_{min,LN}$ by dicing background only experiments using p_N. In section 10 the problem is addressed in this way. Though the underlying hypothesis seems to be a good guess it would be useful to make further investigations during future studies.

8.4.1 Handling of Statistical and Systematic Uncertainties

This section discusses the different systematic uncertainties that have been included in this study. In consistency with the results presented in chapter 10 the assumptions are based on a scenario with an integrated luminosity in the range of $100\,\mathrm{pb}^{-1}$ to $1\,\mathrm{fb}^{-1}$ at a center of mass energy of 10 TeV.

Luminosity

Though it is envisioned to reduce the relative uncertainty on the luminosity to a level of about $1-2\,\%$ [127] this precision can not be reached after recording data correponding to $100\,\mathrm{pb}^{-1}$ to $1\,\mathrm{fb}^{-1}$ of integrated luminosity. At this stage a relative error of 5 % or 10 % provides a realistic estimation [128, page 532, revised IOP version]. Thus a conservative uncertainty of 10% is assumed for this study. The resulting uncertainty on the background mean is implemented by the application of (8.18) for p_N and (8.19) for p_{LN}. The uncertainty is fully correlated for all bins.

Parton Distribution Functions

Uncertainties on cross sections for a given process P in a hadronic collision do not only stem from the interaction between the directly involved elementary particles but also from the restricted knowledge about their contribution to the composite object at a given inelasticity x. By application of the QCD factorization theorem the cross sections for proton proton interactions can be described by equation (8.21) [129].

$$\sigma\left(pp \to X_P\right) \propto \sum_i \sum_j C_{ij}^P \left(x_1, x_2, \alpha_S(Q^2)\right) \otimes f_i\left(x_1, Q^2\right) \otimes f_j\left(x_2, Q^2\right) \tag{8.21}$$

In this formula f_i and f_j are the parton distribution functions (PDFs) of the partons of type i and j within the two interacting protons. Q^2 is the energy scale of the process and the coefficient functions $C_{ij}^P\left(x_1, x_2, \alpha_S(Q^2)\right)$ describe the underlying hard process P. Several approaches to an estimation of the uncertainties on the parton distribution functions have been developed. One of them is the "Hessian formalism". The main idea can be summarized by the following steps. Starting from a χ^2 function to quantify the discrepancy between data and theory one determines a set of best fitting PDFs from a global data set. For this a sufficiently flexible model with d free parameters is assumed. Then a base of d eigenvectors of the Hessian matrix[12] is obtained. In the next step one constructs $2 \cdot d$ PDF sets varying the parameters up and down by a certain quantity in χ^2 along the directions given by the eigenvectors. In this way one receives an estimation of the uncertainties by studying the consequent variations of the considered observables. A detailed discussion of the method can be found in [130].

In practice it may be too extensive to repeat the generation and simulation of a sufficient number of events for every set of PDFs. This is certainly the case for a model unspecific approach with many final state topologies and several observables. Instead one can estimate the uncertainties on a per event basis under the assumption that the PDF contributions to equation (8.21) can be factorized out. In this case one calculates a new weight for each given event and PDF set k [131] by applying equation (8.22). The case $k = 0$ corresponds to the best fit PDF.

$$w_{ij}^k = \frac{f_i^k\left(x_i, Q^2\right) \cdot f_j^k\left(x_j, Q^2\right)}{f_i^0\left(x_i, Q^2\right) \cdot f_j^0\left(x_j, Q^2\right)} \tag{8.22}$$

For a practical justification of the reweighting procedure and an instruction to implementing PDF uncertainties in CMSSW one may refer to [132]. In this work the PDF set CTEQ6.1 [133] has been used. CTEQ6.1 features 40 eigenvector sets for the parton distribution functions and one additional best fit parametrization.

Additional details concerning the inclusion of PDF uncertainties into the MUSiC framework can be found in [134]. For the dicing of hypothetical experiments the resulting variations on final distributions due to differing choices of PDF eigenvector sets are summarized into a single normal or lognormal prior function which is incorporated via equation (8.18) or (8.19).

[12] the matrix one receives by evaluating the second derivatives of χ^2 at the minimum [130]

Cross Sections

Apart from the uncertainties imposed by the parton density functions there are further uncertainties on the LHC cross sections. The quantum field calculations which provide the hard parton-parton cross sections for a given process are always executed at a limited order in perturbation theory. Thus the knowledge about higher order corrections is restricted though it may be possible to make an assessment of the possible differences. Additionally there are uncertainties on the low energy QCD calculations that are used for the decay chains and the shower development of hadronic components. For the moment these effects are combined into an additional global uncertainty of 10% for the total cross sections of all backgrounds. This uncertainty is also applied to all signal samples which are used for the test scenarios presented in chapter 10. As parton density functions are treated seperately, this pragmatic approach is conservative for many of the contributing processes and analysed variables. The framework allows to assume a different uncertainty for each background process if additional information is available. It would also be possible to incorporate dependencies on relevant kinematic variables like the transverse momentum in a future development. The statistical incorporation into the dicing procedures is achieved by the application of equations (8.18) and (8.19). The uncertainty is considered to be fully correlated within each process but independent for different backgrounds.

Efficiencies

Especially when extrapolating to very high energies the efficiency of reconstructing objects is subject to systematic uncertainties. In the case of leptons and photons the W and Z standard candles can be used to estimate the corresponding efficiencies from data. For the case of muons this is described in [135] but the respective uncertainties will further increase when extrapolating to very high energies. For now a global relative uncertainty on the reconstruction efficiency of 2 % for electrons, muons and photons and 1 % for jets is assumed. Errors on different types of objects are assumed to be independent.

For dicing according to strategy D_N a normal prior with relative error $\sigma_{rel,i}$ is used to model the uncertainty on each object of type A_i. For each pseudo-experiment the diced efficiencies are applied to all bins. Then for a class requiring N_i objects of type A_i the effects are combined by applying (8.18) according to $\sigma_{rel,i}$ N_i times for each object type i.

When working with D_{LN} one can assume a lognormal shape for the prior related to $\sigma_{rel,i}$. Then the uncertainty on the reconstruction efficiency of N_i required objects of type A_i is modeled by multiplying the value realized in the respective pseudo experiment N_i times.

Fake Probabilities

Based on simulated events several assessments of fake contributions to reconstructed photons have been presented in section 7.1. Such studies provide information on the rough probability $f_{A\mapsto B}$ that an object of type A is misidentified as an object of type B. Usually the fake probabilities depend on additional parameters like p_T. The estimation of fake contributions from event simulations allows to estimate backgrounds, but the simulated fake probabilities are certainly not precise. The results may vary significantly with respect to inputs like the alignment, the actual detector material budget (e.g. in case of $f_{\gamma\mapsto e}$), the magnetic field and other influences. Thus reliable evaluations of $f_{A\mapsto B}$ can only be established from data. Even then the corresponding relative uncertainties on $f_{A\mapsto B}$ might still be substantial. There are ongoing studies concerning the estimation of fake probabilities within the CMS collaboration (see e.g. [136] or [137]). Once data based results are available from such studies it is possible to correct the weight of simulated background events with fake objects with respect to the measurements. For now this is not the case and only an assessment of the expected uncertainty on such measurements can be taken into account for the evaluation of test scenarios.

The inclusion of the uncertainty on fake contributions into the algorithm is implemented by a reweighting of all included background events. Consider an event with N_i fake objects of type A_i. A fake object is defined to be an object which is not matched to a generated object of the same type within $\Delta R < 0.2$ ($\Delta R < 0.5$ for MET) (see also section 7.1). It is assumed that the weight w of an event is given by

$$w = c \cdot \prod_i f_i^{N_i} \quad ; c \in \mathbb{R} \quad .$$

The parameters f_i are the combined probabilities of faking an object of type i. Assuming standard error propagation the relative change in weight is

$$\sigma_{w,rel} = \sqrt{\sum_i N_i^2 \sigma_{f_i,rel}} \tag{8.23}$$

where $\sigma_{f_i,rel}$ is the relative uncertainty on faking an object of type i. All events are reweighted by applying

$$w' = w\,(1 \pm \sigma_{w,rel}) \quad . \tag{8.24}$$

Then for each bin one can compare between the expected number of events in the reweighted and the original event samples. The resulting difference provides an estimate of the relative error which should be expected due to the uncertainty on fake probabilities. The relative differences are used to model a prior function which is included by application of (8.18) or (8.19). For each pseudo experiment only one value is drawn from the prior function and applied to all bins in all distributions. It is clear that there are several underlying simplifications in this procedure. For example possible kinematic dependencies of f_i are not yet taken into account and the influences of trigger effects are not independently taken into account. For now it is assumed that all fake probabilities are only known up to a factor of two. This implies $f_{i,rel} = 1$.

Jet Energy Scale

For several reasons there is a remaining uncertainty on the jet energy measurement. At the analysis level one is usually interested in the properties of the parton from which the jet originates. Both theoretical and experimental aspects cause problems for a correct calibration of the jet energy scale. The theoretical effects include details of the fragmentation model and aspects related to initial and final state radiation. Further complications are caused by multiple parton interactions and pile-up [138]. A related issue are performance parameters of the jet algorithm like the energy deposited outside of the cone during event reconstruction.

On the experimental side there are uncertainties related to the detector response. Several studies have been dedicated to the preparation of strategies for the jet energy calibration at CMS. They comprise studies evaluating dijet events [139], photon plus jet production [140] and Z events with one additional jet [138]. Though the actual uncertainties certainly depend on parameters like η or p_T the studies suggest that working with a global uncertainty on the jet energy scale of 5% is a conservative assumption when considering scenarios with an integrated luminosity between $100\,\mathrm{pb}^{-1}$ and $1\,\mathrm{fb}^{-1}$. In order to assess the respective uncertainties for the considered distributions the preselected jets are rescaled both up and down by 5%. The MET variable is corrected for the residual change in p_T of the rescaled jet collection. Then the final distributions for each event class are reevaluated. The difference between the scaled samples and the unscaled reference sample is used to model the uncertainty for each bin. The averaged deviation from the unscaled sample is chosen as a $1\,\sigma$ environment for the estimation of a normal prior for D_N and for the estimation a lognormal distribution in the context of equation (8.19). In each pseudo experiment the jet energy scale is fully correlated for all bins. Figure 8.8 shows the influence on the p_T spectrum of a simulated $t\bar{t}$+Jets sample from rescaling the jet energies by 5 % .

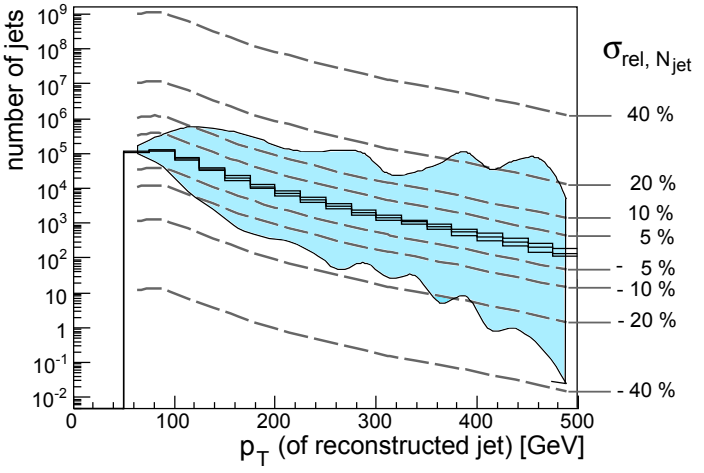

Figure 8.8 Simulated jet p_T spectra at reconstruction level from $t\bar{t}$+Jets for the reference values from the simulated jet reconstruction and for rescaling the jet energy scale once up and once down by 5%. The colored band marks the respective relative change of the number of events per bin, $\sigma_{\mathrm{rel},\,N_{jet}}$, as a function of p_T. The dashed lines label the reference scale for $\sigma_{\mathrm{rel},\,N_{jet}}$.

Statistical Uncertainty on the Number of Simulated Events

The background prediction in each region is based on a limited number of simulated events. Accordingly, the estimation of each simulated background is subject to a statistical uncertainty on the number of events that contribute to a given bin. This aspect is of special importance when large scale factors are involved as it is for example often the case for the QCD multi-jet background.

Though such an uncertainty can always be reduced by the simulation of additional events there are often practical limits due to the finite amount of available computing resources. During the dicing procedure the uncertainties on each process are combined into an assessment of the resulting relative error. The variances of those Poisson distributions that one would expect if taking the original number of events as their respective mean are added in quadrature. For each pseudo experiment the uncertainty is included independently for each bin (eq. (8.18) or eq. (8.19)). Though this kind of evaluation is only approximate it is considered to model the size of the effect correctly and allows for an increased consistency between the dicing and the evaluation of the p-value.

Statistical Uncertainty on the Observed Number of Events

For each pseudo experiment the resulting number of events in a bin is drawn from a Poisson distribution. The mean is the expected value of the background distribution which has been diced according to all considered systematic uncertainties.

8.4.2 Choice of Parameters for p_N and p_{LN}

During the discussion of systematic uncertainties two methods for repetitive dicing of H_0 (the Standard Model expectation) have been outlined. For the background mean in each bin one of them leads to normal distributions for a consistent application of p_N and the other to lognormal distributions for a bin-wise consistent application of p_{LN}. For a given region it is necessary to combine the different uncertainties for the included bins to a combined background prior of

normal shape for the application of p_N and to a prior of lognormal shape for the application of p_{LN}.

Application of p_N

One can make use of the fact that adding independent normal distributions leads again to normal distributions via equation 8.25.

$$\begin{aligned} &\text{if } X \sim N(\mu_1, \sigma_1) \text{ and } Y \sim N(\mu_2, \sigma_2) \text{ then} \\ &X + Y \sim N(\mu_1 + \mu_2, \sqrt{(\sigma_1^2 + \sigma_2^2)}) \end{aligned} \tag{8.25}$$

In this way the parameters according to which the systematic errors have been handled can be combined into one global normal pdf for each region. This pdf is then used as the background prior of p_N.

Application to p_{LN}

Treating errors according to (8.18) one multiplies lognormal distributions. This implies that the combined shape for the pdf can be built by making use of equation (A.14) if the uncertainties are fully correlated between the bins summarized into a region. In this way it is possible to handle the uncertainties on fake probabilities, efficiencies, luminosity, parton density functions and jet energy scales. Uncertainties on cross sections and statistics of simulated events are additive and can cause deviations from the assumed lognormal shape when combined for different processes (cross sections) or bins (number of simulated events). It has been shown in section 8.3 that for sufficiently small uncertainties the normal and lognormal shapes are very similar. However a lognormal background prior is more conservative for a highly significant excesses of observed events if the lognormal pdf is parametrized as described in section 8.3. Taking these results into account one can construct a reasonable solution for dealing with the remaining uncertainties by adding the additively contributing standard deviations in quadrature and building a lognormal pdf with $k = 1 + \sigma_{\text{combined,rel}}$. Such a procedure for handling systematic uncertainties has been proposed as a general solution for sufficiently small uncertainties in [122]. Here it is only applied if a direct combination of uncertainties with equation (A.14) is not feasible.

8.5 Dynamic Binning and Fill-up to Confidence Level

Before letting the algorithm run on a selected distribution of an event class the "binning" needs to be fixed. It is certainly not desirable to choose bin widths which are significantly narrower than the experimental resolution as this would artificially reduce $\tilde{P}$ without increasing the amount of evaluated information. If for the resulting dynamic binning the statistics of simulated events is insufficient it may also be beneficial to require a minimal bin size in order to reduce the possibility of statistically pathological situations. At the moment the dynamic rebinning is applied as explained in the following paragraph [141].

For a $\sum p_T$ distribution the bin width is adjusted according to a worst case scenario. The resolution of the variable is calculated according to the required object with the lowest p_T resolution. Parameters for the resolution of the respective physics objects have been collected from several studies within the CMS collaboration. For the MET variable a simulated MET resolution is applied. As a temporary solution it has been decided to adjust the bin size for M_{inv} on the basis of the $\sum p_T$ variable.
Especially in the high energy tails of some distributions there is a lack in statistics of the simulated Standard Model events. One possibility would be to comprise the complete tail into one bin starting form where Monte Carlo statistics begins to be insufficient. But then one might omit important features of the data distribution. A generic way to address such regions lacking

statistical information for a given background sample is to fill up the region to its confidence level. Here the confidence level corresponds to the mean of a Poisson distribution for which there is an 68% (1σ) probability of not finding an event. According to this strategy one would estimate for a given background that has no contribution for a given region that

$$\mu_b = 1.15 \cdot \text{scalefactor} \quad . \tag{8.26}$$

Unfortunately the background distribution in the affected region is usually unclear so that one needs to fill up every bin in the region which has a data entry according to (8.26) in order to assure conservatism for possible excesses in the data. If the filled up background mean is larger than the data entry the background mean is reduced to the data level to avoid artificial deficits. Sometimes the scalefactor can be so large that filling up every bin containing data will obfuscate the complete region. Thus it should be considered imperative for the MUSiC analysis to provide sufficient background statistics in as many distributions and corresponding bins as practically achievable once running on real data. For the moment the necessary fill-up procedure is implemented by the simple algorithm presented in table 8.3.

	if (bin empty $\wedge$ data entry)
1.	fill up all background processes having entries in the direction of higher energies
2.	fill up background processes from the five closest bins having entries in the direction of higher energies
	if (filled up background mean > data entry)
3.	reduce background mean to data level

Table 8.3 Simple algorithm for a provisional fill-up procedure

Tests suggest that this procedure produces reasonable results in many situations. An practical example in the context of an evaluated test scenario is given in figure 10.8 (see section 10.1). But still it is easy to construct situations in which the performance is not optimal. Thus the fill-up procedure should only be used as a makeshift where no better solution is practically achievable.

8.6 Evaluated Variables

The execution time of the full algorithm grows only linearly with the number of considered distributions. Still, it is reasonable to restrict the MUSiC analysis to a small number of evaluated variables. This is of special importance during the period of early data taking. Though the search of deviations is highly automatized, the evaluation of results is expected to involve a fair amount of educated interpretation. The overall number of evaluated distributions is equal to the number of classes times the number of variables. Even if only a few variables are considered, the resulting number will already be far from small ($\sim O(1000)$ for the current number of classes and variables). Once the corresponding results are well understood it is straightforward to extend the choice of variables while pushing the required understanding of the detector performance and the necessary quality of Standard Model simulation to new limits. The selection of variables should be based on the idea that they might be sensitive to new physics. On the other hand they should not reflect any strongly model specific assumptions. Three variables are analyzed for the test scenarios evaluated in section 10.

$\sum \mathbf{p_T}$ The scalar sum of transverse momenta for all selected objects within an event is a useful variable as its justification is merely based on the assumption that new physics might turn

up in events with high energy objects from a high momentum transfer between partons. A high momentum transfer implies that the interaction in the laboratory frame will tend to take place with a rather balanced momentum between the partons. Hence the resulting high energetic particles have momenta with higher transverse momenta than particles with comparable energy from interactions which correspond to a center of mass system that is more boosted with respect to the laboratory frame. For each event in a given class the considered sum is calculated as

$$\sum p_T = \sum_i \left| p_T^i \right| \quad , \quad i = 1 \ldots N_{\text{selected objetcs}} \quad . \tag{8.27}$$

$\mathbf{M_{inv}}$ The invariant mass is calculated via

$$M_{inv} = \sqrt{\left(\sum_i E^i\right)^2 - \left(\sum_i \vec{p}^{\,i}\right)^2} \quad , \quad i = 1 \ldots N_{\text{selected objects}} \quad . \tag{8.28}$$

M_{inv} is a Lorentz invariant quantity. With respect to the decay products of a decaying on shell particle it corresponds to its rest mass. This can be understood in an easy way by evaluating M_{inv} in the rest frame of the decaying particle. If new physics leads to phenomena which can be interpreted in terms of new decaying particles then calculating the invariant mass of their respective decay products will lead to peaks in the corresponding invariant mass spectrum. For events with missing transverse energy (**MET**) the invariant mass is not well defined as the longitudinal component of the energy is unknown. For these classes the variable M_{inv} is replaced by the transverse invariant mass M_T.

$$M_T = \sqrt{\left(\sum_i E_T^i\right)^2 - \left(\sum_i \vec{p}_T^{\,i}\right)^2} \quad , \quad i = 1 \ldots N_{\text{selected objetcs}} \tag{8.29}$$

MET The calculation of the missing transverse energy for a given event has been outlined in section 7.1.4. Because of the composite structure of the proton the high energetic parton-parton interactions as observed by the detector do usually not take place in the center of mass frame. Particles with low transverse momentum which leave the interaction point along the beam pipe are not detected. Hence the longitudinal energy of particles which interact too weakly to be measured in the CMS experiment (e.g. Standard Model neutrinos) is only poorly restricted. In contrast to this the conservation of 4-momentum implies that the transverse component will usually be balanced by a detectable energy deposition in the detector. The idea of additional contributions to the observable missing transverse energy due to particles like for example gravitons or other kinds of new weakly interacting neutral particles is present in many proposed extensions of the Standard Model. The appeal of such models is often driven by the astrophysical evidence of dark matter which may have its origin in the existence of weakly interacting massive particles. Such particles would increase the MET contributions to those event classes in which the new weakly interacting particles are produced.

8.7 A Remark on Early Data Taking

Once the analysis runs on "real" CMS data observed deviations need to be evaluated in a very careful way. Deviations might not be caused by new physics but can result from an underestimation of the involved uncertainties. The reason for such an underestimation could be a false impression of the performance of some of the many contributing inputs for the analysis. Such problems could be due to the detector response or the simulation of background events.

Additionally there might be unexpected effects in the event reconstruction or simply errors in the software. Though these possibilities remain important during all stages of the MUSiC analysis they are even more prominent during early data taking. It is very likely that at least some of the mentioned factors will result in observed deviations once running on early data. At this stage the main purpose of the MUSiC algorithm will thus be to trace back the deviations to their respective origins. In this way MUSiC can contribute to the understanding of the detector in the role of a "global physics monitor". This has been evaluated in greater detail in [2] and [134]. Once the CMS detector is well understood and after rediscovery of important features of the Standard Model the focus will shift to the aim of finding potential evidence of new physics.

It is a viable question how the MUSiC analysis can restrict the number of observed deviations that present effects of little interest during the period of early data taking. A straightforward approach would be to restrict the range of considered classes and regions. In this way one could blind the algorithm with respect to regions which are considered to be susceptible to misinterpretations. For example one could evaluate only regions with a minimum number of observed or expected events or leave out exclusive classes with high jet multiplicities. A further idea might be to improve the estimation of backgrounds by subdividing the range of the considered variables. Thus one would for example start with evaluating events up to $\sum p_T = 250\,\text{GeV}$. If the corresponding regions are found to be free from significant deviations they can be regarded as signal free regions in a subsequent analysis step evaluating the high energy tails. In this way they could be used for the employment of data-driven techniques resulting in improved background estimates.

9 Coverage Properties of the Applied p-Values

In this section several tests are presented which evaluate the coverage properties of p_N and p_{LN}. The notion of coverage has been introduced in the context of equation (8.8). The idea of the presented tests is to analyze if under controlled assumptions the p-value falsely rejects the null hypothesis as often as implied by interpreting its value as a type I error rate. In section 8.2 it has been explained that the applied p-values should prove be to close to exact or at least have overcoverage (i.e. be conservative) in most situations in order to achieve a good performance of the search algorithm.

In order to check the coverage properties of a p-value one has to fix the probabilities of all parameters that enter the definition of the p-value. This decision about the "statistical truth" defines the specific test. Additionally one has to decide which value of p is tested. This value is labeled Z_{claim} if it is expressed in terms of normal standard deviations. In principle several possibilities of calculating the coverage exist. The straightforward way to proceed is to sample hypothetical experiments according to the made assumptions. For each of those experiments the p-value is evaluated. In the limit of an infinite number of such hypothetical experiments the probability of finding a more extreme result than stated by Z_{claim} stabilizes at the value Z_{true}. In practice one dices enough experiments to ensure that statistical uncertainties on Z_{true} are small compared to the desired exactness of the results. This "dicing method" is fairly robust and can be adapted to a wide range of possible test conditions.

The observed difference $Z_{true} - Z_{claim}$ [1] is the chosen measure of coverage which is stated for a given coverage test. Unfortunately dicing a large number of input values and calculating the corresponding p-values can be quite demanding in terms of computing resources. This makes the evaluation of the coverage difficult when evaluating high significances or scanning test conditions depending on a set of input parameters. In several conceivable scenarios it is also possible to identify the coverage properties with analytical [142] or semi-analytical methods. Such an approach requires a dedicated formulation for each test and might become extremely difficult or even impossible to construct for sophisticated scenarios.

For the purpose of a model unspecific search the intermediate range of significances between $Z = 2$ and $Z = 5$ is of special importance. Additionally the presented studies focuses more on the identification of interesting deviations than on the exact quantification of significances of extreme observations. Thus most results are obtained by dicing large number of hypothetical experiments. Here the range of tests has been restricted to analyzing the coverage of an observed excess of events but the methods can be extended to an evaluation of deficits without additional conceptual difficulties. For the sake of clarity the chosen notation is summarized in table 9.1. It should be emphasized that p-values like p_N and p_{LN} are a common tool in statistical evaluations. The presented results are therefore not only of interest in the context of a model unspecific search but may be transfered to other applications. The notion of coverage for a p-value may be defined for any p-value. But reasonable tests may vary with the definitions of the p-value and the experimental context.

One possibility is to test a scenario in which the truncated normal prior used for p_N truthfully describes the probability of assuming background means μ. In this case μ_b is both the mean of the prior and the true mean of the Poisson distribution. Additionally, the uncertainty one would state under the assumption of μ needs to be defined. In the first test it is assumed that for each

[1] $Z_{true} - Z_{claim} > 0$ means overcoverage (conservative),
$Z_{true} - Z_{claim} < 0$ means undercoverage (liberal), see section 8.1

Z	p-value transformed into normal standard deviations (one sided) according to equation (8.2)
Z_N, Z_{LN}	p-values p_N and p_{LN} (see section 8.3) transformed into normal standard deviations (one sided) according to equation (8.2)
Z_{claim}	level of significance that is tested in terms of a Type I error rate
Z_{true}	fraction of hypothetical experiments in which $Z < Z_{claim}$ is observed transformed via formula (8.2)
n_{on}	number of counted events in a hypothetical experiment
μ_{true}	the mean of the Poisson distribution realized in nature
μ_{diced}	the assumed mean of the Poisson distribution in a given hypothetical experiment
k , b_0	parameters modeling the uncertainty of the background prediction for p_{LN} as defined in equation (8.17)
σ_b , μ_b	parameters modeling the uncertainty of the background prediction for p_N as defined in equation (8.17)

Table 9.1 Listing of notation used in the context of the presented coverage tests

μ an uncertainty of $f \cdot \mu_b$ is used for the calculation of p_N. Thus the parameters f and μ_b need to be defined to fix the parameters of this kind of coverage test. The algorithm evaluating the coverage for given parameters f and μ_b is outlined in table 9.2.

Results for $Z_{claim} = 3$ scanning up to a relative uncertainty of 15% are presented in figure 9.1. Corresponding results for $Z_{claim} = 2$ can be found in figure B.6 in the appendix. One can see that for this type of scenario $Z_{true} - Z_{claim}$ is close to zero (meaning close to exact) for a wide range of possible parameters. The behavior for other tested significances Z_{claim} is similar (tests for single sets of parameters have been performed up to about $Z_{claim} = 5$). Overcoverage is only observed for parameters with small background means $\mu_{b,true}$. In this range also a slightly oscillatory behavior of the coverage is found. This effect is stronger in the region where the uncertainties are very low. Both of these trends become more prominent when evaluating background means smaller than 1 as shown in figure 9.2. The reason for these effects is connected to the properties

1. set value of Z_{claim}
2. fix the background μ_{true} and set $\sigma_b = f \cdot \mu_{true}$

for a sufficient number of times {
3.1 dice the observed number of events n_{on} from a Poisson distribution with mean μ_{true}
3.2 dice the assumed background μ_{diced} for this pseudo-experiment from a truncated normal distribution $N_{trunc}(x, \mu_{true}, \sigma_b)$
3.3 if $\mu_{diced} > n_{on}$ set Z_N to a value smaller than Z_{claim} (only testing for observed excess)
3.4 if $\mu_{diced} < n_{on}$ evaluate $Z_N\,(n_{on}, \mu_{diced}, \sigma_b)$
}

4. calculate Z_{true} from the fraction of hypothetical experiments with $Z_N > Z_{claim}$
5. report coverage by stating $Z_{true} - Z_{claim}$

Table 9.2 Algorithm for testing p_N with fixed absolute uncertainty

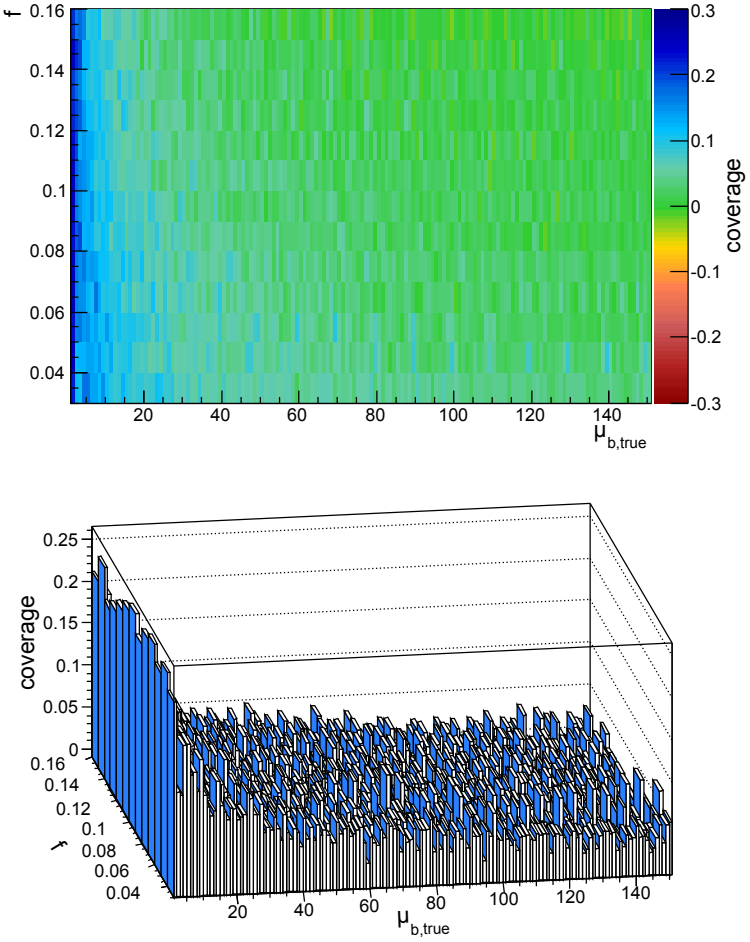

Figure 9.1 Coverage of p_N probing $Z_{claim} = 3$ for the scenario outlined in table 9.2 with $\sigma_b = f \cdot \mu_{true}$

of the Poisson pdf.

As only natural event numbers can be observed, the distribution needs to be discrete. If uncertainties and background mean are small the relevant part of the prior for the background is centered on an interval of backgrounds much smaller than 1. Then the coverage properties are dominated by the Poisson probabilities of observing zero or only few events. Consequently, reducing the true background just by a small amount can cause one of those discrete probabilities (let us for example say $n_{on} = 2$) to contribute to $Z > Z_{claim}$. This can have a large effect on the coverage. This kind of behavior does not strongly depend on the chosen background prior and occurs also for scenarios in the context of p_{LN}.

If one believes that the presented scenario is a good description of the encountered uncertainties one may conclude that p_N has good coverage properties. Unfortunately this is hard to decide in practice.

For example, one can consider a situation in which the assumed relative uncertainties are independent from the background mean that is given in an hypothetical experiment. In this alternative scenario one would assign a relative uncertainty of x % to one's method of estimating μ_b independent of the outcome of the actual measurement.

In this case it is found that one may encounter severe undercoverage if the uncertainties are not negligible. A evaluation of the corresponding findings for $Z_{claim} = 3$ is shown in figure 9.3. The necessary changes that need to be applied to the algorithm are outlined in table 9.3. In principle one could try to think of other scenarios in which one could model different reasonable functions

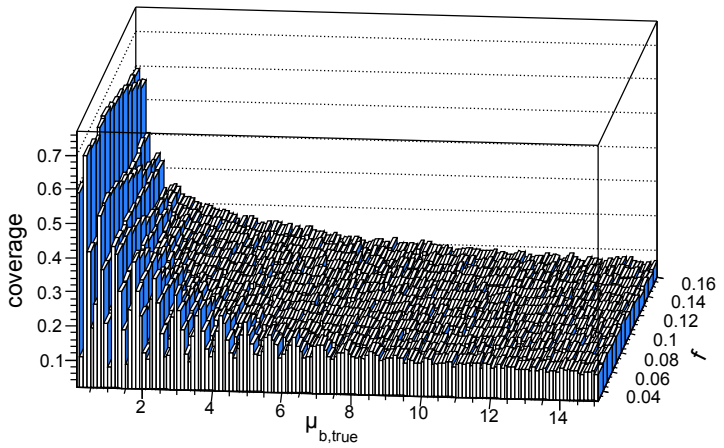

Figure 9.2 Coverage of p_N probing $Z_{claim} = 3$ for scenario 9.2 in the case of small background means

3.4 (modified)	if $\mu_{diced} < n_{on}$ evaluate $Z_N\,(n_{on}, \mu_{diced}, \sigma_{b,rel})$, where $\sigma_{b,rel} = f \cdot \mu_{diced}$

Table 9.3 Changes to the algorithm outlined in 9.2 for testing p_N with fixed relative uncertainty

describing the probability for a given σ_b for a hypothetical experiment. For such a generalization the presented test would then correspond to the special cases $Pr\,(\sigma_b) = \delta\,(\sigma_b - f \cdot \mu_b)$ and $Pr\,(\sigma_b) = \delta\,(\sigma_b - f \cdot \mu_{diced})$.

The results depicted in figures 9.1, B.6, 9.3 and B.7 can be compared to results presented in [117] and are found to be in good agreement. In [117] the corresponding results have been derived in a somewhat different way which is more test specific, but applicable to a wider range in Z_{claim}. Such an alternative procedure for the test outlined in table 9.2 is summarized in table 9.4. Comparing the results presented in figures 9.1 and B.6 with this algorithm provides a good test for a working implementation of the "dicing method". In order to allow for some flexibility for a possible expansion to alternative priors the TMinuit minimization package [143] is used to find μ_{lim} (see table 9.4 for a definition of μ_{lim}). Though in principle using TMinuit extends the applicability to other p-values the starting parameters and the range must be carefully chosen to achieve correct results. Here this approach was mainly intended as a consistency check. Results

1.-2.	as in table 9.2
	loop over (possibly relevant) values of n_{on} *{*
3.1	find value μ_{lim} (if possible by analytical methods) for which $Z\,(n_{on}, \mu_{lim}, \sigma_b) = Z_{claim}$
3.2	integrate probability for $\mu < \mu_{lim}$ (implying $Z_N > Z_{claim}$)
	}
4.	sum up $Pr\,(Z > Z_{claim} \mid n_{on})$ for all n_{on} and calculate Z_{true}
5.	as in table 9.2

Table 9.4 Alternative method for testing p_N with fixed absolute uncertainty

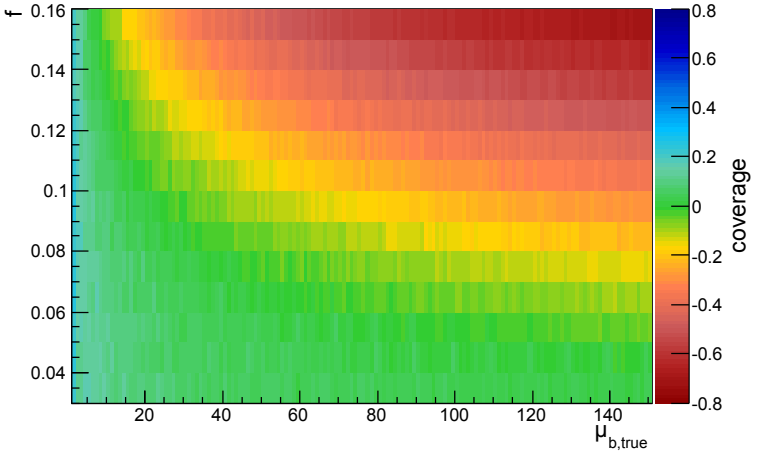

Figure 9.3 Coverage of p_N probing $Z_{claim} = 3$ for a scenario with $\sigma_b = f \cdot \mu_{diced}$

for a range of parameters are shown in figure 9.4 for $Z_{claim} = 2$. As expected no significant differences between the methods can be observed.

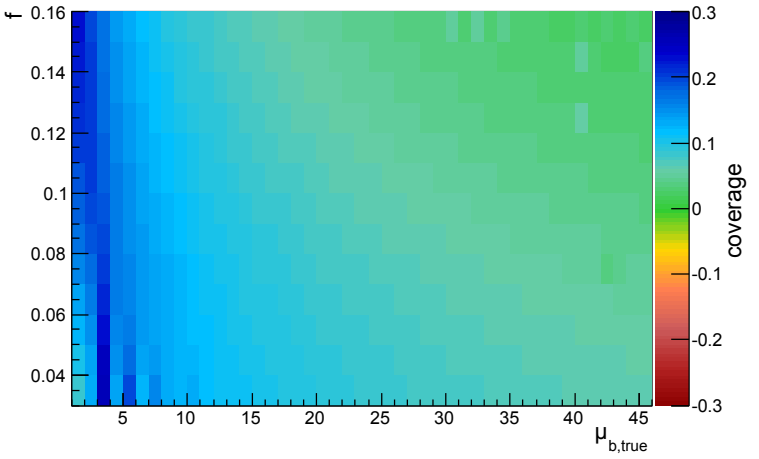

Figure 9.4 Coverage of p_N probing $Z_{claim} = 2$ for the alternative method summarized in table 9.4

It is possible to investigate the coverage for p_{LN} in similar scenarios. For evaluating p_{LN} according to equation (8.17) the parameters b_0 and k for the lognormal distribution need to be chosen for each hypothetical experiment. For a first test it is assumed that the probability distribution of b_0 is indeed given by a lognormal distribution with $k = 1+f$ and median $b_{true} = \mu_{true}$. μ_{true} is the mean of the realized Poisson distribution from which n_{on} is drawn. Additionally it is chosen to set $k = 1 + f$ for each hypothetical experiment. To get a better understanding of this situation one has to recall that in section 8.3 it has been decided to fix $k = 1 +$ relative uncertainty. Therefore this test resembles the situation of testing p_N for a fixed relative error as presented for the example $Z_{claim} = 3$ in figure 9.3. The corresponding algorithm for p_{LN} is summarized in table 9.5 and results for a range of parameters with $Z_{claim} = 3$ are shown in figure 9.5. For comparison the respective results for $Z_{claim} = 2$ are depicted in figure B.8. Comparing figure 9.5 with figure 9.3 one finds that p_{LN} has better coverage properties for a scenario with fixed relative uncertainties. Thus if both a normal and a lognormal parametrization of the background prior

1. define Z_{claim}

2. fix the background μ_{true} and set $k = 1 + f$

for a sufficient number of times {

3.1 dice observed number of events n_{on} from a Poisson distribution with mean μ_{true}

3.2 dice the assumed background $b_{0,diced}$ for this pseudo-experiment from a lognormal distribution $\Lambda_(x, b_0 = \mu_{true}, k)$

3.3 if $b_{0,diced} > n_{on}$ set Z_{LN} to a value smaller than Z_{claim} (only testing for observed excess)

3.4 if $b_{0,diced} < n_{on}$ evaluate $Z_{LN}\,(n_{on}, b_{0,diced}, k)$

}

4. calculate Z_{true} from the fraction of hypothetical experiments with $Z_{LN} > Z_{claim}$

5. report coverage by stating $Z_{true} - Z_{claim}$

Table 9.5 Algorithm for testing p_{LN} with fixed relative uncertainty

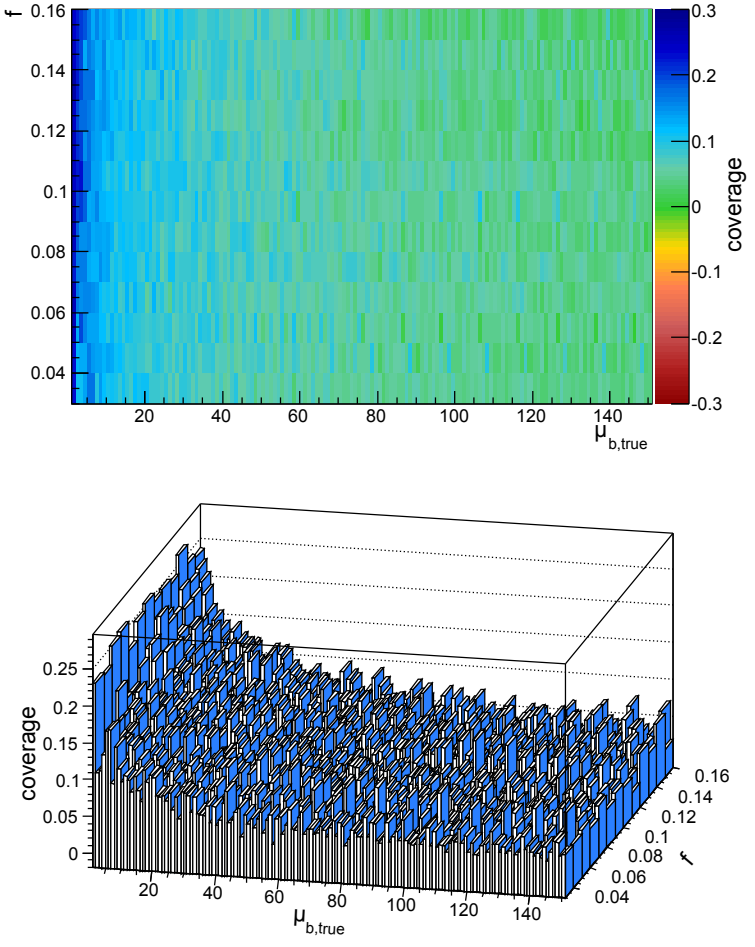

Figure 9.5 Coverage of p_{LN} testing $Z_{claim} = 3$ for the scenario outlined in table 9.5 with $k = 1 + f$

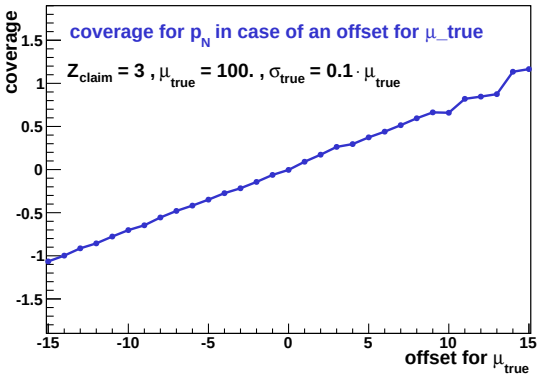

Figure 9.6 Coverage of p_N probing $Z_{claim} = 3$ for a special parameter point in dependence of an offset for the mean μ_{true} according to which μ_b is diced. coverage > 0 signals overcoverage.

seems to be reasonable, using p_{LN} should be the preferred choice if fixed relative uncertainties are found to present a convincing description of additional hypothetical experiments.

Being critical one might argue that scenarios in which the distribution from which the parameters like μ_b are sampled match exactly the assumption of the p-value might still be too much in favor of a good performance in terms of coverage. Hence it makes sense to investigate the effects of loosening the correspondence between the background prior and the test scenario. Straightforward possibilities include an offset for the median, a differing width, or a change in the shape of the distributions according to which the pseudo experiments are sampled.

Figure 9.6 shows an example in which the coverage is evaluated for fixed values of σ_b and μ_{true} as a function of an offset Δ_μ between μ_{true} and the median of the normal distribution from which μ_{diced} is sampled. The result is a test similar to what is outlined in table 9.2. The differences are outlined in table B.1. As expected the coverage deviates from zero for an increasing offset. In the evaluated range the deviation is approximately linear. The graph becomes less smooth for $\Delta_\mu > 10$. In this region a result with $Z > Z_{claim}$ is sufficiently improbable to cause statistical fluctuations that are not negligible with respect to the chosen number of pseudo experiments.

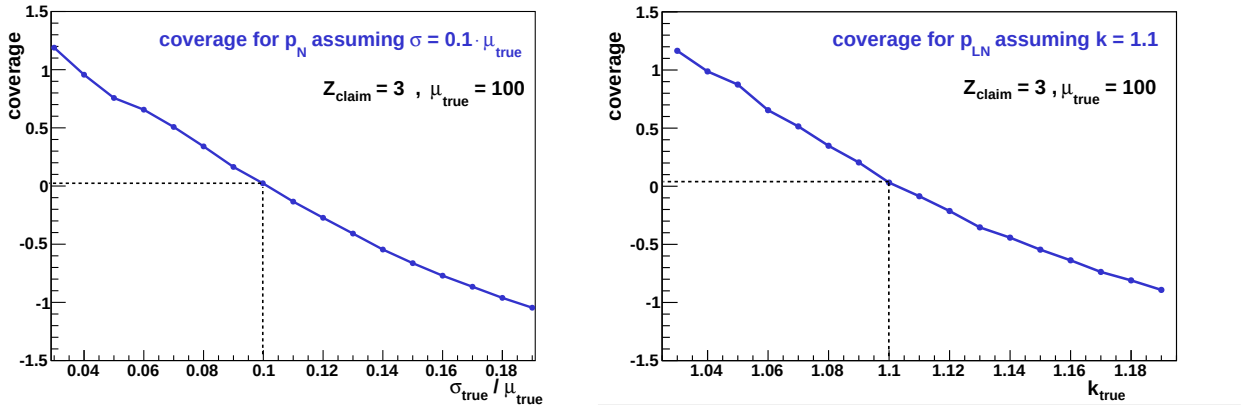

Figure 9.7 Left: coverage of p_N probing $Z_{claim} = 3$ for a special parameter point in dependence of a varying value of the assumed parameter σ_{true} . **Right:** coverage of p_N probing $Z_{claim} = 3$ for a special parameter point in dependence of a varying value of the assumed parameter k_{true} . coverage > 0 signals overcoverage.

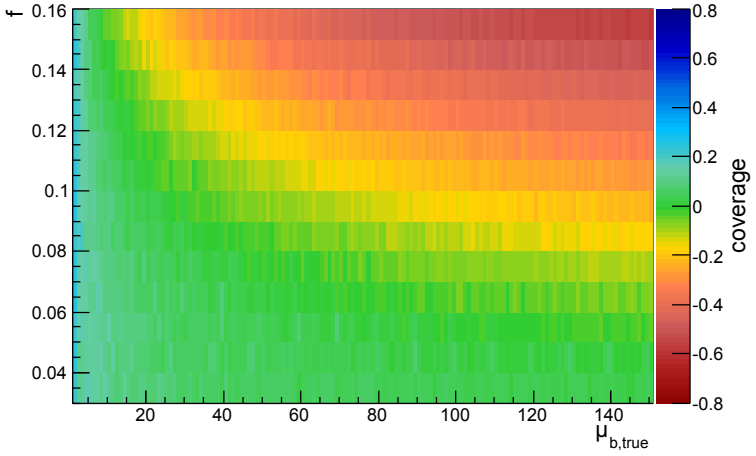

Figure 9.8 Coverage of p_{LN} testing $Z_{claim} = 3$ for the scenario outlined in table B.4

One can also evaluate the varying coverage if there is a difference between the assumed uncertainty on the background mean and the width of the distribution according to which the pseudo-experiments are sampled. The results make an assessment of the significance of an observation if one does not agree with the uncertainty that has been assumed. In figure 9.7 such a scenario is evaluated for a fixed set of parameters while varying k_{true} for p_{LN} and $\sigma_{b,true}$ for p_N. The uncertainties k and σ_b are kept at fixed values with $\Delta_\sigma = \sigma_b - \sigma_{b,true}$ and $\Delta_k = k - k_{true}$. The corresponding changes to the previously prescribed procedures are outlined in tables B.2 and B.3.

There is undercoverage if the uncertainty is underestimated and overcoverage if a larger uncertainty than implied by the dicing procedure is assumed. The shift in coverage in the evaluated region is found to be roughly linear for the chosen parameters.

It is also possible to study the influence on the coverage if the pseudo experiments are diced according to a pdf that is different from the distribution assumed in the evaluation of the p-value. Figure 9.8 presents a scenario in which the test for p_N assuming a fixed relative error (see table 9.3) has been modified towards an evaluation with p_{LN}. Still the background means

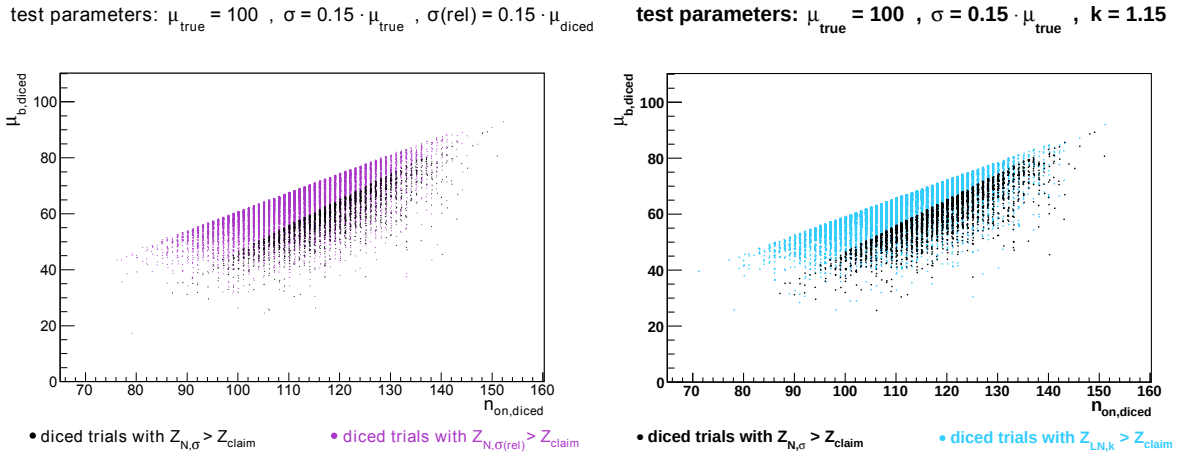

Figure 9.9 Left: comparison between diced parameter points resulting in $Z > Z_{claim}$ for testing p_N in the scenarios with absolute and relative uncertainties. μ_{true} is fixed. **Right:** The same, but comparing p_N in the scenario with absolute uncertainty with the results of employing p_{LN} for the same dicing procedure.

are sampled from a normal distribution but the significance is evaluated with p_{LN} for which it is assumed that $k = 1 + f$ and $b_0 = \mu_{diced}$ (as argued in section 8.3 this is a reasonable choice). The respective modifications to the algorithm are presented in table B.4. The results for this test for $Z_{claim} = 3$ are presented in figure 9.8. They can be compared to what is presented in figure 9.3 as both scenarios assume a fixed relative error for the pseudo experiments. Both examples show undercoverage though the performance in the scenario with p_{LN} is slightly better. The underlying reason for the observed undercoverage is explained in figure 9.9. Here for fixed values of f and μ_{true} both scenarios are compared with the test for p_N assuming a fixed absolute error. The latter shows almost perfect coverage for the chosen parameters. Pseudo-experiments resulting in $Z > Z_{claim}$ are shown in dependence of n_{on} and μ_{diced}. In both cases comparatively lower values of μ_{diced} already result in $Z > Z_{claim}$ for each given value of n_{on}. This causes the observed undercoverage.

It has been demonstrated that the coverage properties of p-values show a rather complex behavior in the different test scenarios. Both p_{LN} and p_N have good coverage properties in some of the evaluated examples. Still one should be aware that there are potential problems which might lead to misinterpretions of the significance of observed deviations in unfavorable cases. Therefore it makes sense to have more than one p-value available for crosschecks. As the coverage properties for a given test scenario clearly depend on the chosen background prior of the p-value, the presented coverage tests provide additional arguments for studying alternatives to normal background priors.

10 Test Scenarios with High p_T Photons

In this section several test signatures are used to evaluate the functionality and the operational liability of the MUSiC analysis framework for signatures with high p_T photons. Showing that the envisioned analysis is sensitive to some of such models gives confidence in the implementation of the workflow and strengthens the credibility of the claim that a model unspecific approach might be able to find hints of new physics. Some of the incorporated signal events are taken from the official CMS signal production. Other samples have been kindly provided by CMS members working on a dedicated analysis or have been produced in cooperation with others.

In an analysis with "real" data the value p_{min} is determined for each distribution. Then p_{min} is used to determine the significance $\widetilde{P}$ in order to account for the look-elsewhere-effect. As explained in section 8.2 this is achieved by dicing a large number of hypothetical background only (B) experiments. If one is testing the algorithm with the help of a signal sample that has been generated with respect to a specific model it is also possible to dice a number of signal plus background experiments ($S + B$). In this case the systematic and statistical uncertainties on the signal sample are also taken into account. This allows to estimate the expected p_{min} distribution for a particular signal plus background scenario. For such tests one defines $\widetilde{P}$ to be the fraction of background only experiments which results in higher significances than given by the median of the signal plus background distribution. This is outlined in equation (10.1).

$$\widetilde{P} = \frac{\text{number of } (B) \text{ experiments wit } p_{min(B)} < p_{min,median(S+B)}}{\text{total number of } (B) \text{ experiments}} \tag{10.1}$$

$\widetilde{P}$ is to be interpreted as the expected significance. For example $\widetilde{P} = 4\,\sigma$ would state that for a particular test model and event class one would expect to find a significance of more than $4\,\sigma$ with 50 % probability. Of course, if the test scenario is found to be highly significant, it will not always be possible or worth the necessary effort to dice enough pseudo experiments for an excact determination of $\widetilde{P}$. In these cases only a lower limit on $\widetilde{P}$ is given by stating the number of (B) experiments that have been diced without finding any case with $p_{min(B)} < p_{min,median(S+B)}$. In principle it is clear that for all the mentioned scenarios there is a statistical uncertainty on $\widetilde{P}$. In order to keep things simple this uncertainty is not explicitely covered, because it is not relevant for the interpration of the evaluated test scenarios. Nevertheless enough pseudo experiments for receiving statistically stable values of $\widetilde{P}$ are diced where this is achievable with reasonable effort.

The feature of dynamic binning is applied in all presented scenarios (see section 8.5). Thus care needs to be taken in order to correctly interpret the bin content in histograms which show the signal and background expectations for each bin (e.g. figure 10.5). To arrive at a correct number of expected events in a given bin one needs to multiply the bin entry with the fraction of bin width and reference width.

$$N_{\text{expected},i} = N_{\text{entry},i} \cdot \frac{\text{width}_i}{\text{width}_{\text{ref}}} \tag{10.2}$$

Though this makes calculating expected event numbers a bit more complicated it preserves the natural shape of the expected distribution while correctly depicting the chosen bin widths.

All evaluated test scenarios assume a center of mass energy of $\sqrt{s} = 10\,\text{TeV}$. Depending on the specific test signal the assumed luminosity is either $1\,\text{fb}^{-1}$ or $100\,\text{pb}^{-1}$. The cross sections of the samples are taken at leading order and have not been corrected for higher order k-factors.

The following scheme is used to label the event classes (see section 6.6) in an unambiguous way: For exclusive event classes the multiplicity of each required object is indicated (e.g. 2e1γ3jet). For inclusive event classes the label $+X$ is appended to the minimal content of objects (e.g. $1\mu 1\gamma + X$).

10.1 Supersymmetric Scenarios with Gauge Mediated Supersymmetry Breaking (GMSB)

Supersymmetric models including the mechanism of Gauge Mediated Supersymmetry Breaking (see section 3.1) present one possibility of solving issues like the hierarchy problem with the help of a supersymmetric approach. In the considered model the lightest neutralino $\widetilde{\chi}_1^0$ is the next to lightest supersymmetric particle (NLSP). The GMSB parameters are fixed according to what has been proposed as the SPS 8 benchmark scenario in [25], leaving Λ as a free parameter. A promising way to detect potential GMSB signals is to look for the decay of the NLSP. The requirement of R-parity imposes the pairwise creation of supersymmetric particles. If the NLSP decay is a probable process within each decay chain, then it is expected to be observed in pairs. For the chosen model the predominant decay channel is $\widetilde{\chi}_1^0 \rightarrow \widetilde{G}\,\gamma$. As a consequence of R-parity conservation the light gravitino $\widetilde{G}$ is stable. The gravitino $\widetilde{G}$ escapes the detector without direct detection. Hence the expected signature of the two NLSP decays is $2\gamma + MET$. The Feynman graph representing the $\widetilde{\chi}_1^0 \rightarrow \widetilde{G}\,\gamma$ decay is shown in figure 10.1.

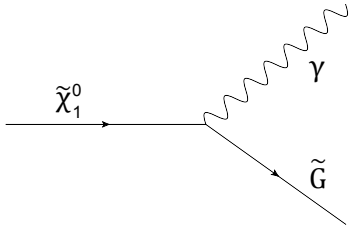

Figure 10.1 Feynman graph depicting the dominant signature of the NLSP decay $\widetilde{\chi}_1^0 \rightarrow \widetilde{G}\,\gamma$ in the evaluated GMSB scenarios

The parameters of the selected benchmark points GM1d, GM1e and GM1f are presented in table 10.1. These benchmark points are not yet experimentally excluded [144, most recent CDF analysis at $2.6\,\mathrm{fb}^{-1}$]. The branching ratios $\widetilde{\chi}_1^0 \rightarrow \widetilde{G}\,\gamma$ vary between 91 % (GM1d) and 85 % (GM1f) [26]. Other decay possibilities include $\widetilde{\chi}_1^0 \rightarrow \widetilde{G}\,Z^0$. The following results assume a $1\,\mathrm{fb}^{-1}$ scenario at 10 TeV. At this stage of data taking one may certainly assume a reasonable understanding of the behavior of the missing transverse energy, including a good performance of the basic corrections outlined in section 7.1.4.

Figure 10.2 depicts the global trigger efficiency for the GM1d signal sample. The events feature high energetic objects which often include two photons. Thus it is not surprising that the chosen MUSiC trigger menu (see table B.7) which incorporates several dedicated photon triggers is found to be fully efficient with respect to the evaluated GMSB model. Details concerning the event and object selection are discussed in chapter 7.1.

Point	Λ	M	$\tan\beta$	N_M	sgn(μ)	C_{grav}	$M(\tilde{\chi}_0^1)$	σ_{LO}[fb]
GM1d	120 TeV	2Λ	15	1	+	1	168 GeV	299
GM1e	140 TeV	2Λ	15	1	+	1	197 GeV	124
GM1f	160 TeV	2Λ	15	1	+	1	226 GeV	58

Table 10.1 Selected GMSB benchmark points

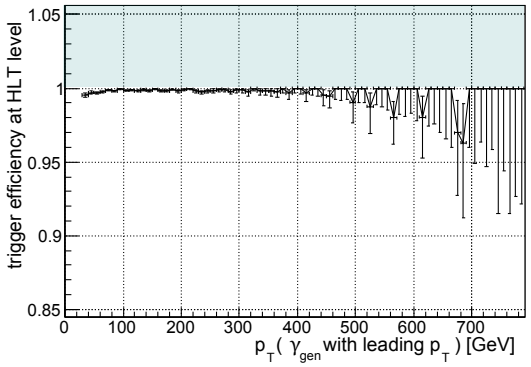

Figure 10.2 Global trigger efficiency for the GM1d signal sample

Because of the potentially complex SUSY decay chains it is natural to assume that the GMSB signal events might feature high jet multiplicities. For GM1d this is checked along with several other variables in figure 10.3. Requiring $p_{T,\text{jet}} > 60\,\text{GeV}$ the found jet multiplicities cover a wide range up to values of about 10. The photon multiplicity of the generated events shows that events with one or two photons are predominant. The multiplicity of the generated photons is depicted after applying a relative isolation cut of 10 % at generator level and cuts on η and p_T to allow for a reasonable comparison with the multiplicity at reconstruction level. Consequently the multiplicity at generator level is lowered with respect to what would be expected from the pure $\widetilde{\chi}_1^0 \rightarrow \widetilde{G}\,\gamma$ branching ratio. As one would expect the average photon multiplicity is reduced after reconstruction. The remaining two plots show the photon p_T spectrum of the photons leading and second in p_T. The p_T distribution of the leading photon is found to be peaking close to 100 GeV.

An event visualization of a GM1d event at an early reconstruction stage with two photons, missing transverse energy and a high multiplicity of additional jets is presented in figure 10.4. The figure is a slightly modified version of what is depicted by the Fireworks event display [145].

Expected to be among the most significant distributions in the considered GMSB scenarios, the $\sum p_T$ distribution of the $2\gamma 1\text{MET} + X$ class is evaluated in some detail. Figure 10.5 depicts the results of scanning the class $2\gamma 1\text{MET} + X$. The upper two plots show the outcome of two $(S+B)$ pseudo experiments providing examples for the application of p_N and p_{LN}. The expected event numbers for the signal and the background contributions are indicated for each bin. They correspond to the evaluation of all included simulated event samples. The GM1d signal is found to be peaking at $\sum p_T \approx 400\,GeV$. It is dominating the background expectation by $1-2$ orders of magnitude over a wide range of the $\sum p_T$ distribution. The determined regions of interest[1] in the pseudo experiments presented in figure 10.5 correspond to the respective range which is delimited with dashed lines. The algorithm has picked a reasonable region in both cases. Extending the region of interest to higher values of $\sum p_T$ would not make sense, since this would increase the background without including additional observed events. Consequently, such a region would be less significant and can therefore not be the region of interest. It is also reasonable that the extension of the region of interest to lower values of $\sum p_T$, where the backgrounds are increasing, will be limited to a range featuring bins in which the number of observed events is sufficiently higher than the respective background expectations.

The main observed backgrounds are due to contributions from the electroweak W+Jets and Z+Jets production and QCD. That these are indeed the most important background processes is supported by results presented in the context of a dedicated GMSB analysis [26]. For the chosen set of selection criteria QCD multijet events seem to be somewhat more important

[1] the region which features the most significant deviation (see section 8)

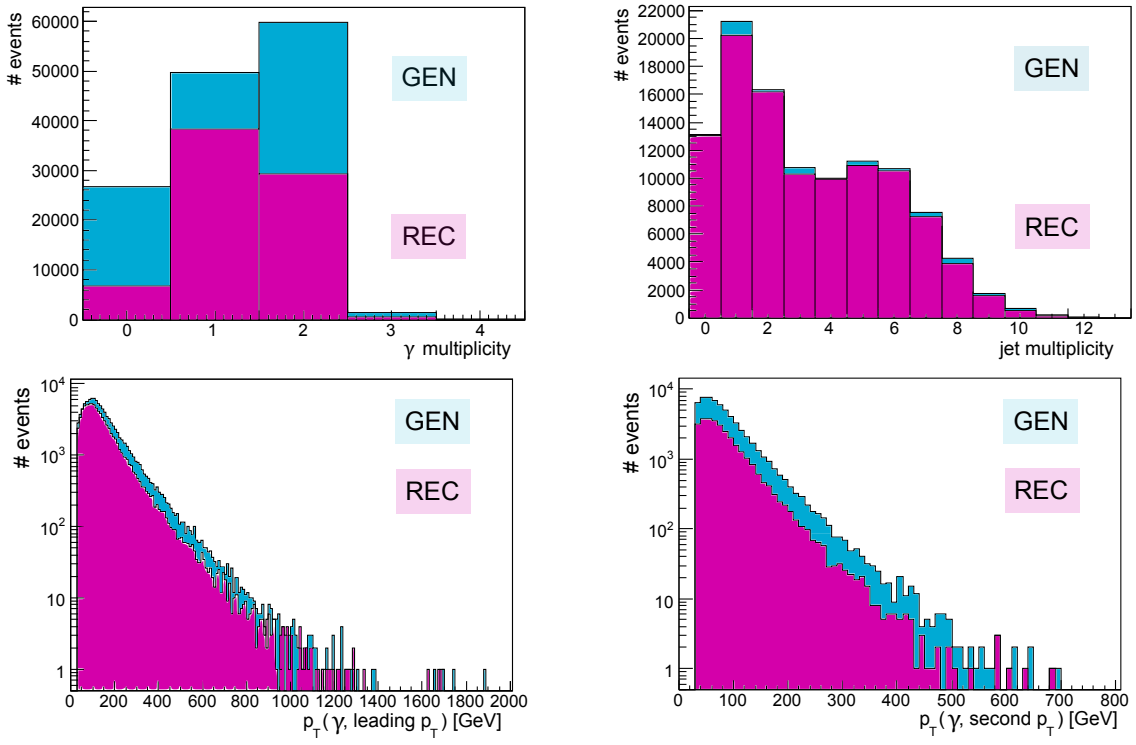

Figure 10.3 **Up left:** photon multiplicity at generator level ($p_T > 30\,\text{GeV}\,, \eta < 2.5\,,$ isolation) (GEN) and after reconstruction and final selection (REC). **Up right:** jet multiplicities for (GEN) ($p_T > 60\,\text{GeV}\,, \eta < 2.5$) and (REC) (final selection). **Down left:** p_T distribution of the photon with highest p_T for (GEN) and (REC). **Down right:** p_T distribution of the photon with next to highest p_T for (GEN) and (REC), all plots referring to **GM1d**

than QCD direct photon production. Though on average reconstructed QCD multijet events in the $2\gamma 1\text{MET} + X$ class need to contain more fake photons this is a reasonable result since the respective cross section in the signal region surpasses the cross section of direct photon production by several orders of magnitude (see table C.2). Direct photon production includes single photon production through quark-gluon Compton scattering $q + g \rightarrow \gamma + q$ and quark-antiquark annihilation $q+\bar{q} \rightarrow \gamma+g$ (see e.g. [146]) and diphoton production which is dominated by the Born diagram $q + \bar{q} \rightarrow 2\gamma$ of quark-antiquark annihilation (see [147]).

Additionally, electroweak W and Z production can cause background contributions. There can be photons from initial and final state radiation or direct photon production like $q\bar{q} \rightarrow W+\gamma$. Final state photons can also stem from by misidentified jets or electrons. Both W and Z events can show missing transverse energy from decay channels with neutrinos in the final state. Of course additional MET contributions can be caused by mismeasurements.

The evaluated uncertainties in the presented $\sum p_T$ distribution have a considerable influence. For the application of p_N this is indicated by the shaded 1σ error band in figure 10.5 (upper left plot). Excluding the statistical Poisson uncertainty the error band summarizes all other uncertainties that are applied in the dicing procedure when evaluating results according to p_N. Consequently one should expect a perceivable difference between the the use of p_N and p_{LN} as discussed in section 8.3.

Even if the systematic uncertainties are taken into account it is still rather obvious that one would expect a high significance in this class. This impression is supported by the values of p_{min} obtained for in the presented examples.

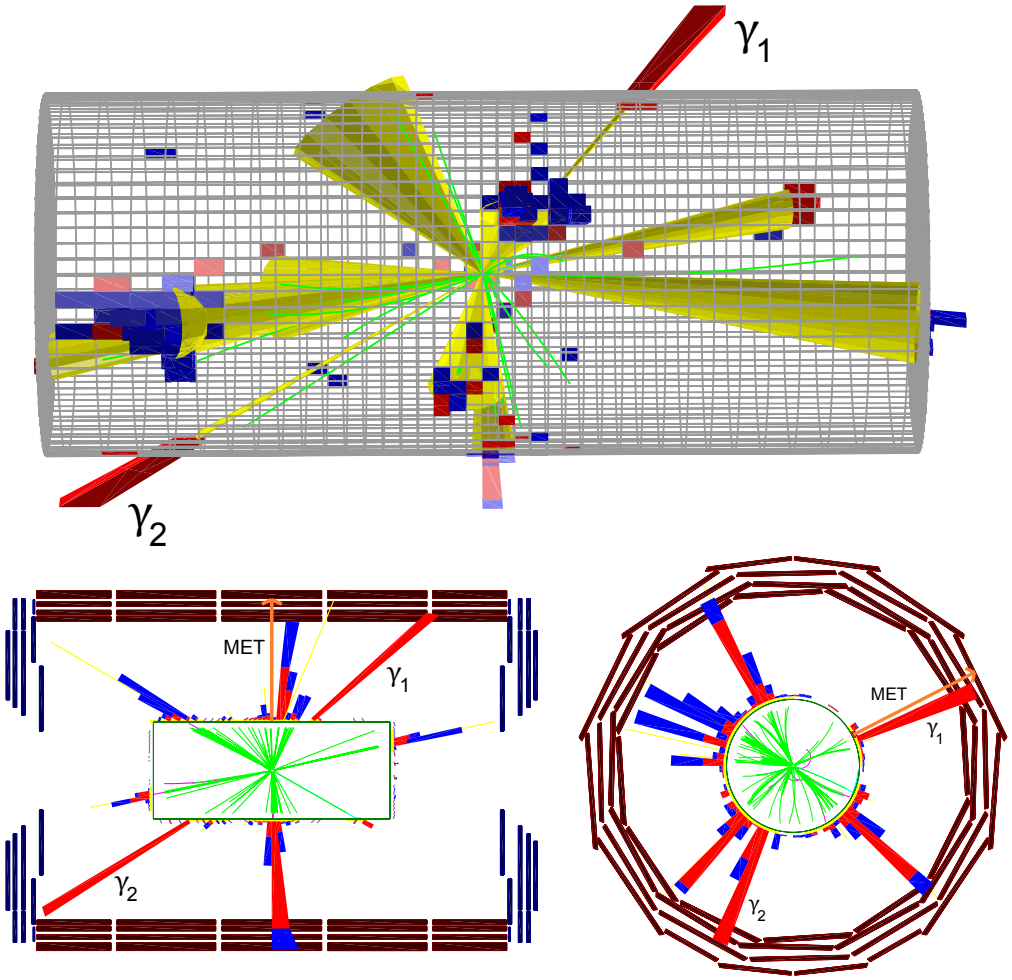

Figure 10.4 CMS Event display of a typical GM1d event with two photons and missing transverse energy generated with Fireworks (modified colors and added labels). Deposits in the ECAL are indicated with red and deposits in the HCAL with blue. The yellow cones correspond to L1 jets. The missing transverse energy (MET $\approx 280\,\text{GeV}$) and the two photons ($\gamma_1 \approx 245\,\text{GeV}$, $\gamma_2 \approx 220\,\text{GeV}$) have been labeled.

The other two histograms in figure 10.5 show the p_{min} distributions for both dicing scenarios. As the background only (B) experiments are always diced according to p_N (see section 8.4) it is not surprising that the corresponding distributions are in full agreement. More interesting are potential differences in the signal plus background distributions. It can be noticed that p_{LN} tends towards higher values of p_{min} and is therefore more conservative. This is in agreement with the explanations in chapter 8.1. The tail of the distribution for low p -values that is observed in the $(S+B)$ distribution of the p_{LN} scenario can be explained by the remaining difficulties which are related to the lognormal dicing procedure. Those have been outlined in section 8.4. In the case of non negligible uncertainties these will lead to a higher probability of dicing large deviations than estimated by the p-value with lognormal prior. However, it should be emphasized that for actual data there is neither need nor possibility to dice $(S+B)$ experiments and thus a consistent implementation of dicing $(S+B)$ experiments is not of crucial importance. In addition to this it is rather unlikely that this problem will have any noteworthy effect on the median of the distribution. With respect to the evaluation of a single event class the median is the only aspect of the $(S+B)$ p_{min} distribution which is used for further investigations. For both p-values $1.5 \cdot 10^4$ signal plus background and one million background only experiments have

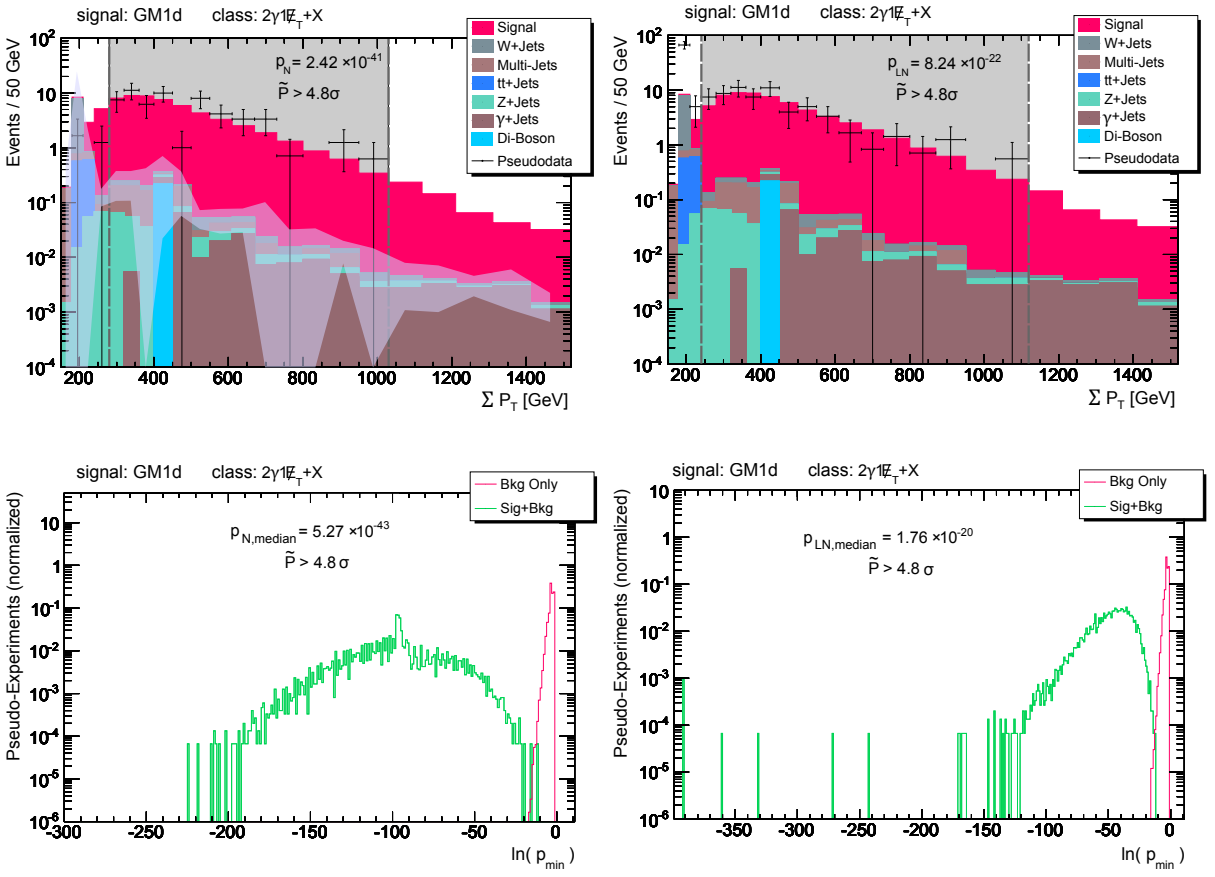

Figure 10.5 Evaluation of the inclusive class with at least two photons and MET for the GM1d test scenario with $L = 1\,\mathrm{fb}^{-1}$ at $\sqrt{s} = 10\,\mathrm{TeV}$. The delimited areas (dashed lines) shaded in grey label the region of interest for a specific pseudo experiment. **Up left:** $(S + B)$ pseudo experiment using normal dicing, (the error band illustrates the uncertainty on the background expectation). **Up right:** $(S + B)$ pseudo experiment using lognormal dicing. **Down left:** p_{min} distribution for (B) (normal dicing) and $(S + B)$ (normal dicing). **Down right:** p_{min} distribution for (B) (normal dicing) and $(S + B)$ (lognormal dicing)

been diced. None of the background only experiments results in $p_{\mathrm{min}}(B) < p_{\mathrm{min}}^{\mathrm{median}}$. Hence it can be conjectured that $\widetilde{P} > 4.8\sigma$ for both p-values.

Figure 10.6 presents the corresponding results of evaluating the class $2\gamma 1\mathrm{MET}+X$ using GM1f as signal instead of GM1d. For the GM1f scenario an interesting difference is found between the two alternative versions of the algorithm. While using p_N to estimate the significance of the $(S + B)$ pseudo experiments leads to $\widetilde{P} \approx 5\,\sigma$ [2] the application of p_{LN} results in a more conservative claim of $\widetilde{P} \approx 4\,\sigma$. The results show that the GM1f benchmark point is close to the threshold in the GMSB parameter space for which one could expect significant deviations in the presented model unspecific approach at $1\,\mathrm{fb}^{-1}$.

The exclusive class $1e2\mathrm{jet}1\mathrm{MET}$ provides a final state without any significant signal contribution from the evaluated GMSB samples. An exemplary pseudo $1e2\mathrm{jet}1\mathrm{MET}$ experiment including the GM1d signal sample is presented in figure 10.7.

The class is correctly labeled as insignificant by both p-values. It can be argued that this specific test provides an example of "successfully finding nothing". This is not trivial as it

[2] That $\widetilde{P} \approx 5\,\sigma$ is clear, because due to the look-elsewhere-effect (section 8.2) $\widetilde{P}$ is expected to be lower than the average of p_{min}.

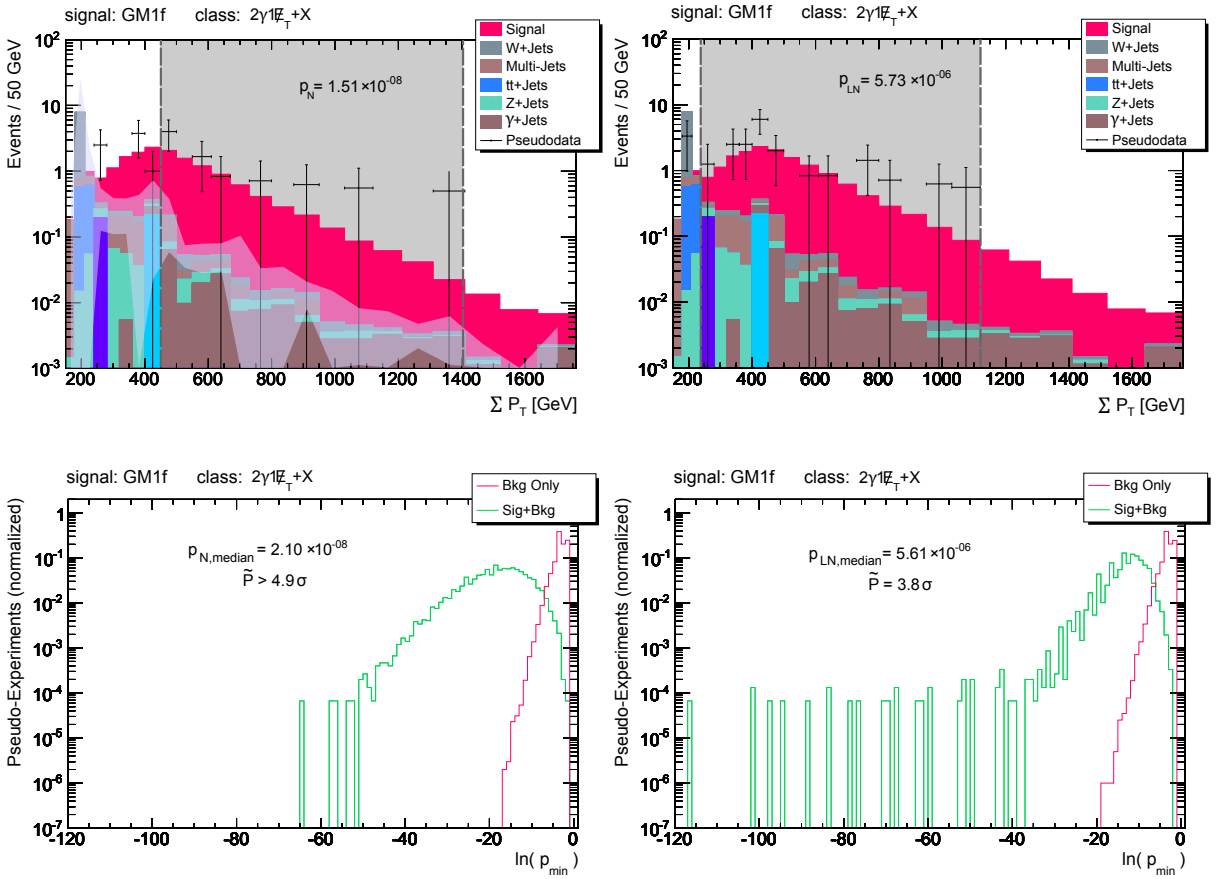

Figure 10.6 Evaluation of the inclusive class with at least two photons and MET for the GM1f test scenario with $L = 1\,\mathrm{fb}^{-1}$ at $\sqrt{s} = 10\,\mathrm{TeV}$. The delimited areas (dashed lines) shaded in grey label the region of interest for a specific pseudo experiment. **Up left:** $(S + B)$ pseudo experiment using normal dicing, (the error band illustrates the uncertainty on the background expectation). **Up right:** $(S + B)$ pseudo experiment using lognormal dicing. **Down left:** p_{min} distribution for (B) (normal dicing) and $(S + B)$ (normal dicing). **Down right:** p_{min} distribution for (B) (normal dicing) and $(S + B)$ (lognormal dicing)

is certainly an important aspect of the applied search algorithm that it should not only spot existing deviations but also correctly identify those classes in which the data is consistent with the background expectation. Another possibility to make crosschecks in a single class with significant signal contribution is to scale down the signal cross section. An example for GM1e in which the calculated cross section is lowered by a factor of ten can be found in figure B.10 in the appendix. Claiming $\tilde{P} = 1.5\,\sigma$ both p-values indicate that the $2\gamma 1\mathrm{MET} + X$ class is insignificant in this consistency check.

The $\sum p_T$ inclusive class $2\gamma 4\mathrm{jet}1\mathrm{MET} + X$ is presented as a second example. Figure 10.8 shows that requiring at least four additional jets makes the class almost background free. While often only the tails of a distribution are affected by lacking statistics due to the limited number of simulated background events, this particular class has not enough simulated events for some samples at lower values of $\sum p_T$. The right histogram in figure 10.8 exemplifies for a specific pseudo experiment that the fill-up procedure introduced in section 8.5 leads to reasonable results.

Though testing the algorithm in specific classes and distributions provides useful insights, it is important to demonstrate that the full analysis is also working in a more realistic scenario in which there is a multitude of classes that need to be evaluated without favoring a specific beyond Standard Model scenario. Apart from testing the scalability of the analysis framework

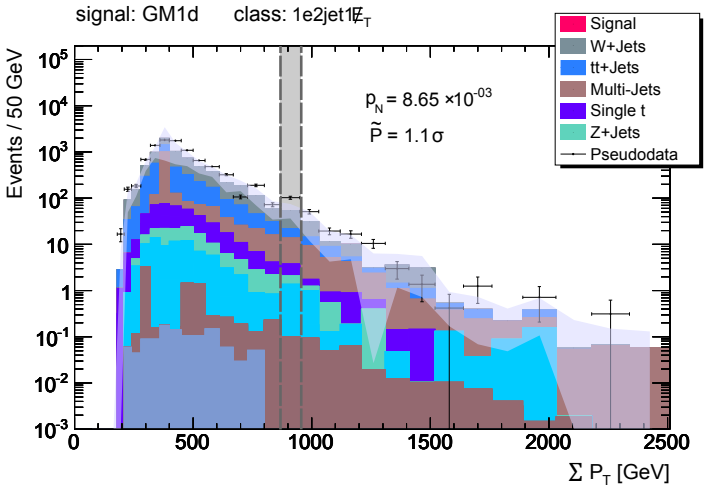

Figure 10.7 Single pseudo experiment in the class 1*e*2jet1MET with signal GM1d at $1\,\mathrm{pb}^{-1}$.

this allows to investigate how signal contributions to a subset of classes can change the global $\widetilde{P}$ distribution that would be expected if the data fully corresponds to the Standard Model. Again, the Standard Model expectation is given by the combination of event and detector simulation. In principle it may also include supplementary inputs from data driven methods.

Corresponding to the scope of this work the evaluated classes are restricted to those requiring at least one selected photon. This criterion was met by about 1300 classes. For each of the up to three distributions in each given class $1.5 \cdot 10^4$ $(S+B)$ experiments and $1 \cdot 10^5$ (B) experiments are diced, corresponding to a limit of about $4.4\,\sigma$ in $\widetilde{P}$. The number of diced $(S+B)$ experiments is high enough to ensure a stable median even if the corresponding p_{min} distribution is rather broad. To speed up the algorithm, regions in which less than three events are observed are skipped if at the same time less than three events are expected. In practice this or a similar requirement helps to reduce the number of apparently significant deviations that are based on only few events. The dicing of the chosen number of pseudo experiments for more than 1000 classes cannot be conducted within a reasonable amount of time on a single desktop machine. Thus for each class a dedicated grid job is sent to the local grid cluster. The full procedure has

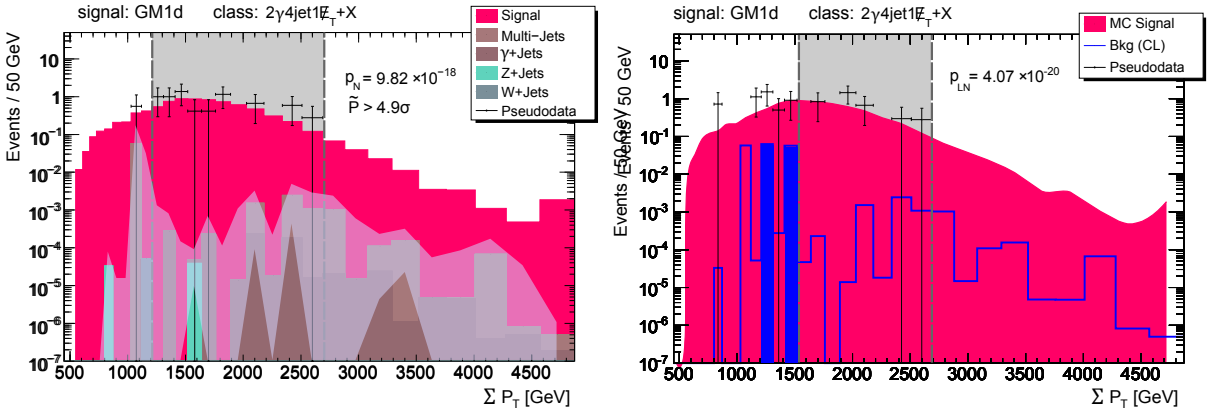

Figure 10.8 Evaluation of the inclusive class with at least two photons, four jets and MET for the GM1d test scenario with $L = 1\,\mathrm{fb}^{-1}$ at $\sqrt{s} = 10\,\mathrm{TeV}$ **Left:** $(S+B)$ pseudo experiment using normal dicing, (the error band illustrates the uncertainty on the background expectation). **Right:** $(S+B)$ pseudo experiment using lognormal dicing exemplifying the fill up procedure (see section 8.5). Empty bins that have been filled up are fully colored in blue

Event Class	$p_{LN,min}$ ($S+B$ median)	$\widetilde{P}$ (Significance)	# SigMC	# BkgMC
2γ5jet1MET+X MET	$1.80 \cdot 10^{-36}$	$> 4.41\,\sigma$	13.3	0.01
2γ2jet1MET+X MET	$2.00 \cdot 10^{-33}$	$> 4.41\,\sigma$	33.5	6.27
2γ1MET+X MET	$3.85 \cdot 10^{-33}$	$> 4.41\,\sigma$	57.7	7.05
2γ5jet1MET+X M_{inv}	$5.85 \cdot 10^{-24}$	$> 4.41\,\sigma$	13.3	0.01
2γ5jet1MET+X $\sum p_T$	$7.63 \cdot 10^{-24}$	$> 4.41\,\sigma$	13.3	0.01
2γ1MET+X M_{inv}	$1.50 \cdot 10^{-23}$	$> 4.41\,\sigma$	57.7	7.05
2γ1MET+X $\sum p_T$	$1.10 \cdot 10^{-19}$	$> 4.41\,\sigma$	57.7	7.05
1e1γ7jet+X $\sum p_T$	$3.35 \cdot 10^{-19}$	$> 4.41\,\sigma$	3.4	0.00
1e1γ7jet+X M_{inv}	$3.35 \cdot 10^{-19}$	$> 4.41\,\sigma$	3.4	0.00
1e1γ5jet1MET+X MET	$7.30 \cdot 10^{-19}$	$> 4.41\,\sigma$	9.0	0.00
1μ1γ2jet1MET+X M_{inv}	$2.91 \cdot 10^{-16}$	$> 4.41\,\sigma$	15.2	20.61
1μ1γ2jet1MET+X $\sum p_T$	$1.76 \cdot 10^{-15}$	$> 4.41\,\sigma$	15.2	20.61
2γ2jet1MET+X M_{inv}	$3.92 \cdot 10^{-13}$	$> 4.41\,\sigma$	33.5	6.27
2γ2jet1MET+X $\sum p_T$	$6.72 \cdot 10^{-12}$	$> 4.41\,\sigma$	33.5	6.27

Table 10.2 A selected subset of the most interesting inclusive event classes with test signal GM1d at $1\,\mathrm{fb}^{-1}$. For this scan lognormal dicing and p_{LN} was used for the evaluation of $(S+B)$. The significance $\widetilde{P}$ refers to the fraction of background only pseudo experiments with $p_{\min}(B) < p_{\min}^{\mathrm{median}}(S+B)$ out of $1 \cdot 10^5$.

been conducted twice, evaluating the $(S+B)$ experiments with p_N in the first cycle and with p_{LN} in the second.

Several details in the practical implementation of the algorithm help to reduce the number of dicing steps and p-value calculations if regions are found to be clearly insignificant. Thus the analysis of many classes terminates quickly. In contrast to this classes with considerable GM1d signal contributions and a good detector resolution leading to a high number of fully evaluated regions run for more than a day until termination. Almost 100% of the grid jobs could be successfully retrieved.

Table 10.2 lists a selection of the most significant inclusive event classes that have been found in the full scan of event classes while dicing $(S+B)$ with p_{LN}. The median of the p_{min} distribution from dicing $(S+B)$ is given along with the value of $\widetilde{P}$. If not enough pseudo experiments could be diced the respective limit on $\widetilde{P}$ is given. The last two columns contain the integrated signal and background expectations for the whole respective distribution.

For all evaluated distributions significant deviations are found in some of the classes. As expected many of the highly significant classes are found to require two photons and missing transverse energy. While requiring additional jets may reduce the signal efficiency it also lowers the expected backgrounds so that some classes requiring additional jets are found to be very significant. Also some of the classes requiring a muon or electron in addition to a photon and missing transverse energy show interesting deviations. This is not unexpected as typical decay chains like $qg \to 5q + 2W + 2\widetilde{\chi}_1^0$ [26] result in the production of W bosons which can lead to isolated muons or electrons through leptonic decays. Additionally there is the already mentioned possibility that the NLSP does not decay into gravitino plus photon. In table 10.3 the results that have been derived by dicing $(S+B)$ with p_N are presented for the same distributions as chosen for table 10.2. With GM1d at $1\,\mathrm{fb}^{-1}$ one is discussing a scenario with highly significant excesses of observed events. Thus it is in agreement with what has been discussed in sections 8.3 and 8.4 that the observed median values of p_{min} are found to be lower for all the distributions

Event Class	$p_{LN,min}$ ($S+B$ median)	$\widetilde{P}$ (Significance)	# SigMC	# BkgMC
2γ1MET+X M_{inv}	$2.09 \cdot 10^{-45}$	$> 4.41\,\sigma$	57.7	7.05
2γ1MET+X MET	$9.43 \cdot 10^{-45}$	$> 4.41\,\sigma$	57.7	7.05
2γ1MET+X $\sum p_T$	$7.03 \cdot 10^{-43}$	$> 4.41\,\sigma$	57.7	7.05
2γ5jet1MET+X MET	$4.26 \cdot 10^{-41}$	$> 4.41\,\sigma$	13.3	0.01
2γ2jet1MET+X MET	$5.69 \cdot 10^{-41}$	$> 4.41\,\sigma$	33.5	6.27
2γ5jet1MET+X M_{inv}	$2.40 \cdot 10^{-27}$	$> 4.41\,\sigma$	13.3	0.01
2γ5jet1MET+X $\sum p_T$	$7.48 \cdot 10^{-27}$	$> 4.41\,\sigma$	13.3	0.01
2γ2jet1MET+X M_{inv}	$1.32 \cdot 10^{-24}$	$> 4.41\,\sigma$	33.5	6.27
2γ2jet1MET+X $\sum p_T$	$7.98 \cdot 10^{-23}$	$> 4.41\,\sigma$	33.5	6.27
1e1γ7jet+X $\sum p_T$	$1.59 \cdot 10^{-19}$	$> 4.41\,\sigma$	3.4	0.00
1e1γ7jet+X M_{inv}	$1.59 \cdot 10^{-19}$	$> 4.41\,\sigma$	3.4	0.00
1e1γ5jet1MET+X MET	$3.66 \cdot 10^{-20}$	$> 4.41\,\sigma$	9.0	0.00
1μ1γ2jet1MET+X M_{inv}	$1.54 \cdot 10^{-17}$	$> 4.41\,\sigma$	15.2	20.61
1μ1γ2jet1MET+X $\sum p_T$	$2.29 \cdot 10^{-17}$	$> 4.41\,\sigma$	15.2	20.61

Table 10.3 A selected subset of the most interesting inclusive classes (corresponding to table 10.2) with test signal GM1d at $1\,\mathrm{fb}^{-1}$ using p_N and normal dicing. The significance $\widetilde{P}$ refers to the fraction of background only pseudo experiments with $p_{\min}(B) < p_{\min}^{\mathrm{median}}(S+B)$ out of $1 \cdot 10^5$.

when dicing according to p_N. As the differences will be dependent on factors like the uncertainty on the background mean in the significant regions the order of the classes with respect to the median $p_{N,min}$ can shift.

Tables 10.4 and 10.5 list the distributions that are found to be most significant for the exclusive event classes. These classes are presented for dicing signal plus background according to p_{LN} and p_N respectively. For exclusive classes the dominant two photon plus missing transverse energy signal is distributed over classes with differing jet multiplicity and thus single classes can not reach the lowest p-values that are observed for inclusive classes. Again dicing $(S+B)$ according to p_{LN} provides results that are somewhat more conservative showing less classes for which none of the $1 \cdot 10^5$ background only experiments features $p_{\min}(B) < p_{\min}^{\mathrm{median}}$. As an example of a single class one can for example consider the 1γ7jet1MET class. An expected significance of $3.35\,\sigma$ is found for the M_{inv} distribution using p_{LN} while $3.74\,\sigma$ are claimed when using p_N. However, it should be emphasized that speaking in general the two applied p-values come to similar results. But if many event classes are evaluated some of them will be subject to high systematic uncertainties and therefore favor notable differences between p_N and p_{LN} (see section 8.1). Taking into account the complete set of event classes the results suggest that though p_{LN} emphasizes the possibility of longer tails for the background prior due to multiplicative uncertainties, both p-values work similarly well for most of the considered distributions.

Evaluating a multitude of event classes one can compare the $\widetilde{P}$ distribution found in the data or in a single set of $(S+B)$ experiments for each class with the expectation in case of a global agreement between (pseudo-)data and background estimation. As $\widetilde{P} = Pr\,(p_{min}(B) < p_{min}(data))$ the $\widetilde{P}$ distribution should be uniform in the case of global agreement. In order to present the results in an accessible way $-\log_{10}\widetilde{P}$ is plotted. Assuming that the $\widetilde{P}$ distribution is normalized to $C \in \mathbb{R}$ the uniform distribution is transformed via

$$X = -\log_{10}\widetilde{P} \quad , \quad f(X)\,\mathrm{d}X \stackrel{!}{=} C\,\mathrm{d}\widetilde{P} \quad . \tag{10.3}$$

Event Class	$p_{LN,min}$ ($S+B$ median)	$\widetilde{P}$ (Significance)	# SigMC	# BkgMC
2γ5jet1MET MET	$1.29 \cdot 10^{-12}$	$> 4.41\,\sigma$	3.7	0.01
1γ7jet1MET MET	$1.30 \cdot 10^{-10}$	$> 4.41\,\sigma$	5.6	0.37
1γ6jet1MET MET	$2.46 \cdot 10^{-09}$	$> 4.41\,\sigma$	7.9	2.38
2γ6jet1MET $\sum p_T$	$1.05 \cdot 10^{-08}$	$> 4.41\,\sigma$	3.1	0.00
2γ6jet1MET M_{inv}	$1.05 \cdot 10^{-08}$	$> 4.41\,\sigma$	3.1	0.00
2γ3jet1MET M_{inv}	$2.20 \cdot 10^{-08}$	$> 4.41\,\sigma$	4.5	0.29
2γ5jet1MET $\sum p_T$	$1.57 \cdot 10^{-07}$	$> 4.41\,\sigma$	3.7	0.01
2γ3jet1MET MET	$3.29 \cdot 10^{-07}$	$> 4.41\,\sigma$	4.5	0.29
2γ6jet1MET MET	$3.11 \cdot 10^{-09}$	$> 4.41\,\sigma$	3.1	0.00
2γ5jet1MET M_{inv}	$3.30 \cdot 10^{-07}$	$> 4.41\,\sigma$	3.7	0.01
2γ1MET $\sum p_T$	$1.07 \cdot 10^{-06}$	$4.41\,\sigma$	5.7	0.35
2γ1MET M_{inv}	$8.08 \cdot 10^{-07}$	$4.26\,\sigma$	5.7	0.35
2γ1MET MET	$2.35 \cdot 10^{-06}$	$4.26\,\sigma$	5.7	0.35
2γ3jet1MET $\sum p_T$	$3.16 \cdot 10^{-07}$	$4.17\,\sigma$	4.5	0.29
2γ2jet1MET MET	$3.46 \cdot 10^{-07}$	$4.05\,\sigma$	6.9	5.66
1γ7jet1MET M_{inv}	$2.16 \cdot 10^{-04}$	$3.35\,\sigma$	5.6	0.37
1γ7jet1MET $\sum p_T$	$2.44 \cdot 10^{-04}$	$3.30\,\sigma$	5.6	0.37
1γ5jet1MET MET	$4.46 \cdot 10^{-04}$	$2.87\,\sigma$	8.3	43.34
1γ6jet1MET $\sum p_T$	$1.19 \cdot 10^{-04}$	$2.81\,\sigma$	7.9	2.38
2γ2jet1MET $\sum p_T$	$3.75 \cdot 10^{-04}$	$2.78\,\sigma$	6.9	5.66
1γ6jet1MET M_{inv}	$1.28 \cdot 10^{-04}$	$2.78\,\sigma$	7.9	2.38
2γ2jet1MET M_{inv}	$6.61 \cdot 10^{-04}$	$2.65\,\sigma$	6.9	5.66

Table 10.4 The 22 most significant exclusive event classes with test signal GM1d at $1\,\mathrm{fb}^{-1}$. For this scan lognormal dicing and p_{LN} was used for the evaluation of $(S+B)$.

To calculate the Jacobian determinant is straightforward.

$$\mathrm{d}\widetilde{P} = \left|\frac{\delta\widetilde{P}}{\delta X}\right| \mathrm{d}X = \left|\frac{\delta 10^{-X}}{\delta X}\right| \mathrm{d}X = \ln(10) \cdot 10^{-X} \mathrm{d}X \tag{10.4}$$

Thus one arrives at

$$f(X) = C \cdot \ln(10) \cdot 10^{-X} \quad . \tag{10.5}$$

Plotting a logarithmically scaled y-axis the $\widetilde{P}$ distribution is expected to fall linearly, because of

$$\log_{10}(f(X)) = \log_{10}(C \cdot \ln 10) - X \quad . \tag{10.6}$$

In figure 10.9 two examples showing $\widetilde{P}$ distributions are plotted with respect to the $\sum p_T$ variable. Each of them compares an $(S+B)$ global pseudo experiment including GM1d with a (B) global pseudo experiment picking $\widetilde{P}$ for one diced value of p_{min}. For the given examples both the background only and the signal plus background distributions are normalized to the number of evaluated event classes. Therefore the number of entries for (B) and $(S+B)$ cannot only differ due to statistical fluctuations in the number of observed classes, but also because an additional signal may tend to increase the number of populated classes.

In concordance with what has been said above an approximately linear evolution is found for the background only scenario representing a global agreement between data and assumptions. In the case of an additional GM1d signal more classes with high deviations are found. They cause a

Event Class	$p_{LN,min}$ ($S+B$ median)	$\widetilde{P}$ (Significance)	# SigMC	# BkgMC
2γ5jet1MET MET	$2.77 \cdot 10^{-13}$	$> 4.41\,\sigma$	3.7	0.01
2γ1jet1MET $\sum p_T$	$1.19 \cdot 10^{-12}$	$> 4.41\,\sigma$	9.3	0.23
2γ1jet1MET M_{inv}	$1.19 \cdot 10^{-12}$	$> 4.41\,\sigma$	9.3	0.23
1γ7jet1MET MET	$4.07 \cdot 10^{-12}$	$> 4.41\,\sigma$	5.6	0.37
2γ1jet1MET MET	$1.23 \cdot 10^{-11}$	$> 4.41\,\sigma$	9.3	0.23
1γ6jet1MET MET	$2.25 \cdot 10^{-10}$	$> 4.41\,\sigma$	7.9	2.38
2γ6jet1MET MET	$5.76 \cdot 10^{-10}$	$> 4.41\,\sigma$	3.1	0.00
2γ6jet1MET $\sum p_T$	$1.48 \cdot 10^{-09}$	$> 4.41\,\sigma$	3.1	0.00
2γ6jet1MET M_{inv}	$1.48 \cdot 10^{-09}$	$> 4.41\,\sigma$	3.1	0.00
2γ3jet1MET M_{inv}	$7.08 \cdot 10^{-09}$	$> 4.41\,\sigma$	4.5	0.29
2γ3jet1MET $\sum p_T$	$8.76 \cdot 10^{-09}$	$> 4.41\,\sigma$	4.5	0.29
2γ4jet1MET MET	$1.38 \cdot 10^{-08}$	$> 4.41\,\sigma$	3.9	0.38
2γ5jet1MET M_{inv}	$2.48 \cdot 10^{-08}$	$> 4.41\,\sigma$	3.7	0.01
2γ5jet1MET $\sum p_T$	$4.16 \cdot 10^{-08}$	$> 4.41\,\sigma$	3.7	0.01
2γ3jet1MET MET	$6.39 \cdot 10^{-08}$	$> 4.41\,\sigma$	4.5	0.29
2γ1MET $\sum p_T$	$3.03 \cdot 10^{-07}$	$> 4.41\,\sigma$	5.7	0.35
2γ1MET MET	$9.13 \cdot 10^{-07}$	$> 4.41\,\sigma$	5.7	0.35
2γ2jet1MET MET	$1.41 \cdot 10^{-07}$	$4.41\,\sigma$	6.9	5.66
2γ1MET M_{inv}	$3.90 \cdot 10^{-07}$	$4.17\,\sigma$	5.7	0.35
2γ2jet1MET $\sum p_T$	$2.15 \cdot 10^{-05}$	$3.42\,\sigma$	6.9	5.66
1γ7jet1MET $\sum p_T$	$2.53 \cdot 10^{-05}$	$3.74\,\sigma$	5.6	0.37
1γ7jet1MET M_{inv}	$2.80 \cdot 10^{-05}$	$3.74\,\sigma$	5.6	0.37

Table 10.5 The 22 most significant exclusive event classes with test signal GM1d at $1\,\mathrm{fb}^{-1}$ using normal dicing and p_N.

change in the $\widetilde{P}$ distribution. In this way the found deviation between data and the background only assumption for a given variable can provide a measure of the global agreement between data and Standard Model expectation. Should signal contributions be spread over many classes

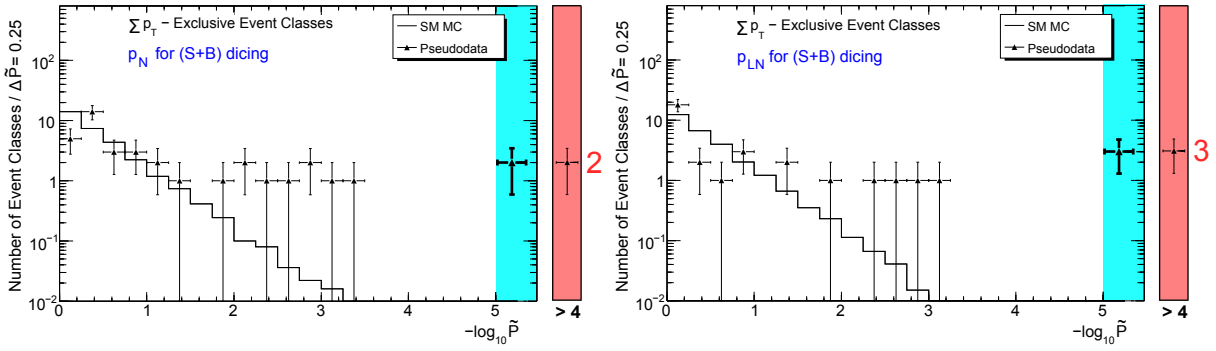

Figure 10.9 Left: A global pseudo experiment with signal GM1d using p_N for $(S+B)$ dicing$\widetilde{P}$ — evaluation of the $\sum p_T$ variable in the exclusive classes. The shaded bin includes all classes in which no background only pseudo experiment out of 100000 is more significant than the selected value p_{min}. The separate bin summarizes all classes with $-\log_{10}\widetilde{P} > 4$. **Right:** the same, but using p_{LN} for $(S+B)$ dicing.

plotting the values of $\widetilde{P}$ may point out deviations that would not cause noteworthy significances in any single class.

In principle it is also possible to draw the $\widetilde{P}$ distribution for a given variable with respect to the inclusive classes. In this class an interpretation of results is more complicated. Inclusive classes are not disjunct and their respective deviations are therefore correlated. This makes a quantification of observed significances difficult. During test scenarios with a simulated signal all classes are evaluated independently for a given pseudo experiment. Two examples of the outcome of inclusive global pseudo experiments in which all distributions are treated as independent are appended in figure B.11.

10.2 Excited Muons

Sections 10.1 and 10.3 focus on scenarios in which the main signatures are composed of photons and missing transverse energy. A composite model with excited muons allows to study a complementary signal with a lepton plus photon signature. The evaluated benchmark scenarios are presented in table 10.6.

compositeness scale Λ	mass parameter m_{μ^*}	cross section σ
6.0 TeV	0.4 TeV	176 fb
6.0 TeV	1.0 TeV	44 fb

Table 10.6 Evaluated benchmark points in a scenario with excited muons (see table C.1, single μ^* production and decay via $\mu^* \to \mu\gamma$)

In section 3.2 it has been discussed that for the selected parameter points the excited muons decay mainly via gauge interactions. According to equations (3.13) and (3.14) about 30% of the excited muons decaying via gauge interaction will decay into a muon and a photon [148, figure 3.4]. The mechanism of the single μ^* production is dominated by contact interactions. The respective Feynman graphs for single μ^* production and the decay $\mu^* \to \mu\gamma$ are shown in figure 10.10.

Again a center of mass energy of 10 TeV is assumed. The simulated scenarios are evaluated at an integrated luminosity of $100\,\mathrm{pb}^{-1}$. The left plot in figure 10.11 shows a representative $(S+B)$ pseudo experiment in the $\sum p_T$ distribution of the $2\mu 1\gamma + X$ inclusive class with $m_{\mu^*} = 0.4\,\mathrm{TeV}$. One can see that the signature of the signal sample is clearly significant. The main background in the signal region is given by the Drell-Yan process. It is found that $\widetilde{P} > 4.8\,\sigma$. For a more detailed discussion of this scenario one might further investigate potential contributions from $t\bar{t}$ and diboson backgrounds in the high energy tail of the distribution (see e.g. [148]).

The right plot in figure 10.11 presents the outcome of an exemplary $(S+B)$ pseudo experiment in the $2\mu 1\gamma$ exclusive class with a μ^* signal at $m_{\mu^*} = 1\,\mathrm{TeV}$.

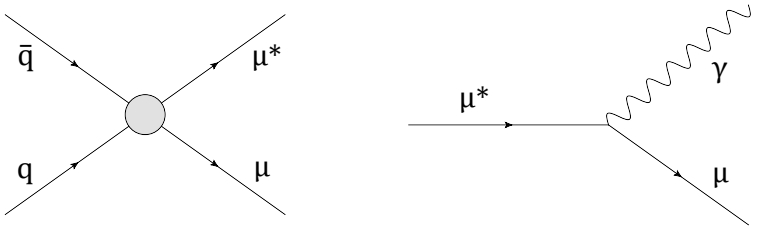

Figure 10.10 **Left:** single μ^* production via contact interaction.
Right: decay $\mu^* \to \mu\gamma$ via gauge interaction

Dicing $1 \cdot 10^6$ background only and $1.5 \cdot 10^4$ signal plus background experiments both p_N and p_{LN} arrive at a significance of $\widetilde{P} = 3.4\sigma$ for $m_{\mu^*} = 1\,\text{TeV}$. This supports the assumption that in many situations the results of the algorithm for intermediate significances will be robust against an exchange of p_N and p_{LN}. With respect to the evaluated example at $m_{\mu^*=1\,\text{TeV}}$ it needs to be emphasized that in many situations one would of course be very cautious if observations are based on only few events.

Further aspects concerning the evaluation of excited muons within the MUSiC framework are discussed in [149]. A dedicated study about the discovery potential of excited muons in CMS is presented in [148].

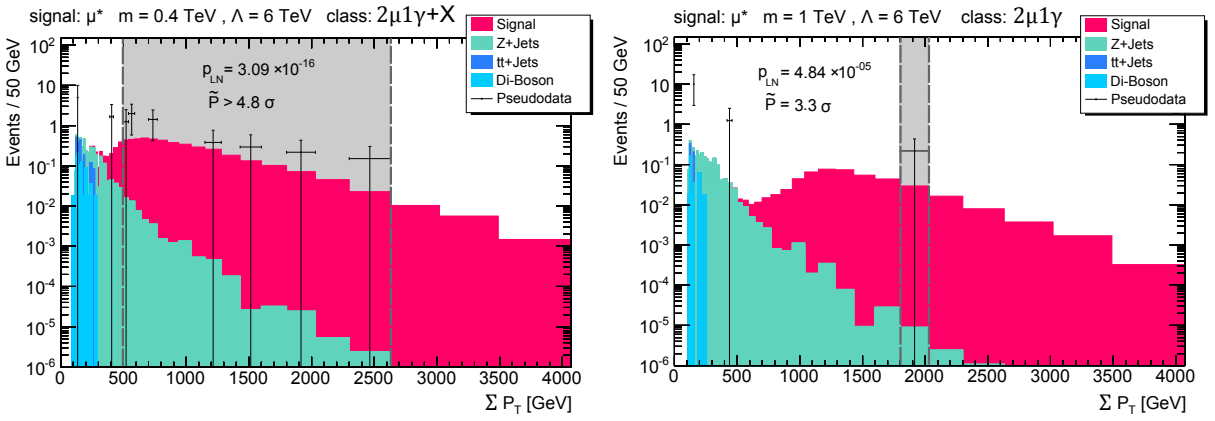

Figure 10.11 **Left:** representative $(S+B)$ pseudo experiment for a scenario with signal μ^* at $m_{\mu^*} = 0.4\,\text{TeV}$. **Right:** representative $(S+B)$ pseudo experiment for a scenario with signal μ^* at $m_{\mu^*} = 1\,\text{TeV}$. One needs to keep in mind the application of dynamic binning to interpret the figures in a correct way.

10.3 Diphoton Signature in a Model with Large Extra Dimensions

It has been outlined in section 3.3 that models extending the usual flat Lorentzian spacetime with additional large extra dimensions provide another source of potential photon signatures. Dedicated studies (e.g. [150] for CMS, see [151] for a short summary) suggest that a certain range of such scenarios is within the LHC discovery reach. Here the considered signature is given by diphoton production via virtual gravitons. It is evaluated if the presented model unspecific approach is capable of finding such a signal.

After discussing parameters of a reasonable benchmark scenario the production of the ADD[3] signal and the diphoton background was conducted using the SHERPA event generator [71]. SHERPA generates events according to a predefined final state. Details concerning the production of ADD samples with SHERPA can be found in [152]. One diphoton sample was produced including virtual graviton production taking into account possible interference between the signal process and QCD diphoton production. As the lower p_T region is expected to be background dominated the sample features a minimal p_T threshold. In this section a p_T threshold refers to a cut on the transverse momentum of the generated photons. Such a threshold allows for modelling the signature with a production of only a moderate number of events. Another sample was produced including QCD diphoton production only. In order to have sufficient event statistics not just for the signal region but also for the range of lower p_T the diphoton background was subdivided into two p_T regions and generated separately. The parameters that are specified for signal plus QCD diphoton background production[4] are given in table 10.7. The selected

[3] see section 3.3

[4] The HLZ convention is used to describe processes with a virtual graviton (see [153, 154]).

benchmark scenario lies in a region of the parameter space that is not experimentally excluded (see [155] for recent exclusion limits). Some details concerning the QCD diphoton background samples are given in table 10.8. The $u\bar{u}$ annihilation Feynman graphs for the virtual graviton signal contribution and the irreducible Born diagram based QCD diphoton background are depicted in figure 10.12. The events are generated one jet inclusive and the underlying event is taken into account. Standard CMSSW simulation and reconstruction steps using the ideal V11 alignment scenario were applied to the generated event samples (see figure 6.1).

M_f	n_{ED}	$p_{T,\mathrm{min}}$	# events	cross section	production
2.5 TeV	4	300 GeV	~ 30000	500fb	1 jet inclusive, with underlying event

Table 10.7 Benchmark point for Diphoton production via a virtual graviton (HLZ convention) in a scenario with large extra dimensions (LED).

Per default the MUSiC framework expects a standalone signal sample. However, the algorithm could be adapted to the needs of this special example. For the dicing of the background only experiments the background samples are used including the two SHERPA based QCD diphoton samples. For $(S+B)$ the high p_T diphoton background sample was replaced by the corresponding diphoton signal plus background sample.

Direct QCD diphoton production is also included in the official photon plus jets production (see table C.2). To avoid double counting of this background contribution, diphoton events have been removed from the photon plus jets samples at generator level. In order to take into account the accidental overlap between the two SHERPA based diphoton background samples the lower p_T sample was made disjunct from the high p_T sample by including an additional p_T cut at generator level.

$p_{T,\mathrm{min}}$	$p_{T,\mathrm{max}}$	# events	cross section	production
100 GeV	400 GeV	~ 30000	260fb	1 jet inclusive, with underlying event
300 GeV	-	~ 30000	9690fb	1 jet inclusive, with underlying event

Table 10.8 QCD diphoton background samples

The selected benchmark scenario can be used to evaluate the respective sensitivity of the presented model unspecific approach in the prominent $2\gamma + X$ inclusive class. In the default setting this class has a rather high bin multiplicity. Thus a minimum bin width of 40 GeV was chosen to speed up the dicing of the pseudo experiments. For this test an integrated luminosity of $100\,\mathrm{pb}^{-1}$ has been assumed. The outcome of an exemplary $(S+B)$ pseudo experiment is shown in figure 10.13. The magenta colored area labeled "signal" marks the effective signal contribution

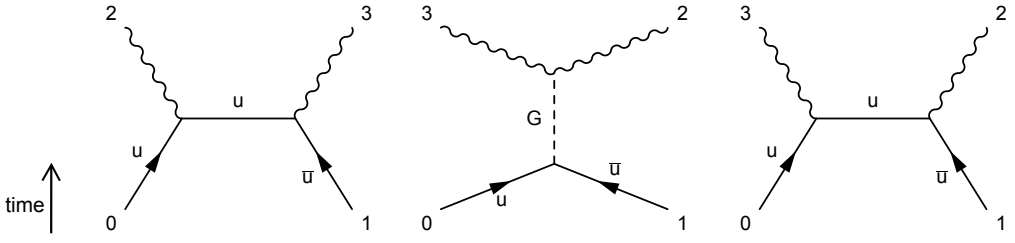

Figure 10.12 Feynman graphs for Born diagram QCD diphoton production and virtual graviton diphoton production for $u\bar{u}$ annihilation

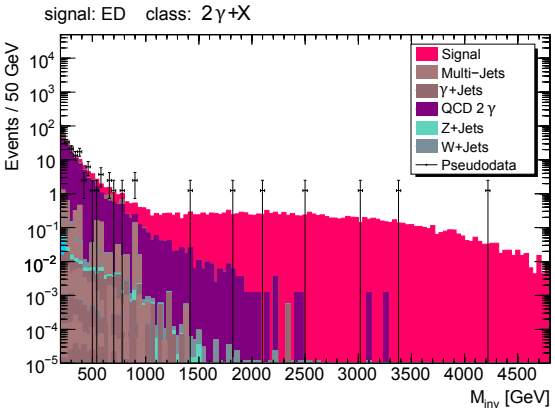

Figure 10.13 Single $(S + B)$ pseudo experiment (Pseudodata) for the M_{inv} distribution of $2\gamma + X$ inclusive class with LED diphoton signal at $100\,\mathrm{pb}^{-1}$. The expectation of signal and background is indicated by the stacked bin entries.

from the ED benchmark scenario. Speaking more technically it is the difference between the contribution from QCD only diphoton production and the virtual graviton plus QCD diphoton production. In the region of lower values of $\sum p_T$ the difference between (B) and $(S + B)$ is smoothly reaching a range where the signal is no longer significant. Hence it is confirmed that a sufficiently low p_T threshold has been chosen for the production of the "signal" sample. The dominant observed background is from direct diphoton production. Additional contributions from QCD multijet events and QCD photon+Jets production seem to be of less importance.

As the generation of the diphoton signal plus background sample and the diphoton background samples is one jet inclusive, no additional k-factors are assumed.

In a dedicated study at CMS [154], in which the signal events are generated exclusively in leading order, k-factors for both the diphoton signal plus background sample and the diphoton background samples have been assumed. It has been estimated that $k = 1.3$ for both samples [154]. Figure 10.14 presents the p_{min} distribution for $5 \cdot 10^5$ (B) and $1 \cdot 10^4$ $(S + B)$ pseudo experiments representing the application of p_N and p_{LN} as already outlined in section 3.1. For the presented scenario the two p-values are found to provide very similar results. Enough pseudo experiments are diced to claim $\widetilde{P} > 4.8\,\sigma$. The p_{min} distributions clearly suggest that the limit would turn out to be higher than 4.8σ if dicing an even higher number of pseudo experiments.

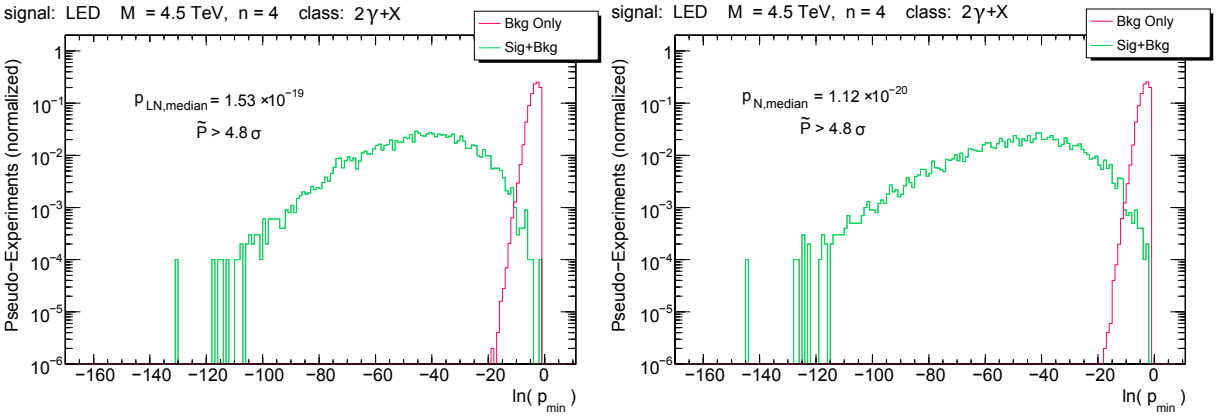

Figure 10.14 p_{min} distributions of $1 \cdot 10^4$ $(S + B)$ and $5 \cdot 10^5$ (B) pseudo experiments in the $2\gamma + X$ invariant mass distribution assuming $L = 100\,\mathrm{pb}^{-1}$ at $\sqrt{s} = 10\,\mathrm{TeV}$. **Reft:** normal dicing for (B) and $(S + B)$. **Right:** normal dicing for (B) and lognormal dicing for $(S + B)$

As a crosscheck the $2\gamma+X$ example was repeated scaling down the luminosity to $20\,\mathrm{pb}^{-1}$. In this case it is found that $\widetilde{P} = 1.5\,\sigma$ indicating that $20\,\mathrm{pb}^{-1}$ would certainly not be enough luminosity to expect any significant deviation.

Due to the differences in the generation of signal events a detailed comparison with the dedicated CMS analysis [154] is not fully straightforward. However, one can reason that for this specific test scenario the $\widetilde{P} > 5\,\sigma$ threshold can be expected to be in a similar range of luminosities as the $5\,\sigma$ discovery limit of the dedicated analysis which is not far from $100\,\mathrm{pb}^{-1}$ [154]. This is not too surprising since the corresponding dedicated analysis uses the high energy tail of the diphoton invariant mass spectrum and applies p_N to estimate the significance. Apparently in some cases the lack of optimized cuts and the penalty factor due to the look-elsewhere-effects can be compensated by the dynamic choice of the region of interest. The typical estimated uncertainty due to systematic uncertainties and limited statistics of the simulated events is about 20% in a typical region of interest. This is not too different from the value of about 30% that has been estimated for the background uncertainty in the signal region of the diphoton invariant mass distribution in the mentioned CMS signal specific diphoton analysis.

11 Conclusion

Several topics related to the extension and continuous development of a Model Unspecific Search at CMS have been presented.

The analysis framework has been successfully extended to new event classes with at least one selected isolated photon. There are several theoretical frameworks for physics beyond the Standard Model with photon based signatures. Consequently, the range of scenarios to which the analysis could be sensitive has been expanded into an important direction.

Cuts and variables for the object selection have been revised with a main focus on photon related aspects. A reasonable compromise between the photon efficiency and the reduction of fake objects has been found. In the context of updating to more recent software versions the CMS Physics Analysis Framework (PAT) has been integrated into the analysis workflow.

For the first time a p-value assuming a lognormal probability distribution function to model the uncertainty on the background mean has been included into the analysis framework of a large scaled model unspecific search. Several statistical aspects of the presented search algorithms have been discussed and coverage related properties of the p-values have been analyzed. It has been found that both p-values show a comparable performance, each of them having their respective advantages. Hence both p-values can be applied in parallel to provide mutual crosschecks and reduce the probability of statistical misinterpretations.

Simulated test scenarios have been analyzed to demonstrate the functionality and operational liability of the analysis framework. Taking into account both statistical and systematic uncertainties several conceivable models postulating the existence of new physics with high p_T photon signatures have been evaluated. A complete scan of photon specific event classes has been successfully conducted for one of those scenarios (GMSB, SPS 8 benchmark point GM1d).

The simulations show that a model unspecific analysis can help to localize significant deviations from the Standard Model expectation. Those could finally lead to the discovery of new physics in photon based event classes.

A Some Basic Properties of the Lognormal Probability Density Function

In the description of the search algorithm the lognormal probability density function (pdf) is justified to be a possible choice for a prior function modelling the uncertainties on the mean of a Poisson counting experiment. Some basic properties of the lognormal pdf are summarized in this section. As most of the following results can be derived by straightforward calculations not all the equations are linked to references.

Let Y be a normal distributed random variable with mean μ and variance σ^2 related to the random variable X via $\mathrm{Y} = \ln \mathrm{X}$. Then the pdf of X will be given by equation (A.1) [90]. We will use the shorthand expressions $N(\mu, \sigma)$ for the normal and $\Lambda(\mu, \sigma)$ for the lognormal distribution. Where it is not considered to be confusing, random variables following those distributions might sometimes be referred to in the same way.

$$f(x) = \frac{1}{\sqrt{2\pi} \cdot \sigma} \cdot \frac{1}{x} \exp\left[-\frac{(\ln x - \mu)^2}{2\sigma^2}\right] \tag{A.1}$$

Using the well-known definition of the error function $\mathrm{erf}(x)$ it is easily shown that the cumulative distribtion function (cdf) is given by (A.2).

$$\mathrm{F}(x') = \int_0^{x'} f(x)\,\mathrm{d}x = \frac{1}{2} + \frac{1}{2} \cdot \mathrm{erf}\left[\frac{\ln x' - \mu}{\sqrt{2} \cdot \sigma}\right] \tag{A.2}$$

Mean and variance of the lognormal pdf are (e.g. [124]):

$$E(\mathrm{X}) = \exp\left[\mu + \frac{1}{2}\sigma^2\right] \tag{A.3}$$

$$V(\mathrm{X}) = \exp\left[2\mu + \sigma^2\right] \cdot \left(\exp\left[\sigma^2\right] - 1\right) \tag{A.4}$$

By direct insertion into equation (A.2) it can be verified that the median is

$$m_{0.5}(\mathrm{X}) = \exp(\mu) \quad . \tag{A.5}$$

Equating the derivation of the lognormal pdf with 0 one finds that the maximum is given by the expression

$$x_{\max} = \exp\left(\mu - \sigma^2\right) \quad . \tag{A.6}$$

In the context of its application in this work it is advantageous to parametrize the lognormal function as it is given in [122].

$$f(b) = \frac{1}{\sqrt{2\pi} \cdot \ln k} \cdot \frac{1}{b} \exp\left[-\frac{\ln^2(b/b_0)}{2\ln^2 k}\right] \tag{A.7}$$

The parameters are related to those of the normal distributed random variable Y through

$$b_0 = \exp(\mu) \quad \text{and} \tag{A.8}$$

$$\sigma = \ln k \quad . \tag{A.9}$$

Hence in this parametrization b_0 is equal to the median and mean and variance read as

$$E\left(\mathrm{X}\right) = \exp\left[\ln b_0 + \frac{1}{2}\ln^2 k\right] = b_0 \cdot \exp\left[\frac{1}{2}\ln^2 k\right] \tag{A.10}$$

$$V\left(\mathrm{X}\right) = \exp\left[2\ln b_0 + \ln^2 k\right] \cdot \left(\exp\left[\ln^2 k\right] - 1\right) \tag{A.11}$$

One can notice that $\lim_{k\to 1} E\left(\mathrm{X}\right) = m_{0.5}\left(\mathrm{X}\right)$, meaning that for k in the vicinity of 1 the mean of the lognormal pdf will be close to its median.

By substitution of $y = \ln x$ one can derive that the cumulated probability within the interval $(\mu - c\cdot\sigma\,,\,\mu + c\cdot\sigma)$ of a pdf $N(y,\mu,\sigma)$ will be equal to that of a lognormal distribution $\Lambda(x,\mu,\sigma)$ in the interval $(m_{0.5}\cdot e^{c\cdot\sigma}\,,\,m_{0.5}/e^{c\cdot\sigma})$. This is summarized in equation (A.12).

$$\int_{\mu-c\cdot\sigma}^{\mu+c\cdot\sigma} N(y,\mu,\sigma)\mathrm{d}y \overset{y=\ln x}{=} \int_{m_{0.5}/e^{c\cdot\sigma}}^{m_{0.5}\cdot e^{c\cdot\sigma}} \Lambda(x,\mu,\sigma)\mathrm{d}x \tag{A.12}$$

An important feature of the lognormal distribution follows from the central limit theorem which states that under fairly general conditions [90, page 49] the sum of n random variables X_i will be distributed like a random variable X following a normal pdf $N\left(\mu,\sigma\right)$ for $n\to\infty$. From this statement it can be motivated that equivalently the product of n random variables will approach a lognormal probability density function as

$$\begin{aligned}
\exp\left[X_{N(\mu,\sigma)}\right] &= \exp\left[\lim_{n\to\infty}\sum_i X_i\right] \\
&= \lim_{n\to\infty}\exp\left[\sum_i X_i\right] \\
&= \lim_{n\to\infty}\prod_i \exp\left[X_i\right] \\
&= \lim_{n\to\infty}\prod_i Y_i \\
&= Y_{\Lambda(\mu,\sigma)} \quad .
\end{aligned} \tag{A.13}$$

For more technical formulations of the central limit theorem for products of random variables and an explicit statement of the underlying conditions see [124, page 13 et seq.].

The lognormal distribution has the nice reproductive property that

" If $\{X_j\}$ is a sequence of independent Λ-variates, where X_j is $\Lambda\left(\mu_j,\sigma_j\right)$, $\{b_j\}$ a sequence of constants and $c = e^a$ a positive constant, then provided $\sum_j b_j\mu_j$ and $\sum_j b_j^2\sigma_j^2$ both converge the product $c\prod_j X_j^{b_j}$ is $\Lambda\left(a + \sum_j b_j\mu_j, \sqrt{\sum_j b_j^2\sigma_j^2}\right)$ " (adapted notation) [124, theorem 2.3]

As a special case it can be derived from this theorem that the product of n lognormal random variables X_i will itself follow a lognormal distribution, because of

$$\prod_i X_{\Lambda(\mu_i,sigma_i)} = \Lambda\left(\sum_j \mu_j, \sqrt{\sum_j \sigma_j^2}\right) \quad . \tag{A.14}$$

In [122, page 10] it is mentioned that for a variable modeled by a lognormal pdf its shape will approach that of a normal pdf if the errors on the variable are relatively small (implying a small parameter σ in equation (A.1)) but no reference or proof is given there. Thus the statement is motivated and further specified by the following argument.

Consider a lognormal distribution (A.1) and a normal distribution $N\left(\mu',\sigma'\right)$. Both are unimodal with maxima given by μ' and equation (A.6) respectively. Clearly the maxima will

approach each other in the case of small σ if $\mu = \ln \mu'$. In this case the absolute value of the difference between the two distributions will be given by

$$\begin{aligned} \mathrm{I} &= \left| \frac{1}{\sqrt{2\pi}\cdot\sigma'} \exp\left[-\frac{(x-\mu')^2}{2\sigma'^2}\right] - \frac{1}{\sqrt{2\pi}\cdot\sigma}\cdot\frac{1}{x}\exp\left[-\frac{(\ln x-\mu)^2}{2\sigma^2}\right]\right| \\ &= \left| \frac{1}{\sqrt{2\pi}\cdot\sigma'} \exp\left[-\frac{(x-\mu')^2}{2\sigma'^2}\right] - \frac{1}{\sqrt{2\pi}\cdot\sigma}\cdot\frac{1}{x}\exp\left[-\frac{(\ln x-\ln\mu')^2}{2\sigma^2}\right]\right| \\ &< \max\left(\frac{1}{\sqrt{2\pi}\cdot\sigma'} \exp\left[-\frac{(x-\mu')^2}{2\sigma'^2}\right],\ \frac{1}{\sqrt{2\pi}\cdot\sigma}\cdot\frac{1}{x}\exp\left[-\frac{(\ln x-\ln\mu')^2}{2\sigma^2}\right]\right) \quad . \end{aligned} \tag{A.15}$$

Let us choose $\sigma = \sigma'$. Then because of the exponential factors it follows that for each $x \neq \mu'$ and $\delta > 0$ there exists a value $\sigma_0(x)$ so that

$$\mathrm{I} < \delta \quad \forall \quad \sigma < \sigma_0(x) \quad .$$

For the special case $x = \mu'$ we get

$$\mathrm{I} < \max\left(\frac{1}{\sqrt{2\pi}\cdot\sigma},\ \frac{1}{\sqrt{2\pi}\cdot\sigma}\cdot\frac{1}{\mu'}\right) \quad .$$

Again one can make I arbitrarily small by choosing a sufficiently small σ. Summarizing one can find a $\sigma_{\mathrm{lim}}(\delta_0, \delta_1)$ so that for a given neighborhood of radius r surrounding μ I will be smaller than δ_0 and the difference between the maxima will be smaller than δ_1 for all $\sigma < \sigma_{\mathrm{lim}}$.

B Supplementary Figures and Tables

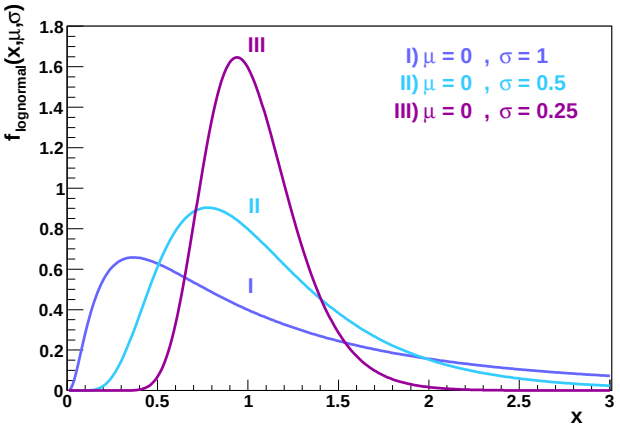

Figure B.1 Lognormal pdf with parametrization (A.1) for different values of σ if μ is fixed to 0

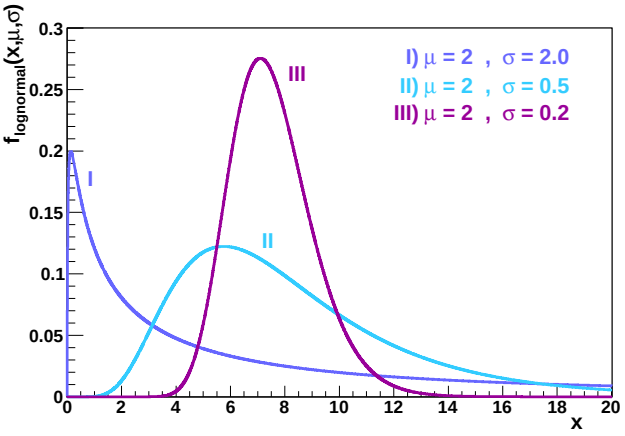

Figure B.2 Lognormal pdf with parametrization (A.1) for different values of σ if μ is fixed to 2

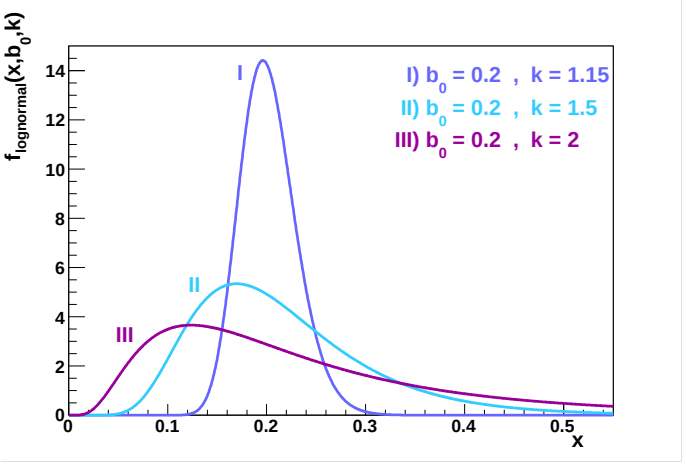

Figure B.3 Lognormal pdf with parametrization (A.7) for different values of k if b_0 is fixed to 0.2

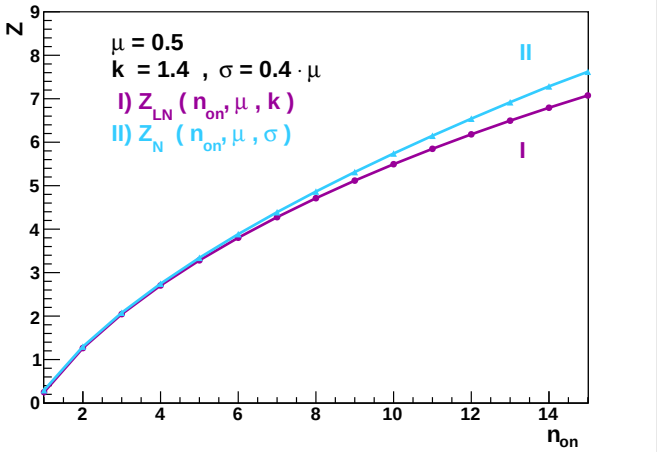

Figure B.4 p_{LN} and p_N as a function of the observed number of events n_{on} (other parameters fixed)

3.2 (modified)	dice the assumed background μ_{diced} for this pseudo-experiment from a truncated normal distribution $N_{trunc}(x, \mu_{true} + \Delta_\mu, \sigma_b)$ where $\Delta_\mu = \mu_{mean,diced} - \mu_{true}$ is the applied offset in μ

Table B.1 Modification to table 9.2 for testing p_N with an additional offset in μ

2. (modified)	fix the background μ_{true} and set $\sigma_{b,true} = f \cdot \mu_{true}$
3.2 (modified)	dice the assumed background μ_{diced} for this pseudo-experiment from a truncated normal distribution $N_{trunc}(x, \mu_{true}, \sigma_{b,true})$
3.4 (modified)	if $\mu_{diced} < n_{on}$ evaluate $Z_N\,(n_{on}, \mu_{diced}, \sigma_b)$, where $\sigma_b = \sigma_{b,true} + \Delta_\sigma$

Table B.2 Modification to the test procedure outlined in table 9.2 for testing p_N with a deviation in the assumed uncertainty

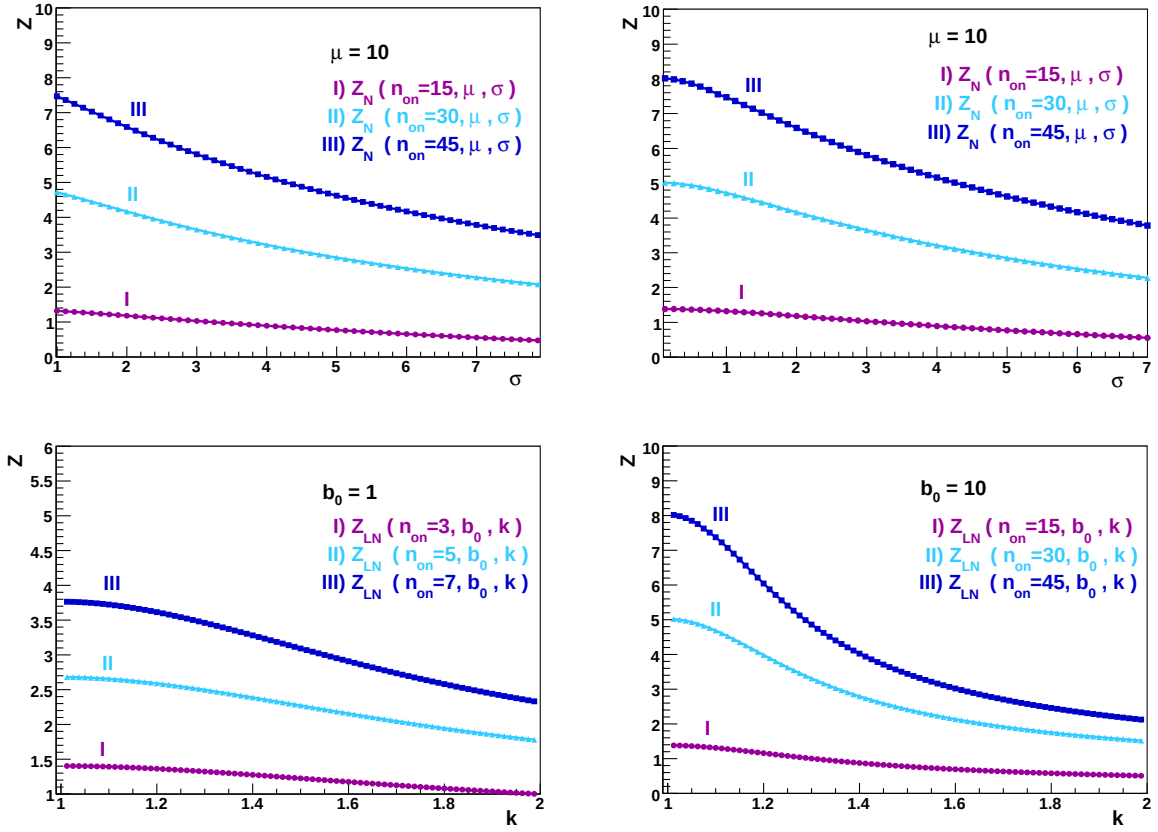

Figure B.5 **Up left and up right:** p_N as a function of the uncertainty σ on the median of the background prior (other parameters fixed). **Down left and down right:** p_{LN} as a function of the uncertainty k on the median of the background prior (other parameters fixed)

2. (modified)	fix the background μ_{true} and set $k_{true} = 1 + f$
3.2 (modified)	dice the assumed background $b_{0,diced}$ for this pseudo-experiment from a lognormal distribution $\Lambda_(x, b_0 = \mu_{true}, k_{true})$
3.4 (modified)	if $b_{0,diced} < n_{on}$ evaluate $Z_{LN}\,(n_{on}, b_{0,diced}, k)$, where $k = k_{true} + \Delta_k$

Table B.3 Modification to the test procedure outlined in table 9.5 for testing p_{LN} with a deviation in the assumed uncertainty

3.4 (modified)	if $\mu_{diced} < n_{on}$ evaluate $Z_{LN}\,(n_{on}, b_0, k)$, where $b_0 = \mu_{diced}$ and $k = 1 + f$

Table B.4 Changes to the algorithm outlined in table 9.2 for testing p_{LN}. Assuming $k = 1 + f$ and $b_0 = \mu_{diced}$ the background mean μ_{diced} is diced according to a normal distribution.

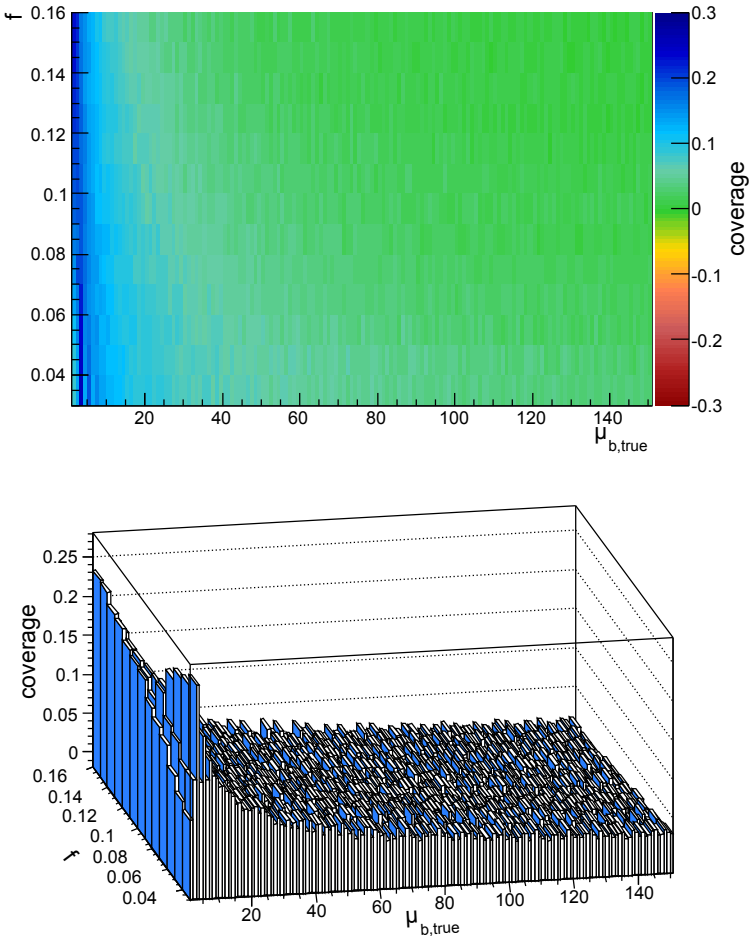

Figure B.6 Coverage of p_N probing $Z_{claim} = 2$ for the scenario outlined in table 9.2 with $\sigma_b = f \cdot \mu_{true}$

category 1	$0.8 < E/p_{in} < 1.2$	$\wedge$	$f_{\text{brem}} \begin{cases} < 0.06 & \text{, barrel} \\ < 0.1 & \text{, endcap} \end{cases}$
category 2	$0.8 < E/p_{in} < 1.2$	$\wedge$	$f_{\text{brem}} \begin{cases} > 0.06 & \text{, barrel} \\ > 0.1 & \text{, endcap} \end{cases}$
category 3	$E/p_{in} < 0.8$	$\vee$	$E/p_{in} > 1.2$

Table B.5 Definition of the different electron categories used for identification cuts. p_{in} is the estimated momentum at the track vertex. $f_{\text{brem}} = (p_{in} - p_{out})/p_{in}$ is a measure of the effects due to bremsstrahlung

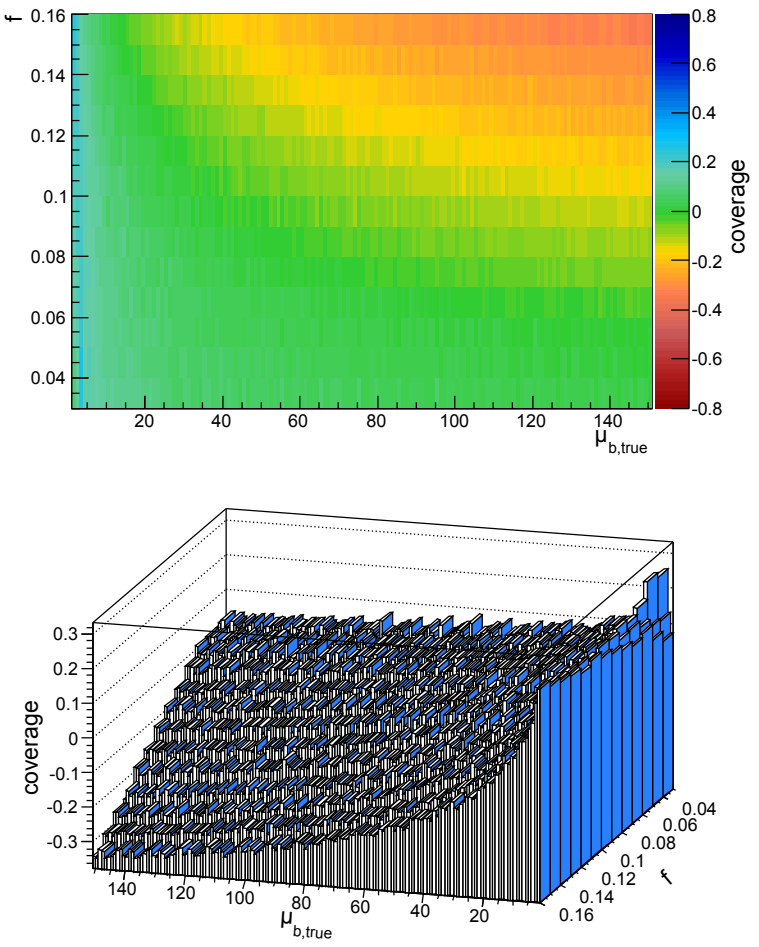

Figure B.7 Coverage of p_N probing $Z_{claim} = 2$ for a scenario with $\sigma_b = f \cdot \mu_{diced}$

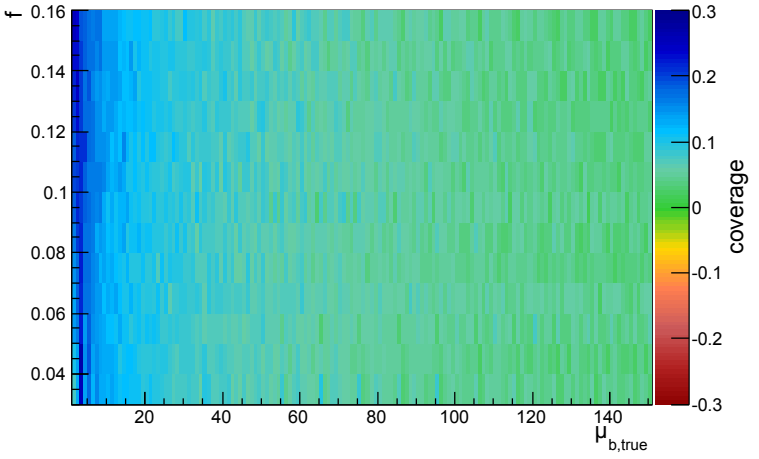

Figure B.8 Coverage of p_{LN} testing $Z_{claim} = 2$ for the scenario outlined in table 9.5 with $k = 1 + f$

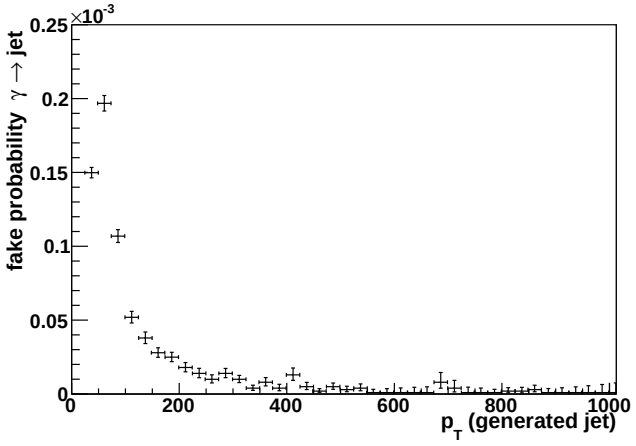

Figure B.9 Fake probability jet $\longrightarrow \gamma$ for generated jets as function of p_T of the generated jet. The jets are taken from the complete QCD multijet sample (see table C.2).

variable	region	category 1	category 2	category 3
H/E	barrel	< 0.042	< 0.050	< 0.045
	endcap	< 0.037	< 0.055	< 0.050
$\sigma_{\eta\eta}$	barrel	< 0.011	< 0.0125	< 0.01
	endcap	< 0.0252	< 0.0265	< 0.026
$\Delta\eta_{\text{in}}$	barrel	< 0.003	< 0.0055	< 0.0065
	endcap	< 0.0055	< 0.006	< 0.0075
$\Delta\phi_{\text{in}}$	barrel	< 0.016	< 0.032	< 0.0525
	endcap	< 0.035	< 0.025	< 0.065
$E_{\text{seed}}/p_{\text{in}}$	barrel	> 0.94	> 0.24	> 0.11
	endcap	> 0.83	> 0.32	> 0

Table B.6 "Tight" set of cuts for photon identification. For $E/p_{\text{in}} > 1.5$ the $\Delta\phi_{\text{in}}$ cut is replaced by 0.09 in the barrel region and 0.092 in the endcaps.

name	L1 prerequisite	basic requirement
muon trigger		
IsoMu15	$p_T > 10\,\text{GeV}$	$p_T > 15\,\text{GeV}$muon candidate, isolation criterion
Mu15	$p_T > 10\,\text{GeV}$	$p_T > 15\,\text{GeV}$ muon candidate
DoubleIsoMu3	two $p_T > 3\,\text{GeV}$	$p_T > 3\,\text{GeV}$ muon candidates, isolation criterion
DoubleMu3	two $p_T > 3\,\text{GeV}$	$p_T > 3\,\text{GeV}$ muon candidates
electron trigger		
IsoEle18 L1R	e/γ with $p_T > 15\,\text{GeV}$	$p_T > 18\,\text{GeV}$ electron candidate, isolation criterion
Ele15 LW L1R	e/γ with $p_T > 10\,\text{GeV}$	$p_T > 15\,\text{GeV}$ electron candidate
DoubleIsoEle12 L1R	two e/γ with $p_T > 10\,\text{GeV}$	two $p_T > 12\,\text{GeV}$ electron candidates, isolation criterion
DoubleEle10 LW OnlyPixelM L1R	two e/γ with $p_T > 5\,\text{GeV}$	two $p_T > 10\,\text{GeV}$ electron candidates, pixel match ≥ 1
photon trigger		
IsoPhoton20 L1R	e/γ candidate with $p_T > 15\,\text{GeV}$	$p_T > 20\,\text{GeV}$ photon candidate, track isolation criterion
Photon25 L1R	e/γ candidate with $p_T > 15\,\text{GeV}$	photon candidate with $p_T > 25\,\text{GeV}$
DoubleIsoPhoton20 L1R	two e/γ candidates with $p_T > 8\,\text{GeV}$	two $E_T > 20\,\text{GeV}$ photon candidates, track isolation criterion
DoubleIsoPhoton20 L1I	two e/γ candidates with $p_T > 10\,\text{GeV}$	two $E_T > 20\,\text{GeV}$ photon candidates, track isolation criterion
ECAL activity trigger		
EM80	e/γ candidate with $p_T > 15\,\text{GeV}$	$E_T > 80\,\text{GeV}$ EM candidate
EM200	e/γ candidate with $p_T > 15\,\text{GeV}$	$E_T > 200\,\text{GeV}$ EM candidate

Table B.7 List of the HLT triggers currently used in the MUSiC analysis

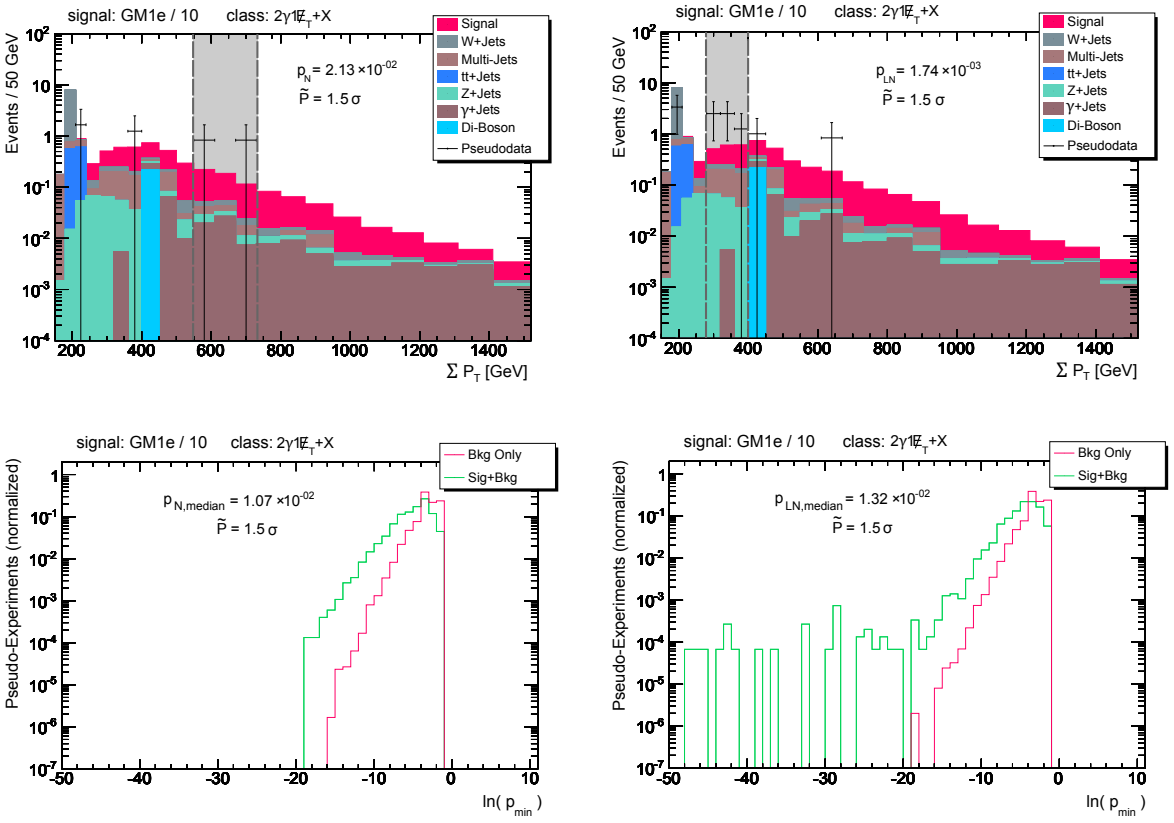

Figure B.10 Evaluation of the inclusive class with at least two photons and MET for the GM1e test scenario with $L = 1\,\mathrm{fb}^{-1}$ at $\sqrt{s} = 10\,\mathrm{TeV}$ in a scenario with GM1e in which the signal cross section was scaled down by a factor of 10. **Up left:** $(S + B)$ pseudo experiment using normal dicing. **Up right:** $(S + B)$ pseudo experiment using lognormal dicing. **Down left:** p_{min} distribution for B (normal dicing, 600000 pseudo experiments) and $(S + B)$ (normal dicing, 15000 pseudo experiments). **Down right:** p_{min} distribution for B (normal dicing, 500000 pseudo experiments) and $B + S$ (lognormal dicing, 15000 pseudo experiments)

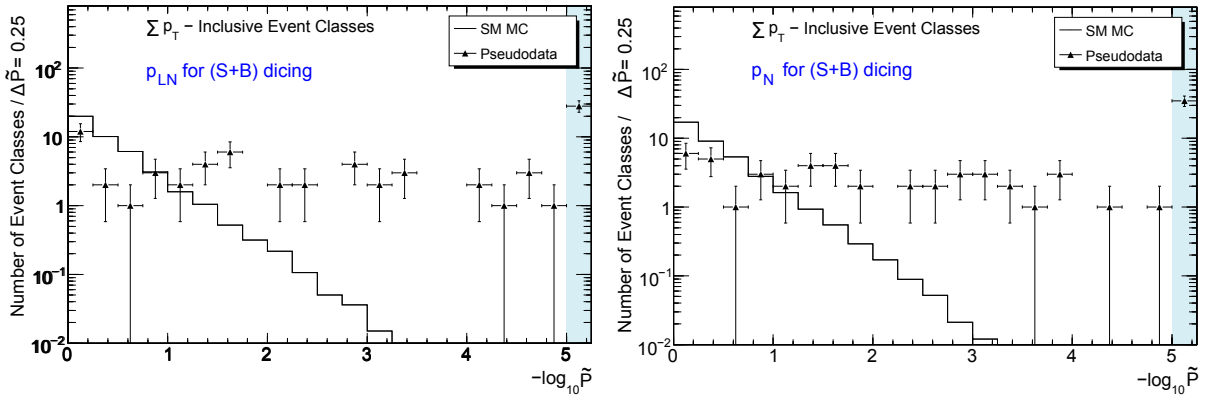

Figure B.11 Left: A global pseudo experiment with signal GM1d using p_{LN} for (S+B) dicing$\widetilde{P}$ — evaluation of the $\sum p_T$ variable in the inclusive classes (all classes diced independently) **Right:** the same but using p_N for (S+B) dicing

n_{on}	Z_{LN}	Z_N	Z_{LN}	Z_N
	$\mu = b_0 = 0.1$, $k = 1.2$, $\sigma_{\mathrm{rel}} = 0.2$		$\mu = b_0 = 0.1$, $k = 1.6$, $\sigma_{\mathrm{rel}} = 0.6$	
1	1.30	1.31	1.25	1.28
2	2.57	2.58	2.46	2.48
3	3.57	3.58	3.36	3.41
5	5.17	5.18	4.75	4.89
7	6.50	6.51	5.82	6.12
10	8.19	8.22	7.10	7.68
	$\mu = b_0 = 1$, $k = 1.2$, $\sigma_{\mathrm{rel}} = 0.2$		$\mu = b_0 = 1$, $k = 1.6$, $\sigma_{\mathrm{rel}} = 0.6$	
1	-	-	-	-
4	2.01	2.02	1.70	1.75
7	3.61	3.63	2.90	3.11
10	4.92	4.96	3.77	4.24
15	6.75	6.84	4.83	5.84
20	8.29	8.46	5.65	7.22
	$\mu = b_0 = 10$, $k = 1.2$, $\sigma_{\mathrm{rel}} = 0.2$		$\mu = b_0 = 10$, $k = 1.6$, $\sigma_{\mathrm{rel}} = 0.6$	
15	1.15	1.18	0.69	0.65
20	2.22	2.27	1.27	1.32
25	3.15	3.25	1.74	1.94
30	3.98	4.16	2.13	2.53
35	4.73	5.01	2.47	3.10
50	6.61	7.73	3.25	4.67
	$\mu = b_0 = 50$, $k = 1.2$, $\sigma_{\mathrm{rel}} = 0.2$		$\mu = b_0 = 50$, $k = 1.6$, $\sigma_{\mathrm{rel}} = 0.6$	
60	0.77	0.78	0.35	0.27
75	1.83	1.91	0.82	0.76
90	2.74	2.98	1.20	1.24
110	3.78	4.33	1.63	1.86
130	4.67	5.60	1.99	2.48
150	5.45	6.82	2.29	3.08

Table B.8 Values for Z_N and Z_{LN} (one-sided) in case of an observed excess for several sets of parameters with a relative error of 20% or 60% on the estimated background (numbers without rounding)

n_{on}	Z_{LN}	Z_N	Z_{LN}	Z_N
	$\mu = b_0 = 20$, $k = 1.2$, $\sigma_{\mathrm{rel}} = 0.2$		$\mu = b_0 = 20$, $k = 1.6$, $\sigma_{\mathrm{rel}} = 0.6$	
0	5.17	4.40	3.83	2.32
3	3.85	3.35	2.62	1.65
5	3.19	2.81	2.10	1.38
9	2.11	1.89	1.31	0.95
13	1.21	1.09	0.72	0.57
17	0.42	0.37	0.24	0.23
	$\mu = b_0 = 50$, $k = 1.2$, $\sigma_{\mathrm{rel}} = 0.2$		$\mu = b_0 = 50$, $k = 1.6$, $\sigma_{\mathrm{rel}} = 0.6$	
0	7.66	5.00	5.20	2.67
1	7.24	4.79	4.76	2.42
5	6.04	4.23	3.63	1.95
10	4.93	3.65	2.82	1.61
22	2.97	2.40	1.55	1.05
35	1.39	1.19	0.68	0.55
	$\mu = b_0 = 100$, $k = 1.2$, $\sigma_{\mathrm{rel}} = 0.2$		$\mu = b_0 = 100$, $k = 1.6$, $\sigma_{\mathrm{rel}} = 0.6$	
5	8.45	4.69	4.92	2.27
10	7.47	4.39	4.11	2.00
15	6.67	4.11	3.53	1.81
30	4.81	3.31	2.36	1.39
50	3.00	2.29	1.39	0.95
65	1.93	1.57	0.87	0.67
	$\mu = b_0 = 300$, $k = 1.2$, $\sigma_{\mathrm{rel}} = 0.2$		$\mu = b_0 = 300$, $k = 1.6$, $\sigma_{\mathrm{rel}} = 0.6$	
3	13.08	5.15	7.95	2.82
15	11.16	4.77	5.83	2.32
45	8.35	4.21	3.83	1.83
95	5.52	3.35	2.38	1.37
120	4.49	2.93	1.90	1.18
160	3.15	2.26	1.31	0.90

Table B.9 Values for Z_N and Z_{LN} (one-sided) in case of an observed deficit for several sets of parameters with a relative error of 20% or 60% on the estimated background (numbers without rounding)

C Signal and Background Samples

Process	cross section (fb)	# events	path in dbs
GMSB GM1b	$2.97 \cdot 10^3$	$1 \cdot 10^5$	/Exotica_GMSB_GM1b/Summer08_IDEAL_V11_redigi_v1
GM1c	$8.43 \cdot 10^2$	$1 \cdot 10^5$	/Exotica_GMSB_GM1c/Summer08_IDEAL_V11_redigi_v1
GM1d	$2.99 \cdot 10^2$	$1 \cdot 10^5$	/Exotica_GMSB_GM1d/Summer08_IDEAL_V11_redigi_v1
GM1e	12.4	$1 \cdot 10^5$	/Exotica_GMSB_GM1e/Summer08_IDEAL_V11_redigi_v1
GM1f	5.82	$1 \cdot 10^5$	/Exotica_GMSB_GM1f/Summer08_IDEAL_V11_redigi_v1
$\mu^\star$ M 200	Λ dependent	$2 \cdot 10^4$	produced for dedicated Exotica analysis
$\mu^\star$ M 400	Λ dependent	$2 \cdot 10^4$	
$\mu^\star$ M 1000	Λ dependent	$2 \cdot 10^4$	
LED Diphoton	$5.0 \cdot 10^2$	$3 \cdot 10^4$	private production with SHERPA $M_f = 2.5$,$n = 4$, 1 jet inclusive

Table C.1 Used signal samples

Process	cross section (fb)	# events	path in dbs
Photon+Jets	$2.89 \cdot 10^8$	$9 \cdot 10^5$	/PhotonJetPt15/Summer08_IDEAL_V12_redigi v1
	$3.22 \cdot 10^7$	$9 \cdot 10^5$	/PhotonJetPt30/Summer08_IDEAL_V12_redigi_v1
	$1.01 \cdot 10^6$	$8 \cdot 10^5$	/PhotonJetPt80/Summer08_IDEAL_V12_redigi_v1
	$5.14 \cdot 10^4$	$9 \cdot 10^5$	/PhotonJetPt170/Summer08_IDEAL_V12_redigi_v1
	$4.19 \cdot 10^3$	$1 \cdot 10^6$	/PhotonJetPt300/Summer08_IDEAL_V12_redigi_v1
	$4.52 \cdot 10^2$	$1 \cdot 10^6$	/PhotonJetPt470/Summer08_IDEAL_V12_redigi_v1
	$2.00 \cdot 10^1$	$1 \cdot 10^6$	/PhotonJetPt800/Summer08_IDEAL_V12_redigi_v1
	0.27	$1 \cdot 10^6$	/PhotonJetPt1400/Summer08_IDEAL_V12_redigi_v1
	$1.5 \cdot 10^{-3}$	$1 \cdot 10^6$	/PhotonJetPt2200/Summer08_IDEAL_V12_redigi_v1
QCD	$1.46 \cdot 10^{12}$	$7 \cdot 10^6$	/QCDpt15/Summer08_IDEAL_V11_redigi_v3
	$1.09 \cdot 10^{11}$	$3 \cdot 10^6$	/QCDpt30/Summer08_IDEAL_V11_redigi_v1
	$1.93 \cdot 10^9$	$3 \cdot 10^6$	/QCDpt80/Summer08_IDEAL_V11_redigi_v1
	$6.26 \cdot 10^7$	$3 \cdot 10^6$	/QCDpt170/Summer08_IDEAL_V11_redigi_v1
	$3.66 \cdot 10^6$	$3 \cdot 10^6$	/QCDpt300/Summer08_IDEAL_V11_redigi_v1
	$3.16 \cdot 10^5$	$3 \cdot 10^6$	/QCDpt470/Summer08_IDEAL_V11_redigi_v1
	$1.19 \cdot 10^4$	$3 \cdot 10^6$	/QCDpt800/Summer08_IDEAL_V11_redigi_v2
	$1.72 \cdot 10^2$	$5 \cdot 10^5$	/QCDpt1400/Summer08_IDEAL_V11_redigi_v1
	1.42	$2 \cdot 10^6$	/QCDpt2200/Summer08_IDEAL_V11_redigi_v2
	$8.60 \cdot 10^{-3}$	$5 \cdot 10^5$	/QCDpt3000/Summer08_IDEAL_V11_redigi_v1
Z+Jets	$3.7 \cdot 10^6$	$1 \cdot 10^6$	/ZJets-madgraph/Summer08_IDEAL_V11_redigi_v1
Z+Jets FastSim	$3.7 \cdot 10^6$	$1 \cdot 10^7$	/Zjets-madgraph/Winter09_IDEAL_V11_FastSim_v1
W+Jets	$4.0 \cdot 10^7$	$9 \cdot 10^6$	/WJets-madgraph/Summer08_IDEAL_V11_redigi_v1
TT+Jets	$3.17 \cdot 10^5$	$1 \cdot 10^6$	/TTJets-madgraph/Fall08_IDEAL_V11_redigi_v10
W+2 Photons	10.4	$1 \cdot 10^5$	/Wgg-madgraph/Fall08_IDEAL_V11_redigi_v1
Z+2 Photons	5.1	$1 \cdot 10^5$	/Zgg-madgraph/Fall08_IDEAL_V11_redigi_v1
WW inclusive	$4.48 \cdot 10^4$	$2 \cdot 10^5$	/WW/Summer08_IDEAL_V11_redigi_v1
ZZ inclusive	$7.1 \cdot 10^3$	$2 \cdot 10^5$	/ZZ/Summer08_IDEAL_V11_redigi_v1
WZ inclusive	$1.74 \cdot 10^4$	$2 \cdot 10^5$	/WZ_incl/Summer08_IDEAL_V11_redigi_v1
DrellYan $\mu\mu$	$1.10 \cdot 10^3$	$1 \cdot 10^4$	/DYmumuM200/Summer08_IDEAL_V11_redigi_v2
	44.88	$1 \cdot 10^4$	/DYmumuM500/Summer08_IDEAL_V11_redigi_v2
	2.55	$1 \cdot 10^4$	/DYmumuM1000/Summer08_IDEAL_V11_redigi_v2
	$5.58 \cdot 10^{-2}$	$1 \cdot 10^4$	/DYmumuM2000/Summer08_IDEAL_V11_redigi_v2
W $e\nu$		$5 \cdot 10^4$	special samples from the Wprime Working Group
SingleTop t	$5.53 \cdot 10^4$	$3 \cdot 10^5$	/SingleTop_tChannel/Summer08_IDEAL_V11_redigi_v3
SingleTop tW	$2.73 \cdot 10^4$	$2 \cdot 10^5$	/SingleTop_tWChannel/Summer08_IDEAL_V11_redigi_v3
SingleTop s	$1.66 \cdot 10^3$	$1 \cdot 10^4$	/SingleTop_sChannel/Summer08_IDEAL_V11_redigi_v3
QCD Diphoton	$2.6 \cdot 10^2$	$3 \cdot 10^4$	private production with SHERPA E_T 100 to 400 GeV
	$9.6 \cdot 10^3$	$3 \cdot 10^4$	private production with SHERPA E_T 300 to 10000 GeV

Table C.2 Applied RECO background samples (mainly from Summer08 and Fall08 production)

D Units and Conventions

Throughout this work units are mainly given in the SI system or in natural units. While in the SI system the constants $\hbar$ and c are

$$\hbar = 1.055 \cdot 10^{-34}\,\mathrm{Js} \quad \text{and} \quad \mathrm{c} = 2.998 \cdot 10^{8}\,\frac{\mathrm{m}}{\mathrm{s}} \quad , \tag{D.1}$$

they are set equal to 1 in natural units:

$$\hbar \equiv c \equiv 1 \quad . \tag{D.2}$$

This implies that mass, inverse length, and inverse time have the dimension of energy. In particle physics energies are usually stated in terms of eV. This energy unit is related to energies in the SI system via

$$1\,\mathrm{eV} = 1.60 \cdot 10^{-34}\,\mathrm{J} \quad . \tag{D.3}$$

Cross sections are given using the unit convention

$$1\,\mathrm{b} = 1 \cdot 10^{-28}\,\mathrm{m}^2 \quad . \tag{D.4}$$

Equal indices denote implicit summation. For example the product of two Lorentz 4-vectors is given by

$$a^\mu b_\mu = \sum_\mu a^\mu b_\mu = \sum_\mu a_\mu\, g_{\mu\mu}\, b_\mu = a_0\, b_0 - \vec{a}\,\vec{b} \quad . \tag{D.5}$$

The γ-matrices are used in the standard representation

$$\gamma^0 = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix} \quad , \quad \gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3 = \begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix} \tag{D.6}$$

and

$$\gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix} \quad , i = 1 \ldots 3 \tag{D.7}$$

with the Pauli matrices σ^i (see e.g. [18]).

Bibliography

[1] R. Penrose. *The Road to Reality: A Complete Guide to the Laws of the Universe*. Vintage, 2005.

[2] P. Biallass, T. Hebbeker, C. Hof, A. Meyer, and S. A. Schmitz. MUSiC - An Automated Scan for Deviations between Data and Monte Carlo Simulation, 2008. CMS AN-2008/065.

[3] T. Hebbeker. A Global Comparison between L3 Data and Standard Model Monte Carlo - a first attempt, 1998. L3 note 2305 `http://web.physik.rwth-aachen.de/~hebbeker/l3note_2305.pdf`.

[4] A. Aktas and others (H1 Collaboration). A general search for new phenomena in ep scattering at HERA. *Physics Letters B*, 602:14, 2004. online version: arXiv:hep-ex/0408044v1.

[5] T. Aaltonen and others (CDF collaboration). Model-Independent and Quasi-Model-Independent Search for New Physics at CDF. *Physical Review D*, 78(012002), 2008. online version: arXiv:0712.1311v2 [hep-ex].

[6] T. Aaltonen and others (CDF collaboration). Global Search for New Physics with 2.0 fb^{-1} at CDF. *Physical Review D*, 79(011101), 2009. online version: arXiv:0809.3781v2 [hep-ex].

[7] The H1 Collaboration. A General Search for New Phenomena at HERA, 2009. DESY-08-173 , online version: arXiv:0901.0507v1 [hep-ex].

[8] K. Zuber. *Neutrino Physics*. IOP Publishing, 2004.

[9] W. Bernreuther. Skript zur Vorlesung Quantenfeldtheorie 2, 2008. RWTH Aachen University.

[10] published licensed under Creative Commons Attribution 3.0 Unported License by MissMJ `http://commons.wikimedia.org/wiki/File:Standard_Model_of_Elementary_Particles.svg`.

[11] W. Buchmüller and C. Lüdeling. Field Theory and the Standard Model, 2006. arXiv:hep-ph/0609174v1.

[12] J. L. Rosner. The Standard Model in 2001, 2006. arXiv:hep-ph/0609174v1.

[13] W. Bernreuther. Skript zur Vorlesung Quantenfeldtheorie 1, 2008. RWTH Aachen University.

[14] M. E. Peskin and D. V. Schroeder. *An Introduction to Quantum Field Theory*. Westview Press, 1995.

[15] S. Weinberg. *The Quantum Theory of Fields: Volume 1, Foundations*. Elsevier, 2005.

[16] C. Berger. *Elementarteilchenphysik*. Springer, 2nd edition, 2006.

[17] P. Langacker. Introduction to the Standard Model and Electroweak Physics, 2009. arXiv:hep-ph/09010241v1.

[18] F. Halzen and A. D. Martin. *Quarks and Leptons: An Introductory Course in Modern Particle Physics*. John Wiley & Sons, 1984.

[19] A. Signer. ABC of SUSY. *J.Phys.G*, 36(073002), 2009. online version: arXiv:0905.4630v1 [hep-ph].

[20] S. P. Martin. A Supersymmetry Primer, 1997. arXiv:hep-ph/9709356v5 (2008).

[21] I. J. R. Aitchison. Supersymmetry and the MSSM: An Elementary Introduction, 2005. arXiv:hep-ph/0505105v1.

[22] M. Drees. An Introduction to Supersymmetry, 1996. arXiv:hep-ph/9611409v1.

[23] C. Amsler and others (Particle Data Group). Review of Particle Physics. *Physics Letters B*, 667:1, 2008. and 2009 partial update for the 2010 edition `http://pdg.lbl.gov`.

[24] G. F. Giudice and R. Rattazzi. Theories with Gauge-Mediated Supersymmetry Breaking, 1998. arXiv:hep-ph/9801271v2 (1999).

[25] B. C. Allanach et al. The Snowmass Points and Slopes: Benchmarks for SUSY Searches. *Eur.Phys.J.C*, 25:113–123, 2002. online version: arXiv:hep-ph/0202233v1.

[26] F. Santanastasio. *Search for Supersymmetry with Gauge-Mediated Breaking using high energy photons at CMS experiment*. PhD thesis, Sapienza Universita di Roma, 2007. available online at `http://www.roma1.infn.it/cms/tesiPHD/santanastasio.pdf`.

[27] A. Belyaev et al. Production and Decay of Excited Electrons at the LHC. *Eur.Phys.J. C*, 32S2:1–18, 2004. online version arXiv:hep-ph/0212006v2.

[28] R. Sekhar Chivukula. Lectures on Technicolor and Compositeness, 2000. arXiv:hep-ph/0011264v1.

[29] U. Baur, M. Spira, and P. M. Zerwas. Excited-quark and -lepton production at hadron colliders. *Physical Review D*, 42(3):815–824, 1990.

[30] O. J. P. Éboli, S. M. Lietti, and P. Mathews. Excited leptons at the CERN Large Hadron Collider. *Physical Review D*, 65(075003), 2002.

[31] J. W. Coenen. Search for excited muons in $p\bar{p}$ collision at $\sqrt{s} = 1.96\,\text{TeV}$, 2005. Diplomarbeit `http://web.physik.rwth-aachen.de/~hebbeker/theses/coenen_diploma.pdf`.

[32] V. M. Abazov and others (DØ collaboration). Search for excited muons in $p\bar{p}$ collision at $\sqrt{s} = 1.96\,\text{TeV}$. *Physical Review D*, 73(111102(R)), 2006.

[33] C. Ewerz. Extra Dimensions, 06/07. Lecture course at Heidelberg University `http://www.thphys.uni-heidelberg.de/~ewerz/scriptum/dimensions_ce.pdf`.

[34] R. Gabadadze. ICTP Lectures on Large Extra Dimensions, 2003. CERN-TH/2003-157, arXiv:hep-ph/0308112v1.

[35] G. D. Kribs. TASI 2004 Lectures on the Phenomenology of Extra Dimensions, 2006. arXiv:hep-ph/0605325v1.

[36] E. G. Adelberger et al. Tests of the Gravitational Inverse-Square Law below the Dark-Energy Length Scale. *Phys.Rev.Lett.*, 98(021101), 2007. online version: arXiv:hep-ph/0611184v1.

[37] N. Arkani-Hamed, S. Dimopoulos, and G. Dvali. The hierarchy problem and new dimensions at a millimeter. *Phys.Lett.B*, 429:263–272, 1998. online version: arXiv:hep-ph/9803315v14.

[38] R. Rattazzi. Cargese Lectures on Extra Dimensions, 2006. CERN-PH-TH/2006-029 arXiv:hep-ph/0607055v1.

[39] G. F. Giudice, R. Rattazzi, and J. D. Wells. Quantum gravity and extra dimensions at high-energy colliders. *Nucl.Phys. B*, 544:3–38, 1999. online version: arXiv:hep-ph/9811291v2.

[40] G. Landsberg. Collider Searches for Extra Dimensions, 2004. arXiv:hep-ex/0412028v2.

[41] Technological Challenges for the LHC, 2003. talks given at the CERN Academic Training Lecture Regular Programme `http://indico.cern.ch/conferenceDisplay.py?confId=a032516`.

[42] L. Evans. The Large Hadron Collider. *New Journal of Physics*, 9(335), 2007.

[43] display of LHC cooling status. (hardware comissioning coordination) `http://hcc.web.cern.ch/hcc/field.php`.

[44] D. Green. *The Physics of Particle Detectors*. Cambridge University Press, 2000.

[45] C. Grupen and B. Shwartz. *Particle Detectors*. Cambridge University Press, 2nd edition, 2008.

[46] T. Fließbach. *Allgemeine Relativitätstheorie*. Elsevier, 5th edition, 2006.

[47] The CMS Collaboration. The CMS experiment at the CERN LHC. *JINST*, 3(S08004), 2008.

[48] T. Hebbeker. Skript zur Vorlesung Elementarteilchenphysik 1, 2007. RWTH Aachen University.

[49] The CMS Collaboration. The CMS electromagnetic calorimeter project : Technical Design Report. 1997. CERN-LHCC-97-033, online version at: `http://cmsdoc.cern.ch/ftp/TDR/ECAL/ecal.html`.

[50] C. I. Rovelli. *The CMS electromagnetic calorimeter and the search for the Higgs boson in the decay channel* $H \rightarrow WW^{(*)} \rightarrow 2e2\nu$. PhD thesis, École Polytechnique, 2006. available online at `http://www.imprimerie.polytechnique.fr/Theses/Files/Rovelli.pdf`.

[51] E. Tournefier. The Preshower Detector of CMS at LHC. 2000. CMS Conference Report, CMS CR 2000/010.

[52] The CMS Collaboration. The CMS Collaboration. CMS Physics TDR Volume I: "Detector Performance and Software". 2006. CERN-LHCC-2006-001, online version at: `http://cmsdoc.cern.ch/cms/cpt/tdr/`.

[53] The CMS Collaboration. The hadron calorimeter project: technical design report. 1997. CERN-LHCC-97-031, online version at: `http://cmsdoc.cern.ch/ftp/TDR/HCAL/hcal.html`.

[54] V. V. Abramavov and others (The CMS-HCAL collaboration). Studies of the Response of the Prototype CMS Hadron Calorimeter, Including Magnetic Field Effects, to Pion, Electron, and Muon Beams. *Nucl.Instrum.Meth.A*, 457:75–100, 2001. online version: arXiv:hep-ex/0007045v1.

[55] The CMS Collaboration. The CMS magnet project : Technical Design Report, 1997. CERN-LHCC-97-031, online version at: http://cmsdoc.cern.ch/ftp/TDR/MAGNET/magnet.html.

[56] The CMS Collaboration. The Tracker Project: Technical Design Report, 1998. CERN/LHCC 98-6, online version at: http://cmsdoc.cern.ch/ftp/TDR/TRACKER/tracker.html.

[57] M. Wingham. *Commissioning of the CMS tracker and preparing for early physics at the LHC*. PhD thesis, Imperial College London, 2008. available online at http://cdsweb.cern.ch/record/1133150?ln=eng.

[58] R. Adolphi. *Construction and Calibration of the Laser Alignment System for the CMS Tracker*. PhD thesis, RWTH Aachen University, 2006. available online at http://accms04.physik.rwth-aachen.de/~schael/CMS_Thesis.html.

[59] F. Simon. Teilchendetektoren in der hochenergiephysik, 2009. lectures given at the "41. Herbstschule für Hochenergiephysik Maria Laach" http://maria-laach.physik.uni-siegen.de/Folien/2009/Simon/ML_TeilchenDetektoren.pdf.

[60] P. (on behalf of the CMS Collaboration) Azzuri. Track Reconstruction Performance in CMS, 2008. online version: arXiv:0812.5036v1 [physics.ins-det].

[61] M. Atac et al. Beam test results of the US-CMS forward pixel detector. *Nuclear Instruments and Methods in Physics Research A*, 488:271–281, 2002.

[62] F.. Sauli. Principles of Operation of Multiwire Proportional and Drift Chambers, 1977. CERN Report 77-09.

[63] D. Teyssier. CMS muon system performance, 2009. CMS CR -2009/101.

[64] P. A. Biallass. *Commissioning of the CMS Muon Detector and Development of Generic Search Strategies for New Physics*. PhD thesis, RWTH Aachen University, 2009. available online at http://web.physik.rwth-aachen.de/~hebbeker/theses/biallass_phd.pdf.

[65] H. Schwarthoff. *Simulationen in Konzeption und Bau der zentralen Myondriftkammern am CMS-Detektor*. PhD thesis, RWTH Aachen University, 1997.

[66] The CMS Collaboration. The Muon Project: Technical Design Report, 1997. CERN/LHCC 97-32 http://cms.cern.ch/iCMS/jsp/page.jsp?mode=cms&action=url&urlkey=CMS_TDRS.

[67] The CMS Collaboration. CMS trigger and data-acquisition project : Technical Design Report, Volume I: The Level-1 Trigger, 2000. CERN/LHCC 2000-038.

[68] C. Foudas. The CMS Level-1 Trigger at LHC and Super-LHC, 2008. presented at ICHEP08, arXiv:0810.4133v1 [physics.ins-det].

[69] The CMS Collaboration. CMS trigger and data-acquisition project, Volume II: Technical Design Report : Data Acquisition & High-Level Trigger, 2002. CERN/LHCC 2002-26.

[70] S. Agostinelli et al. Geant4 - a simulation toolkit. *Nuclear Instruments and Methods in Physics Research A*, 506:250–303, 2003.

[71] T. Gleisberg et al. SHERPA 1.α, a proof-of-concept version. *JHEP*, 065(2), 2004.

[72] http://doc.cern.ch//archive/electronic/cern/others/PHO/photo-cms/gen/gen-2007-002.pdf.

[73] W. Adam et al. PAT: the CMS Physics Analysis Toolkit, 2009. CMS CR -2009/083.

[74] PAT twiki, 2009. `https://twiki.cern.ch/twiki/bin/view/CMS/SWGuidePAT`, revision 23.

[75] V. Silva. *Grid Computing for Developers*. Charles River Media, 1st edition, 2006.

[76] The CMS Collaboration. CMS Technical Design Report : The Computing Project, 2005. CERN/LHCC 2002-26, available online at `http://cdsweb.cern.ch/record/838359/files/lhcc-2005-023.pdf`.

[77] D. Spiga et al. The CMS Remote Analysis Builder (CRAB). In *Procs. of the XI Int.W. on Advanced Computing and Analysis Techniques in Physics Research (PoS ACAT07)*, pages 580–586, 2007.

[78] G. Codispoti. Distributed analysis with CRAB: the client-server architecture evolution and commissioning, 2008. talk given at ACAT 2008, slides online at `http://indico.cern.ch/materialDisplay.py?contribId=34&sessionId=4&materialId=slides&confId=34666`.

[79] O. Actis, M. Erdmann, R. Fischer, A. Hinzmann, M. Kirsch, T. Klimkovich, G. Mueller, M. Plum, and J. Steggemann. Visual Physics Analysis (VISPA) - Concepts and First Applications, 2008. arXiv:0810.3609v1[physics.data-an].

[80] O. Actis et al. Physics eXtension Library (PXL) User's Guide, Version 1.0, 2008. available online at http://pxl.sourceforge.net.

[81] ROOT webpage. `http://root.cern.ch/drupal/`.

[82] ROOT An Object-Oriented Data Analysis Framework Users Guide 5.24, 2009. `http://root.cern.ch/drupal/content/users-guide`.

[83] G. Welch and G. Bishop. An Introduction to the Kalman Filter, 2005. `http://www.cs.unc.edu/~welch/kalman/kalmanIntro.html`.

[84] R. Frühwirth, D. Regler, R.K. Bock, H. Grote, and D. Notz. *Data Analysis Techniques for High-Energy Physics*. Cambridge University Press, 2nd edition, 2000.

[85] R. Frühwirth. Application of Kalman Filtering to Track and Vertex Fitting. *Nucl.Instrum.Meth.A*, 262:444–450, 1987.

[86] W. Adam, R. Frühwirth, A. Strandlie, and T. Todorov. Reconstruction of electrons with the Gaussian-sum filter in the CMS tracker at LHC, 2003. Talk from the 2003 Computing in High Energy and Nuclear Physics (CHEP03), online version: arXiv:physics/0306087v1 [physics.data-an].

[87] W. Adam, R. Frühwirth, A. Strandlie, and T. Todorov. Reconstruction of electrons with the Gaussian-sum filter in the CMS tracker at LHC. *J. Phys. G: Nucl. Part. Phys*, 31:N9–N20, 2005.

[88] R. Frühwirth and S. Frühwirth-Schnatter. On the treatment of energy loss in track fitting. *Computer Physics Communications*, 110:80–86, 1998.

[89] S. Baffioni et al. Electron Reconstruction in CMS, 2006. CMS NOTE 2006/040.

[90] F. James. *Statistical Methods in Experimental Physics*. World Scientific Pulishing, 2nd edition, 2006.

[91] N. Adam et al. Electron Reconstruction at Low p_t, 2009. CMS AN -2009/074.

[92] G. Abbiendi et al. Muon Reconstruction in the CMS Detector, 2008. CMS AN-2008/97.

[93] G. P. Salam and G. Soyez. A practical Seedless Infrared-Safe Cone jet algorithm. *JHEP*, 05:086, 2007. online version: arXiv:0704.0292v2 [hep-ph].

[94] D. Acosta and others (CDF collaboration). Study of Jet Shapes in Inclusive Jet Production in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96tev$. *Physical Review D*, 71(112002), 2005. online version: arXiv:hep-ex/0505013v2.

[95] A. Bhatti et al. Performance of the SISCone Jet Clustering Algorithm, 2008. CMS AN -2008/002.

[96] G. C. Blazey et al. Run II Jet Physics: Proceedings of the Run II QCD and Weak Boson Physics Workshop, 2000. arXiv:hep-ex/0005012v2.

[97] S. Esen et al. $\not\!\!E_t$ Performance in CMS, 2008. CMS AN -2007/041.

[98] G. Landsberg and F. Moortgat. MET Reconstruction Performance and Validation, 2008. CMS AN -2008/089.

[99] J. Branson et al. A cut based method for electron identification in CMS, 2008. CMS AN-2008/082.

[100] O. Charaf et al. Electron ID at High Energies, 2008. CMS NOTE 2008/45v2.

[101] N. Amapane et al. Muon Identification in CMS, 2008. CMS AN-2008/098.

[102] A. Bhatti et al. Plans for Jet Energy Corrections at CMS, 2007. CMS AN-2007/055.

[103] M. Wessels. *General Search for New Phenomena in ep Scattering at HERA*. PhD thesis, RWTH Aachen University, 2004. available online at `http://mozart.physik.rwth-aachen.de/dipldrcd/martin.wessels.dissertation.pdf`.

[104] T. Aaltonen and others (CDF collaboration). Model-Independent and Quasi-Model-Independent Search for New Physics at CDF. *Physical Review D*, 78(012002), 2008. online version: arXiv:0712.1311v2 [hep-ex].

[105] G. Choudlakis. *Model Independent Search For New Physics At The Tevatron*. PhD thesis, Massachusetts Institute of Technology, 2008. arXiv:0805.3954v1 [hep-ex].

[106] B. Abbott and others (D$\emptyset$ collaboration). Search for new physics in $e\mu X$ data at D$\emptyset$ using SLEUTH: A quasi-model-independent search strategy for new physics. *Physical Review D*, 62(092004), 2000. online version: arXiv:hep-ex/0006011v2.

[107] E. L. Lehmann. The Fisher,Neyman-Pearson Theories of Testing Hypothesis: One Theory or Two? *Journal of the American Statistical Association*, 88(424):1242–1249, 1993.

[108] M. J. Bayarri, J. O. Berger, and T. Sellke. Calibration of P-Values for Testing Precise Null Hypotheses. *The American Statistician*, 55(1):62–71, 2001.

[109] V. Blobel and E. Lohrmann. *Statistische und numerische Methoden der Datenanalyse*. Teubner, Leipzig, 1998.

[110] L. Demortier. P Values and Nuisance Parameters. In H. B. Prosper, L. Lyons, and A. De Roeck, editors, *PHYSTAT LHC Workshop on Statistical Issues for LHC Physics*, pages 23–41, 2008. (PHYSTAT 2007) `http://phystat-lhc.web.cern.ch/phystat-lhc/`.

[111] L. Demortier. P Values: What They Are and How to Use Them, 2006. version 3.00, available online at `http://www-cdf.fnal.gov/~luc/statistics/cdf0000.ps`.

[112] F. James. overview of principles of statistics. In L. Lyons and M. R. Whalley, editors, *Advanced statistical techniques in particle physics : Proceedings*, 2002.

[113] James O. Berger and Thomas Sellke. Testing a Point Null Hypothesis: The Irreconcilability of P Values and Evidence. *Journal of the American Statistical Association*, 82(397):112–122, 1987.

[114] M.J. Bayarri, T. Sellke, and J. O. Berger. Calibration of P-values for Testing Precise Null Hypotheses. *The American Statistician*, 55(1):62–71, 2001. an online version is available at: `http://www.stat.duke.edu/~berger/papers/99-13b.ps`.

[115] M. J. Bayarri and Hubbard R. Confusion Over Measures of Evidence (p's) Versus Errors (α's) in Classical Statistical Testing. *The American Statistician*, 57(3):171–178, 2003.

[116] V. Bartsch and G. Quast. Expected signal observability at future experiments, 2005. CMS NOTE 2005/004.

[117] Robert D. Cousins, James T. Linnemann, and Jordan Tucker. Evaluation of three methods for calculating statistical significance when incorporating a systematic uncertainty into a test of the background-only hypothesis for a Poisson process. *Nuclear Instruments and Methods in Physics Research A*, 595:480–501, 2008. arXiv:physics/0702156v4 [physics.data-an].

[118] J. Conrad and F. Tegenfeldt. On Bayesian treatment of systematic uncertainties in confidence interval calculation. *Nucl.Instrum.Meth. A*, 539(1-2):407–413, 2004. online version: arXiv:physics/0408039v2 [physics.data-an].

[119] B. Knuteson. Systematic analysis of HEP collider data. *Nuclear Instruments and Methods in Physics Research*, 534:7–14, 2004.

[120] N. L. Johnson and S. Kotz. *Continuous Univariate Distributions -1.* Houghton Mifflin, 1970.

[121] M. K. Simon. *Probability Distributions Involving Gaussian Random Variables: A Handbook for Engineers and Scientists.* Springer, 2002.

[122] S. Abdullin et al. Search Strategy for the Standard Model Higgs Boson in the $H \rightarrow ZZ^{(*)} \rightarrow 4\mu$ Decay Channel using M (4μ)-Dependent Cuts, 2006. CMS NOTE 2006/122.

[123] A. Korytov. private communications, 2009.

[124] J. Aitchison and J. A. C. Brown. *The Lognormal Distribution.* Cambridge University Press, 1957. reprinted 1981.

[125] Mehta N. et al. Approximating the Sum of Correlated Lognormal or Lognormal-Rice Random Variables, 2006. Mitsubishi Electric Research Laboratories TR2006-010 `http://www.merl.com/reports/docs/TR2006-010.pdf`.

[126] A. Safak. Statistical Analysis of the Power Sum of Multiple Correlated Log-Normal Components. *IEEE Transactions on Vehicular Technology*, 42(1):58–61, 1993.

[127] M. W. Krasny, J. Chwastowski, and K. Slowikowski. Luminosity measurement method for LHC: The theoretical precision and the experimental challenges. *Nucl.Instrum.Meth. A*, Volume 584(1):42–52, 2008. online version: arXiv:hep-ex/0610052v1.

[128] The CMS Collaboration. The CMS Collaboration. CMS Physics TDR Volume II: "Physics Performance". *J. Phys. G: Nucl. Part. Phys.*, Volume 584(34):995–1579, 2007. online version at: `http://cmsdoc.cern.ch/cms/cpt/tdr/`.

[129] R. S. Thorne. The role of uncertainties in parton distribution functions. In *PHYSTAT LHC Workshop on Statistical Issues for LHC Physics Proceedings*, 2007. (PHYSTAT 2007) online version: arXiv:0711.2986v1 [hep-ph].

[130] J. Pumplin and others (CDF collaboration). Uncertainties of predictions from parton distribution functions II: the Hessian method. *Physical Review D*, 65(014013), 2001. online version: arXiv:hep-ph/0101032v3.

[131] A. M. Cooper-Sarkar, C. Gwenlan, and A. Tricoli. Uncertainties on W and Z production at the LHC. In *HERA@LHC workshop 2005 Proceedings*, 2005. online version: arXiv:hep-ex/0509002v1.

[132] P. Biallass, T. Hebbeker, C. Hof, A. Meyer, and H. Pieta. Parton Distribution Uncertainty Determination within CMSSW, 2009. CMS AN-2009/048.

[133] D. Stump et al. Inclusive Jet Production, Parton Distributions, and the Search for New Physics. *JHEP*, 0310:046, 2003. online version: arxiv:hep-ph/0303013.

[134] C. Hof. *Implementation of a Model-Independent Search for New Physics with the CMS Detector exploiting the World-Wide LHC Computing Grid.* PhD thesis, RWTH Aachen University, 2009. available online at `http://web.physik.rwth-aachen.de/~hebbeker/theses`.

[135] D. Acosta et al. Measuring Muon Reconstruction Efficiency from Data, 2006. CMS NOTE 2006/060.

[136] W. Andrews et al. Data-driven methods to estimate the electron and muon fake contributions to lepton analyses, 2009. CMS AN -2009/041.

[137] N. Marinelli. Track finding in gamma conversions in CMS. In *10th International Conference on Advanced Technology and Particle Physics (ICATPP 07)*, 2007. online version: arXiv:0710.2818v1.

[138] V. Büge et al. Calibration of the absolute jet energy scale with $Z(\rightarrow \mu^+\mu^-) + \text{jet}$ events at CMS, 2009. CMS AN 2009/057.

[139] R. Harris and K. Kousouris. Determination of the Relative Jet Energy Scale at CMS from Dijet Balance., 2008. CMS AN -2008/031.

[140] D. del Re and M. Voutilainen. Jet energy calibration with photon+jet events, 2009. CMS AN -2009/012.

[141] H. Pieta. private communications, 2009.

[142] R. D. Cousins. private communications, 2009.

[143] TMinuit. `http://root.cern.ch/root/html/TMinuit.html`.

[144] The CDF collaboration. Setting Limits on GMSB Models in the $\gamma\gamma + \not\!\!E_t$ Final State at CDF, June 2009. CDF/PUB/CDFR/PUBLIC/9625 Version 2.0 `http://www-cdf.fnal.gov/physics/exotic`.

[145] D. Kovalskyi et al. Fireworks: A Physics Event Display for CMS, 2007. CMS CR -2009/111, presented at CHEP09, Fireworks online workbook: `https://twiki.cern.ch/twiki/bin/view/CMS/WorkBookFireworks`.

[146] S. Bhattacharya et al. Study of Direct Photon plus Jet production in cms experiment at $\sqrt{s} = 14\,\text{TeV}$. *Eur.Phys.J.C*, 53:49–58, 2008. online version: arXiv:0705.2740v4 [hep-ex].

[147] I. Akin et al. Search for Scalar and Tensor Unparticles in the Diphoton Final State, 2009. CMS AN -2009/096.

[148] C. Zeidler. Studie zum Nachweis angeregter Myonen mit dem CMS-Detektor, 2006. Diplomarbeit `http://web.physik.rwth-aachen.de/~hebbeker/theses/zeidler_diploma.pdf`.

[149] K. Padeken. Studie von angeregten Myonen im CMS-Experiment, 2009. bachelor thesis.

[150] L. Benucci et al. Search for Mono-Jet from ADD Extra-Dimensions at $\sqrt{s} = 10\,\text{TeV}$, 2009. CMS AN -2009/109.

[151] L. Benucci. Searching for extra-dimensions at CMS, 2009. CMS CR -2009/033.

[152] M. Ata. Monte Carlo Studies of Models with Extra Dimensions in CMS, 2009. Diplomarbeit (in preparation) `http://web.physik.rwth-aachen.de/~hebbeker/theses`.

[153] T. Han, J. D. Lykken, and R. Zhang. On Kaluza-Klein States from Large Extra Dimensions. *Physical Review D*, 59(105006), 1999. online version: arXiv:hep-ph/9811350v4.

[154] J. P. Chou et al. Search for Large Extra Dimensions in the Diphoton Final State, 2009. CMS AN -2009/062.

[155] V. M. Abazov and others (D$\emptyset$ collaboration). Search for large extra spatial dimensions in the dielectron and diphoton channels in $p\bar{p}$ collisions at $\sqrt{s} = 1.96\,\text{TeV}$. *Phys. Rev. Lett.*, 102, 2009. online version: arXiv:0809.2813v1 [hep-ex].

[156] W. Shakespeare. *William Shakespeare - The Complete Works*. Oxford University Press, 1988. A Midsummer Night's Dream, Act 5, Epilogue, 14.

Acknowledgments and Closing Quote

First and foremost I want to thank my father and my mother for their continuous support. They are the best parents I could ever have wished for.

In the direct context of this work I would like to begin by thanking Prof. Thomas Hebbeker for assigning me to a fascinating topic and offering help and advice whenever they were needed. Thanks to Prof. Martin Erdmann for agreeing to be the second referee for this work.

Special thanks go also to Philipp Biallass and Carsten Hof, developers of the MUSiC project from its start, who were always supportive and helped me to gain ground in the field of model independent search. They have been an inspiration to work with.

Moreover I would like to thank my office mates Walter Bender and Holger Pieta for creating a great working atmosphere, offering their valuable opinions and for helping me with computer related challenges whenever necessary. Many thanks also to the all those people, inside and outside of our local CMS working group, who helped with many pieces of relevant advice and frequent fruitful discussions. In this context I would like to mention by name Metin Ata, Michael Bontenackels, Jens Frangenheim, Kerstin Hoepfner, Hendrik Jansen, Carsten Magass, Markus Merschmeyer, Arnd Meyer, Paul Papacz, Daniel Teyssier and Clemens Zeidler.

Regarding activities not directly related to physics during the past years of studying I would like to thank my piano teacher Ralph Rotzoll and iaido sensei Michael Moritz for their patient instructions, which have been an a source of recreation and creative momentum. I would also like to thank Martin Hermus for many nice evenings with sometimes serious, sometimes purely entertaining discussions during the last years. Finally I would like to thank Mr. and Mrs. Georgiou for their great hospitality during my Erasmus semester at Queen Mary College, London.

It seems like the final words of theses are often written late at night. At least this holds true for this work. Being rather tired I would have almost forgotten that there is a common urge among physics students writing their thesis to include at least one quote which suggests that they might also have a bit of knowledge of non-scientific literature or even poetry (though it may often be safe to assume questionable profundity). I have considered borrowing from the poem "The Choice" by W. B. Yeates but such humor might be misunderstood too easily. Instead, let me avoid anything complicated or pretentious. "So, good night unto you all." [156]

Selbständigkeitserklärung

Hiermit versichere ich, die Arbeit selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt sowie Zitate kenntlich gemacht zu haben.