The Dissertation Committee for Yuan-Pao Yang certifies that this is the approved version of the following dissertation:

Techniques for Discovery and Mass Measurement Using the Boundary of Many-Body Phase Space

Committee:

Can Kılıç, Supervisor

Arno Bohm

Volker Bromm

Jacques Distler

Sonia Paban

Techniques for Discovery and Mass Measurement Using the Boundary of Many-Body Phase Space

by

Yuan-Pao Yang,

Dissertation
Presented to the Faculty of the Graduate School of
The University of Texas at Austin
in Partial Fulfillment
of the Requirements
for the degree of
Doctor of Philosophy

The University of Texas at Austin
August, 2019

Dedicated to Jessie Lin, who joined me, and Chun-Mu Liang, who left me behind,
along this journey

Acknowledgement

Along this journey, I have received blessing from people more than I ever thought. On the day I left Taiwan, the Californian gentleman next to me was amazed by the decision to study in a country without family and friends. But now, the United States has turned into home and the land of friends for me.

First of all, I would like to thank my adviser, Can Kılıç, for the guidance and mentorship during the past six years as well as the encouragement for the future. I also want to thank Arno Bohm, Volker Bromm, Jacques Distler, and Sonia Paban. This dissertation would have not been published without your support.

I also want to express my thanking to Dipsikha Debnath, James S. Gainer, Doojin Kim, and Konstantin T. Matchev, for you the Gators have accomplished so many fantastic ideas and researches with us the Longhorns, as Matthew D. Klimek, Sivaramakrishnan Swaminathan, Cynthia Trendfilova, Niral Desai, and Taewook Youn.

I want to thank specially to Janice Duffy and AJ Bunyard, for your help in administrative and mischievous stuff in the theory group. I shall also be grateful to Josiah Couch, who has been a very helpful friend especially in Europe, and also other people in the theory group including, Phuc Nguyen, Brandon DiNunno, Yinan Zhu, Fei Yan, Minglei Xiao, Stefan Eccles, Ravi Mohan, Robbie Rosati, Behzat Ergun, Tyler Guglielmo, Anderson Misobuchi, Qianyu Hao, Vikas Aragam, and Joshua Ziegler. It was a pleasure to work with you.

Also thanks Taiwanese people here in the University of Texas as well as Austin Taiwanese Presbyterian Church. All of you have brought me the taste of Taiwan in the opposite side of the Earth.

Finally, I would like to thank Jessie Lin, for you have given up all possible chances and future in the New York city just for keeping me company in Texas. Now it's time to fulfill my promise to return there together.

Lí tùi lí ê lô-po̍k só si it-chhè ê chû-ài kap it-chhè sêng-si̍t, chiū chì-bî-sòe ê góa iā bōe kham-tit tit-ti̍oh; góa chêng kia̍h góa ê phuê-siu^{n}, kè chit ê Thài-pîng-iû^{n}; ta^{n} góa í-keng chiâ^{n}-chòe nn̄g-ê huè-kuī. Góa chêng hī-khang thia^{n}-kì^{n} Lí, Ta^{n} góa ê ba̍k-chiu khòa^{n}-kì^{n} Lí; Goān Lí ê ông-kok lâi-lîm, Goān Lí ê chí-ì chāi tē-siōng si̍t-hêng chhin-chhiū^{n} chāi thian-siōng.

Techniques for Discovery and Mass Measurement Using the Boundary of Many-Body Phase Space

Yuan-Pao Yang, Ph.D.
The University of Texas at Austin, 2019

Supervisor: Can Kılıç

Despite several independent clues indicating new physics at the TeV scale, a lack of conclusive evidence for new physics so far suggests that future discoveries may only come from small numbers of new signal events. Therefore, it is crucial to find optimal methods with which to accurately measure the properties of new physics despite low statistics. In this scenario, the traditional kinematics variables such as edges and endpoints, which are merely one dimensional projections of the kinematic boundaries of a higher dimensional phase space, may not be the most efficient way to utilize limited information, especially in events consisting of cascade decays with large Standard Model backgrounds.

In this dissertation, adopting a benchmark decay topology with one invisible particle in four-body final states proceeding through a sequence of two-body decays via intermediate resonances, I will focus on techniques which utilize the boundary of the allowed four-body phase space as well as the symmetric Gram determinant of the momenta of final state particles Δ_4, which is essential for parametrizing the phase space in a manifestly Lorentz-invariant way, to improve mass measurements and discovery sensitivity significantly over more conventional variable choices and techniques.

Contents

List of Figures

Chapter One: Introduction

1.1 The Standard Model

To this day, the Standard Model of particle physics has been the best description of particles and interactions between them, which constitute our universe. All components of the Standard Model have been observed including the discovery of the Higgs Boson, and their properties have been measured to be consistent with Standard Model predictions, with only one exception in the neutrino sector, where at least two flavors of neutrinos have non-zero masses.

The dynamics of the Standard Model is described by quantum field theory, which is a quantum mechanical theory invariant under the Poincaré spacetime symmetry. The Standard Model is composed of three type of fields:

- Gauge fields, which are vector fields and the connections of local internal (i.e., non-spacetime) symmetries. The gauge structure of the Standard Model is the symmetry group $SU(3)_c \times SU(2)_L \times U(1)_Y$, with with corresponding charges called color, weak isospin and hypercharge.
- Matter fields, which are spinor fields and defined by their transformation properties under both the internal (global as well as local) and spacetime symmetries. They can be further divided into subgroups called quarks and leptons according to whether they carry color or not.
- Higgs field, which is a scalar field with nonzero vacuum expectation value (VEV), and this VEV breaks the $SU(2)_L \times U(1)_Y$ into $U(1)_{em}$ at low energy.

The Standard Model particles are summarized in 1.1. The Standard Model contains 19 parameters whose numerical values are taken as inputs rather than explained. There are 3 masses for leptons, 6 masses for quarks, 3 CKM mixing angles, 1 CKM CP-violation phase, 3 gauge couplings, 1 Higgs mass, 1 Higgs VEV, and 1 QCD vacuum angle.

Category	Symbol	Name	$SU(3)_c$	$SU(2)_L$	$U(1)_C$	B	L
Gauge Fields	G	Gluon	**8**	**1**	0	0	0
	W	W Boson	**1**	**3**	0	0	0
	B	B Boson	**1**	**1**	0	0	0
Fermion Fields	q_L	Left-handed Quark	**3**	**2**	$\frac{1}{3}$	$\frac{1}{3}$	0
	u^C	Left-handed Anti-u-Quark	$\bar{\mathbf{3}}$	**1**	$-\frac{4}{3}$	$-\frac{1}{3}$	0
	d^C	Left-handed Anti-d-Quark	$\bar{\mathbf{3}}$	**1**	$\frac{2}{3}$	$-\frac{1}{3}$	0
	ℓ_L	Left-handed Lepton	**1**	**2**	-1	0	1
	ℓ_L^C	Left-handed Anti-Lepton	**1**	**1**	2	0	-1
Higgs Field	H	Higgs	**1**	**2**	1	0	0

Table 1.1: Standard Model Fields Summary

1.2 The limitations of the Standard Model and TeV scale Supersymmetry

Despite the fact that Standard Model has achieved tremendous success in predicting experimental results, there are still a number of phenomena not yet fully understood, including, but not limited to: gravity, neutrino masses, dark matter, the hierarchy problem, the flavor problem, and so on. Among those deficits, the puzzles that seemed promising to be solved at the TeV scale are:

- **Hierarchy Problem**. The Standard Model contains only one dimensional parameter, the Higgs mass, which is of order 100 GeV at the EW energy scale. In the language of effective field theory, Standard Model is the IR theory of some UV theory at a higher scale, which is usually taken to be the (10^{19} GeV). However, in the framework of Quantum Field Theory, the Higgs mass is UV sensitive without any protection unlike the fermion case. Therefore, given the Planck scale at 10^{19} GeV, why the Higgs mass is only 100 GeV is a fine tuning problem. Some very popular theories, including Supersymmetry, aim to solve this problem by introducing some TeV scale particle to balance the running of Higgs mass contributed from Standard Model Particles.

- **Dark Matter**. Many independent astronomical observations indicate the existence of dark matter and it is well acknowledged that dark matter composes $80-85\%$ of matter in the universe and rarely interacts with baryonic particles.

However, there is no such particle in the Standard Model which could provide the corresponding relic abundance that is required by studies of cosmological models. The easiest way for obtaining the correct abundance of dark matter today via thermal production predicts a new particle a new particle in the 100 GeV mass range that interacts via some force of electroweak strength, and this phenomenon is called " the WIMP miracle" accordingly. Partners to neutral Standard Model particles provided by supersymmetry have long been prime WIMP candidates.

1.3 Purpose and Outline of this Thesis

Despite the study of Supersymmetry and other theories aimed to extend the Standard Model to solve its deficits in the TeV scale, the lack of conclusive evidence for new physics so far suggests that future discoveries may only possess small numbers of signal events. Therefore, it is crucial to find optimal methods with which we can accurately measure the properties of new physics despite low statistics. In this scenario, the traditional kinematics variables such as edges and endpoints, which are merely the one dimensional projection of the kinematic boundaries of a higher dimensional phase space, may not be the most efficient way to utilize limited information.

In previous research[1], it has been demonstrated that mass measurements for new particles appearing in decay chains can be improved by determining the boundary of the available phase space in its full dimensionality rather than by using one-dimensional kinematic features for each stage of the cascade decay. However, this result assumed an ideal situation where only signal events were present, no combinatorics showed up, all detectors possessed perfect resolution, and all particles in the decay chain decayed isotropically. This thesis focuses on how to implement a viable mass measurement methodology for kinematic boundaries in a more realistic experimental environment, and highlights the advantages of using Δ_4, which will be explained latter, as a discovery variable. The Chapter 2 will be a self-contained mathematical description to Phase Space, and the following chapters will be the application of it.

Chapter Two: Phase Space

2.1 Introduction

To study a decay process from an initial state of single particle to a final state of n particles, it is necessary to describe the initial as well as the final states, and also the transition rate from an given initial state to a final state.

To describe a physical system, it is common to use the *momentum-space* of a state with n particles in three dimensional space, which is the the set of all possible momenta, $\mathbf{p}_f$, which is a $3n$ space defined by the n momenta.

Now consider a system of a decay process which is composed of n particles with energy E_f and momenta $\mathbf{p}_r$ as the decay products of a single particle with energy E_0 and momentum $\mathbf{p}_0$. The physical allowed region is not the unconstrained $3n$ space anymore: One has to impose the condition of momentum conservation

$$\begin{aligned} E_0 &= E_1 + E_2 + \cdots + E_n = \sum_{i=1}^{n} E_i \\ \mathbf{p}_0 &= \mathbf{p}_1 + \mathbf{p}_2 + \cdots + \mathbf{p}_n = \sum_{i=1}^{n} \mathbf{p}_i \end{aligned} \tag{2.1}$$

where (according to dispersion relation)

$$E_i^2 = \mathbf{p}_i^2 + m_i^2$$

on the final state momentum vectors. As a result, the n momentum vectors $\mathbf{p}_f$ cannot vary arbitrarily for a fixed initial state, but have to satisfy the four conditions 2.1. The conditions 2.1 define a $3n-4$ dimensional surface called the *phase-space* inside the momentum-space.

Now for the transition rate. According to the quantum mechanics, the transition rate from an initial state i to a final state f is described by *Fermi's golden rule*, which utilizes the *matrix element*:

$$\langle f \mid S \mid i \rangle \equiv \langle f \mid (\mathbb{1} + iT) \mid i \rangle \tag{2.2}$$

Particularly, in the case of particle decay, one then has:

$$\begin{aligned}\langle \mathbf{p}_1 \cdots \mathbf{p}_n \mid S \mid \mathbf{p}_0 \rangle &= \langle \mathbf{p}_1 \cdots \mathbf{p}_n \mid (iT) \mid \mathbf{p}_0 \rangle \\ &\equiv (2\pi)^4 \delta^{(4)}(\mathbf{p}_0 - \sum \mathbf{p}_f) \cdot i\mathcal{M}(\mathbf{p}_0 \to \{\mathbf{p}_f\})\end{aligned} \tag{2.3}$$

and then the Fermi's golden rule states that the differential decay rate is

$$d\Gamma = \frac{1}{2m_0} \mid \mathcal{M}(\mathbf{p}_0 \to \{\mathbf{p}_f\}) \mid^2 d\Pi_n(\mathbf{p}_0; \mathbf{p}_f) \tag{2.4}$$

where $d\Pi_n(\mathbf{p}_0; \mathbf{p}_f)$ is the element of phase space, and is determined by

$$d\Pi_n(\mathbf{p}_0; \mathbf{p}_f) = (2\pi)^4 \delta^{(4)}(\mathbf{p}_0 - \sum \mathbf{p}_f) \prod_{f=1}^{n} \frac{d^3 p_f}{(2\pi)^3} \frac{1}{2E_f} \tag{2.5}$$

and one may integrate over the phase space to obtain the total decay rate.

More generally, one can change $\mathbf{p}_f$ into a set of variables $\mathbf{\Phi}$[2]. Then 2.4 will have the form:

$$d\Gamma = \frac{1}{2m_0} d\mathbf{\Phi}\ \rho_n(\mathbf{\Phi})\ \mid \mathcal{M}(\mathbf{\Phi}) \mid^2 \tag{2.6}$$

where $d\mathbf{\Phi}$ is a volume element in the $3n-4$ dimensional phase space and the *phase space density* $\rho_n(\mathbf{\Phi})$ contains all factors arising from transforming from the $\mathbf{p}_i$ to the variables $\mathbf{\Phi}$. When $\mathbf{p}_1, \cdots, \mathbf{p}_n$ vary over the whole phase space, the set $\mathbf{\Phi}$ varies over a $3n-4$ dimensional region which is the *physical region of* $\mathbf{\Phi}$. As we will see shortly, it is very convenient to choose a set of variables $\mathbf{\Phi}$ that is Lorentz invariant.

2.1.1 Phase Space Integral and Two Body Final State Phase Space

If the matrix square, $\mid \mathcal{M} \mid^2$ is identically 1, then the decay rate:

$$\Gamma = \frac{1}{2m_0} \Pi_n \tag{2.7}$$

is said to be given by *phase space*. Moreover, the probability density function $\frac{1}{\Gamma}\frac{d\Gamma}{d\phi}$ are called *phase space distributions*[2].

In the case of two body phase space, this integration is extremely easy [3]. The phase space integral is:

$$\begin{aligned}\Pi_2 &= \int (2\pi)^4 \delta^{(4)}(p_0 - p_1 - p_2) \frac{d^3 p_1}{(2\pi)^3} \frac{1}{2E_1} \frac{d^3 p_2}{(2\pi)^3} \frac{1}{2E_2} \\ &= (4\pi)^{-2} \int \delta^{(3)}(\mathbf{p}_0 - \mathbf{p}_1 - \mathbf{p}_2) \delta(E_0 - E_1 - E_2) \frac{d^3 p_1}{E_1} \frac{d^3 p_2}{E_2}\end{aligned} \tag{2.8}$$

In the rest frame of the initial particle, $\mathbf{p}_0 = (m_0, 0, 0)$, therefore,

$$\begin{aligned}
\Pi_2 &= (4\pi)^{-2} \int \delta^{(3)}(\mathbf{p}_1 + \mathbf{p}_2)\delta(m_0 - \sqrt{\mathbf{p}_1^2 + m_1^2} - \sqrt{\mathbf{p}_2^2 + m_2^2}) \frac{d^3p_1}{\sqrt{\mathbf{p}_1^2 + m_1^2}} \frac{d^3p_2}{\sqrt{\mathbf{p}_2^2 + m_2^2}} \\
&= (4\pi)^{-2} \int \frac{\delta(m_0 - \sqrt{\mathbf{p}_1^2 + m_1^2} - \sqrt{\mathbf{p}_2^2 + m_2^2})}{\sqrt{\mathbf{p}_1^2 + m_1^2}\sqrt{\mathbf{p}_1^2 + m_2^2}} d^3p_1 \\
&= (4\pi)^{-2} \int \frac{\delta(m_0 - \sqrt{\mathbf{p}_1^2 + m_1^2} - \sqrt{\mathbf{p}_2^2 + m_2^2})}{\sqrt{\mathbf{p}_1^2 + m_1^2}\sqrt{\mathbf{p}_1^2 + m_2^2}} p_1^2 dp_1 d\Omega \\
&= (4\pi)^{-1} \int \frac{\delta(m_0 - \sqrt{\mathbf{p}_1^2 + m_1^2} - \sqrt{\mathbf{p}_2^2 + m_2^2})}{\sqrt{\mathbf{p}_1^2 + m_1^2}\sqrt{\mathbf{p}_1^2 + m_2^2}} p_1^2 dp_1
\end{aligned} \tag{2.9}$$

It can be shown that the final integral gives

$$\Pi_2 = \frac{1}{8\pi} \frac{\lambda^{1/2}(m_0^2, m_1^2, m_2^2)}{m_0^2} \tag{2.10}$$

where $\lambda(x, y, z) \equiv x^2 + y^2 + z^2 - 2xy - 2yz - 2zx$ is the *basic three-particle kinematic function*[2]. Therefore, the phase space decay rate of two final state is:

$$\Gamma_2 = \frac{1}{16\pi} \frac{\lambda^{1/2}(m_0^2, m_1^2, m_2^2)}{m_0^3} \tag{2.11}$$

2.2 General Multi-body Phase Space

2.2.1 Three Body Phase Space

In the case of more-than-two body final states, one can further reduce the number of final state variables by mapping them to the rest frame of the mother particle. Therefore, three variables are trivial and there remains $3n - 7$ *essential final state variables*.

In the case of three body final state, the number of essential final state variables is two. It is both conventional and convenient to choose two invariant masses $m_{12}^2 \equiv (\mathbf{p}_1 + \mathbf{p}_2)^2$ and $m_{23}^2 \equiv (\mathbf{p}_2 + \mathbf{p}_2)^3$ as the variables.
A strict mathematical description of the three body final state decay can be found at [2]. Here we only extract important results:

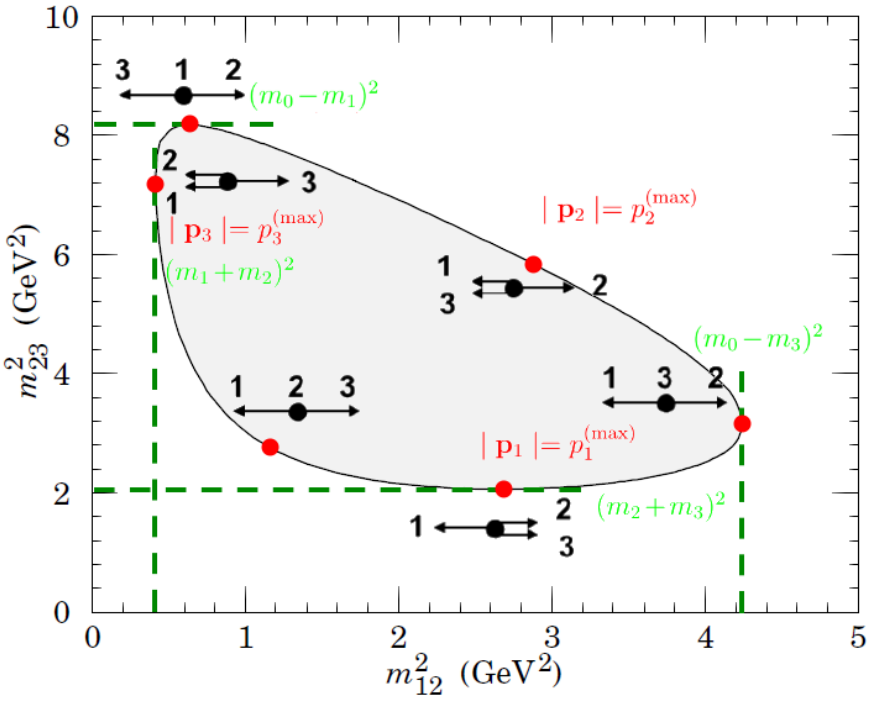


Figure 2.1: Configurations of momentum vectors in the center of mass of the mother particle on the boundary of the Dalitz plot.[4]

- The physical region of m_{12}^2 and m_{23}^2 for three body final state decays can be described as

$$G(m_{12}^2, m_{23}^2, m_0^2, m_2^2, m_1^2, m_3^2) \leq 0 \tag{2.12}$$

where $G(x, y, z, u, v, w)$ is the *basic four-particle kinematic function*, and can be expressed as a Cayley determinant:

$$G(x, y, z, u, v, w) = -\frac{1}{2}\begin{vmatrix} 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & v & x & z \\ 1 & v & 0 & u & y \\ 1 & x & u & 0 & w \\ 1 & z & y & w & 0 \end{vmatrix} \tag{2.13}$$

One notable result is that the final state momenta $\mathbf{p}_1$, $\mathbf{p}_2$, and $\mathbf{p}_3$ are collinear on the boundary of the Dalitz plot (Fig.2.1).

- Within the physical region, the phase space distribution is:

$$\frac{d^2\Pi_3}{dm_{12}^2 dm_{23}^2} = \frac{1}{16m_0^2} \tag{2.14}$$

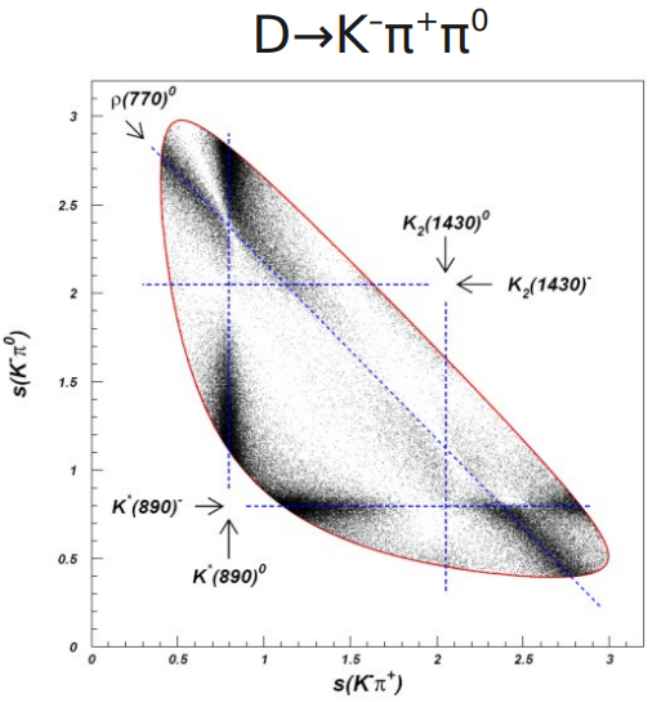


Figure 2.2: An example for the application of the Dalitz plot. The non-uniform distribution is due to the interference between the intermediate resonances[5].

That is, the phase space distribution density within the physical region is flat.

2.2.1.1 Dalitz Plot

The Dalitz Plot is defined as the physical region of $\mathbf{p}_0 \to \mathbf{p}_1 + \mathbf{p}_2 + \mathbf{p}_3$ in the $m_{12}^2 m_{23}^2$ plane. The boundary and the phase space density is given by 2.12 and 2.14, respectively. Since the decay rate is

$$d\Gamma = \frac{1}{(2\pi)^3} \frac{1}{8m_0^3} dm_{12}^2 dm_{23}^2 \ \mid \mathcal{M}(m_{12}^2, m_{23}^2) \mid^2 \tag{2.15}$$

within the physics region, Dalitz Plot is occasionally used to identify the matrix element, especially for resonant processes, in which the particle decays into two decay products, with one of those decay products immediately decaying into two additional decay products. In this case, the Dalitz plot will show a non-uniform distribution, with a peak around the mass of the resonant decay (see Fig.2.2 for an example.).

2.2.2 Four and More Body Phase Space

Again, as in the case of three body final states, invariant masses $m_{ij}^2 \equiv (\mathbf{p}_i + \mathbf{p}_j)^2$ are a very natural and convenient choice for parametrizing phase space. However, unlike the case of three body final states, which has a flat distribution within the physical region, the phase space boundary and distribution of more-than-three body phase space is not as trivial as the three body case. Here we will summarize the corresponding results from [6], which has a very detailed discussion for n-body final states in an m-dimensional space.

For n-body final states, first define the matrix

$$\mathcal{Z} = \{z_{ij}\} \quad \text{with} \quad z_{ij} = p_i \cdot p_j \,, \tag{2.16}$$

where the $\{p_i\}$ are the four momenta of the final state particles. The variables Δ_i can then be defined as

$$\det\left[\lambda I_{n\times n} - \mathcal{Z}\right] \equiv \lambda^n - \left(\sum_{i=1}^{n} \Delta_i \lambda^{n-i}\right). \tag{2.17}$$

Among these variables, Δ_4 will play a special role in the four dimensional spacetime.

- **Boundary**

  As described in ref. [6], the kinematically allowed region is given by $\Delta_{1,2,3,4} \geq 0$ (and Δ_i, $i > 4$ will be zero automatically in four dimensions), with the boundary located at

  $$\Delta_4 = 0, \qquad \Delta_{1,2,3} \geq 0\,. \tag{2.18}$$

- **Distribution**

  The phase space distribution in the region of $\Delta_4 \geq 0$ is

  $$d\Pi_n = (\text{Const.}) \times (m_0)^{-2}\Big(\prod_{i<j} dm_{ij}\Big)\Delta_4^{(n-5)/2}\Theta(\Delta_4)\delta(\Delta_5)\cdots\delta(\Delta_n)\delta\Big(\sum_{i<j} m_{ij} - K\Big) \tag{2.19}$$

  where $K = m_0^2 + 2\sum_{i=1}^{n} m_i^2$.

In the rest of this dissertation, we will focus on the case of four body final states. From 2.19, the general four-body phase space distribution density element is given by

$$d\Pi_4 = \left(\prod_{i<j} dm_{ij}^2\right) \frac{8}{(4\pi)^{10} m_0^2 \Delta_4^{1/2}} \, \delta\left(\sum_{i<j} m_{ij}^2 - \left(m_0^2 + 2\sum_{i=1}^{4} m_i^2\right)\right), \tag{2.20}$$

where $m_{ij}^2 = (p_i + p_j)^2 = 2z_{ij} + m_i^2 + m_j^2$. One has to keep in mind that the m_{ij}^2 are not all independent variables: the δ-function in 2.20 eliminates one degree of freedom from the six $m_i j^2$, therefore, only five degrees of freedom remain. This matches the formula of $3n - 7$ for $n = 4$ case.

As we can see, this phase space density is not flat as in the case of the three-body phase space, but instead depends on the phase space coordinates through the factor of Δ_4 which is a function of the m_{ij}^2 . Since the boundary of the kinematically allowed region is defined by $\Delta_4 = 0$, this means that the phase space density is enhanced near the boundary of the physical region. The apparent singularity is an integrable one, and it is not canceled by the presence of the δ-function enforcing momentum conservation, as its argument is linear in the m_{ij}^2 and does not lead to any nontrivial Jacobian factors.

For the purpose of application, we assume the following final state particles:

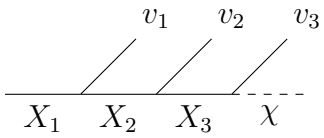
- One massive but missed particle, which represents a dark matter candidate.
- Three visible but effectively massless particles, since visible jets or leptons are far less massive than dark matter candidates.

Note that there are only a few possibilities for how a cascade decay to four final state particles may proceed in stages. Denoting the initial particle by X_1, any intermediate resonances by X_2, X_3 etc., the visible final state particles as v_1, v_2, and v_3, and particle χ always being the invisible one, the possibilities are (Fig 2.3)

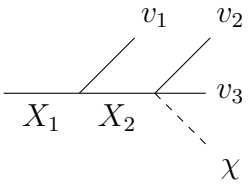
- $X_1 \to v_1 + X_2$, $X_2 \to v_2 + X_3$, $X_3 \to v_3 + \chi$

  This is a three stage cascade decay, each stage being a two-body decay. We will refer to this possibility as the "2+2+2" topology.

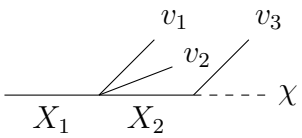
- $X_1 \to v_1 + X_2$, $X_2 \to v_2 + X_3 + \chi$



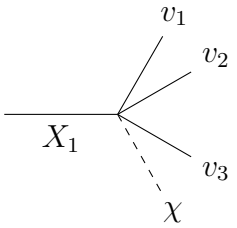
(a) "2+2+2" topology



(b) "2+3" topology



(c) "3+2" topology



(d) genuine four-body decay

Figure 2.3: Topologies of four final state decay

This is a two stage cascade decay, with a two-body decay followed by a three-body decay, and is referred to as the "2+3" decay topology.

- $X_1 \to v_1 + v_2 + X_2$, $X_2 \to X_3 + \chi$

  This is a two stage cascade decay, with a three-body decay followed by a two-body decay, and is referred to as the "3+2" decay topology.

- $X_1 \to v_1 + v_2 + v_3 + \chi$

  A genuine four-body decay with no intermediate resonances.

In the remainder of this chapter, we will focus on our benchmark decay topology (Fig.2.3a), which is composed of three steps of sequential two-body decays. Each intermediate state will eliminate one degree of freedoms and therefore the number of essential final state variables remains $5 - 2 = 3$.

2.3 Phase Space of "2+2+2" decay with one missing final state

2.3.1 Boundary of the Phase Space: $\Delta_4 = 0$

Now we focus on our benchmark decay topology, which is composed of three steps of sequential two-body decays (Fig 2.3a).

2.3.1.1 Mathematical Description

Before any formal derivation, we can already get some qualitative result:

- Each intermediate resonance will eliminate one degree of freedom. therefore the number of essential final state variables remains $5 - 2 = 3$, which matches the number of observable m_{ij}^2 (that is, m_{12}^2, m_{13}^2, and m_{23}^2 since $\mathbf{p}_\chi$ is not visible.)
- Since m_{12}^2, m_{13}^2, and m_{23}^2 are a good set of final state variables, we can use the well-known edges/endpoints of them to get the approximate location of the phase space:

$$\begin{aligned}(m_{12}^2)^{(\max)} &= \frac{(m_{X_1}^2 - m_{X_2}^2)(m_{X_2}^2 - m_{X_3}^2)}{m_{X_2}^2} \\ (m_{13}^2)^{(\max)} &= \frac{(m_{X_1}^2 - m_{X_3}^2)(m_{X_3}^2 - m_\chi^2)}{m_{X_3}^2} \\ (m_{23}^2)^{(\max)} &= \frac{(m_{X_2}^2 - m_{X_3}^2)(m_{X_3}^2 - m_\chi^2)}{m_{X_3}^2}\end{aligned} \tag{2.21}$$

As a result, we can see that the the physical allowed region is actually contained in a rectangular cuboid in the first octant with three edges of length $(m_{12}^2)^{(\max)}$, $(m_{13}^2)^{(\max)}$, and $(m_{23}^2)^{(\max)}$.

For a quantitative derivation: First, it can be shown that the physical region for the $2 + 2 + 2$ topology is bounded by $\Delta_4 = 0$ in the first octant. Therefore, we would like to get an analytic form of Δ_4. For $2 + 2 + 2$ topology, we have two more conditions:

$$\begin{aligned}(m_{23}^2 + m_{24}^2 + m_{34}^2) - (m_{X_2}^2 + m_2^2 + m_3^2 + m_\chi^2) &= 0 \\ m_{34}^2 - m_{X_3}^2 &= 0\end{aligned} \tag{2.22}$$

Since we have assumed $m_1 = m_2 = m_3 = 0$, we have

$$\begin{aligned} m_{12}^2 + m_{13}^2 + m_{14}^2 + m_{23}^2 + m_{24}^2 + m_{34}^2 - (m_{X_1}^2 + 2m_\chi^2) &= 0 \\ m_{23}^2 + m_{24}^2 + m_{34}^2 - (m_{X_2}^2 + m_\chi^2) &= 0 \\ m_{34}^2 - m_{X_3}^2 &= 0 \end{aligned} \tag{2.23}$$

Or

$$\begin{aligned} m_{14}^2 &= (m_{X_1}^2 - m_{X_2}^2 + m_\chi^2) - (m_{12}^2 + m_{13}^2) \\ m_{24}^2 &= (m_{X_2}^2 - m_{X_3}^2 + m_\chi^2) - m_{23}^2 \\ m_{34}^2 &= m_{X_3}^2 \end{aligned} \tag{2.24}$$

Use those three equations, we can eliminate unwanted m_{14}^2, m_{24}^2, and m_{34}^2. Furthermore, we define $\xi_{ij} = m_{ij}^2/(m_{ij}^2)^{\max}$ (*i.e.*, scale m_{ij}^2 to be between 0 and 1). It can be shown that

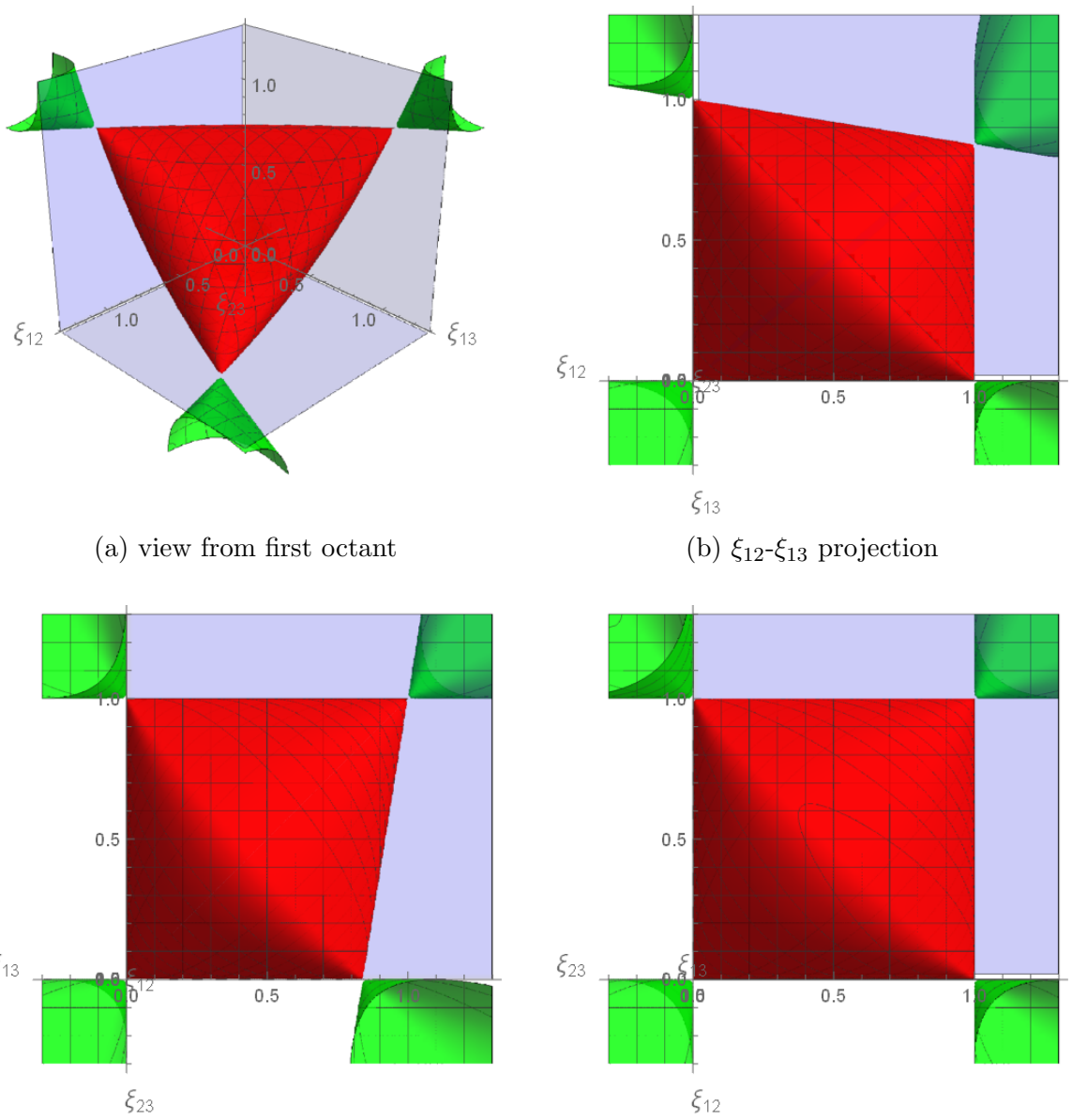
$$\Delta_4 = \Delta_4^{\max}\Big[16\xi_{12}(1-\xi_{12})\xi_{23}(1-\xi_{23}) - 4\big(\frac{m_{X_2}}{m_{X_3}}(\xi_{13}-\xi_{23}+\xi_{12}\xi_{23}) - \frac{m_{X_3}}{m_{X_2}}\xi_{12}(1-\xi_{23})\big)^2\Big] \tag{2.25}$$

with

$$\Delta_4^{(\max)} = \frac{1}{64}\frac{(m_{X_1}^2 - m_{X_2}^2)^2(m_{X_2}^2 - m_{X_3}^2)^2(m_{X_3}^2 - m_{X_4}^2)^2}{m_{X_2}^2 m_{X_3}^2} \tag{2.26}$$

From this equation, we can see some properties of the physical region of m_{12}^2, m_{13}^2, and m_{23}^2:

- After the scaling, the shape of the physical region only depends on the ratio m_{X_3}/m_{X_2}
- The physical region has a samosa-like shape inside the box in the first octant (See Fig.2.4).
- The value of Δ_4 will be the maximal at $(\xi_{12}, \xi_{13}, \xi_{23}) = (\frac{1}{2}, \frac{1}{4}(1 + (\frac{m_{X_3}}{m_{X_2}})^2), \frac{1}{2})$, which we will call the 'center' of the phase space.
- The region $\Delta_4 > 0$ has four 'horns' outside the first octant (of course, these are unphysical) connected to the samosa at points $(0,0,0)$, $(1, (\frac{m_{X_3}}{m_{X_2}})^2, 0)$, $(1,0,1)$ and $(0,1,1)$



(a) view from first octant

(b) ξ_{12}-ξ_{13} projection

(c) ξ_{13}-ξ_{23} projection

(d) ξ_{23}-ξ_{12} projection

Figure 2.4: The $\Delta_4 \geq 0$ region with different view point. The red "samosa" is the physical region, and the green "horns" are the regions with $\Delta_4 \geq 0$ but not physical (that is, with $\Delta_3 < 0$). The transparent blue is the boundary of the first octant.

2.3.1.2 Boundaries of the Phase Space with different mass spectrum

For an ideal case where only signal events are present, the boundary of the phase space is a good tool to determine the mass spectrum: for any mass spectrum hypothesis,

if any signal events fall outside the phase space boundary, then the mass hypothesis must be wrong. However, the reverse is not quite true: If for a mass spectrum hypothesis, there is not any events fall outside the boundary, the hypothesis is not necessarily true. However, we can still draw some conditions on the mass hypothesis.

Now consider a true spectrum m_{X_1}, m_{X_2}, m_{X_3}, and m_{X_4}, which has the edges/endpoints $(m_{12}^2)^{(\max)}$, $(m_{13}^2)^{(\max)}$, and $(m_{23}^2)^{(\max)}$, and the hypothetical mass spectrum $\tilde{m}_{X_1}$, $\tilde{m}_{X_2}$, $\tilde{m}_{X_3}$, $\tilde{m}_{X_4}$ with edges/endpoints $(\tilde{m}_{12}^2)^{(\max)}$, $(\tilde{m}_{13}^2)^{(\max)}$, and $(\tilde{m}_{23}^2)^{(\max)}$. Furthermore, define $y = m_{X_3}/m_{X_2}$, $\tilde{y} = \tilde{m}_{X_3}/\tilde{m}_{X_2}$. Now the Δ_4 for the hypothetical mass spectrum is:

$$\begin{aligned}\tilde{\Delta}_4 &= (\tilde{\Delta}_4)^{(\max)}\Big[16\tilde{\xi}_{12}(1-\tilde{\xi}_{12})\tilde{\xi}_{23}(1-\tilde{\xi}_{23}) - 4\big(\tilde{y}^{-1}(\tilde{\xi}_{13}-\tilde{\xi}_{23}+\tilde{\xi}_{12}\tilde{\xi}_{23}) - \tilde{y}\tilde{\xi}_{12}(1-\tilde{\xi}_{23})\big)^2\Big] \\ &= (\tilde{\Delta}_4)^{(\max)}\Big[16R_{12}\xi_{12}(1-R_{12}\xi_{12})R_{23}\xi_{23}(1-R_{23}\xi_{23}) \\ &\qquad - 4\big(\tilde{y}^{-1}(R_{13}\xi_{13}-R_{23}\xi_{23}+R_{12}\xi_{12}R_{23}\xi_{23}) - \tilde{y}R_{12}\xi_{12}(1-R_{23}\xi_{23})\big)^2\Big]\end{aligned} \tag{2.27}$$

With $\tilde{\xi}_{ij} \equiv m_{ij}^2/(\tilde{m}_{ij}^2)^{(\max)} = R_{ij}\xi_{ij}$ and $R_{ij} \equiv (m_{ij}^2)^{(\max)}/(\tilde{m}_{ij}^2)^{(\max)}$.

If the signal events are contained inside the hypothetical phase space, $\tilde{\Delta}_4$ must be positive for all events. Therefore, as special cases, all four special points we mentioned about should also carry positive $\tilde{\Delta}_4$:

- $(0,0,0)$:
$$\tilde{\Delta}_4(0,0,0) = 0 \tag{2.28}$$

- $(1,y^2,0)$:
$$\tilde{\Delta}_4(1,y^2,0) = -4(\tilde{\Delta}_4)^{(\max)}\big(\tilde{y}^{-1}R_{13}y^2 - \tilde{y}R_{12}\big)^2 \geq 0 \tag{2.29}$$

- $(1,0,1)$:
$$\begin{aligned}\tilde{\Delta}_4(1,0,1) &= (\tilde{\Delta}_4)^{(\max)}\Big[16R_{12}(1-R_{12})R_{23}(1-R_{23}) \\ &\qquad - 4\big(\tilde{y}^{-1}(-R_{23}+R_{12}R_{23}) - \tilde{y}R_{12}(1-R_{23})\big)^2\Big] \\ &= -4(\tilde{\Delta}_4)^{(\max)}y^{-2}(R_{23}(1-R_{12}) - \tilde{y}^2R_{12}(1-R_{23}))^2 \geq 0\end{aligned} \tag{2.30}$$

- $(0,1,1)$:
$$\tilde{\Delta}_4(0,1,1) = -4(\tilde{\Delta}_4)^{(\max)}\big(\tilde{y}^{-1}(R_{13}-R_{23})\big)^2 \geq 0 \tag{2.31}$$

To satisfy 2.29 and 2.31, the only possibility is $R_{12}(\tilde{y}/y)^2 = R_{13} = R_{23}$. Now to satisfy 2.30, we must have:

$$R_{12} = 0 \text{ or } R_{12} = \frac{1-y^2}{1-\tilde{y}^2} \tag{2.32}$$

The first simply says the true spectrum is totally degenerate. For the second option, from the definition of R_{23}:

$$R_{12} = \frac{(m_{X_1}^2 - m_{X_2}^2)(m_{X_2}^2 - m_{X_3}^2)/m_{X_2}^2}{(\tilde{m}_{X_1}^2 - \tilde{m}_{X_2}^2)(\tilde{m}_{X_2}^2 - \tilde{m}_{X_3}^2)/\tilde{m}_{X_2}^2} = \frac{(m_{X_2}^2 - m_{X_3}^2)/m_{X_2}^2}{(\tilde{m}_{X_2}^2 - \tilde{m}_{X_3}^2)/\tilde{m}_{X_2}^2} \tag{2.33}$$

or equivalently, $m_{X_1}^2 - m_{X_2}^2 = \tilde{m}_{X_1}^2 - \tilde{m}_{X_2}^2$. Combine this with the condition we have before, we will have:

$$\begin{aligned} m_{X_1}^2 - m_{X_2}^2 &= \tilde{m}_{X_1}^2 - \tilde{m}_{X_2}^2 \\ m_{X_2}^2 - m_{X_3}^2 &= \tilde{m}_{X_2}^2 - \tilde{m}_{X_3}^2 \\ m_{X_3}^2 - m_{X_4}^2 &= \tilde{m}_{X_3}^2 - \tilde{m}_{X_4}^2 \end{aligned} \tag{2.34}$$

In other words, the mass square gap between the true spectrum and hypothetical spectrum is the same (or the hypothetical one is just an overall shift of the true one.).

The condition above is merely for the four vertices to stay inside (actually, on) the boundary of the phase space of a hypothetical spectrum. Now we consider the boundary of the true spectrum (on which signal events lays) on which $\Delta_4 = 0$. Then we can solve:

$$\xi_{13}^{\pm} = (\sqrt{\xi_{23}(1-\xi_{12})} \pm y\sqrt{\xi_{12}(1-\xi_{23})})^2 \tag{2.35}$$

Setting $m_{X_2}^2 - \tilde{m}_{X_2}^2 = m_{X_3}^2 - \tilde{m}_{X_3}^2 = \Delta m^2$ since 2.34, it can be shown that:

$$\tilde{\Delta}_4 = \Delta m^2 \frac{16(\tilde{\Delta}_4)^{(\max)}}{m_{X_2}^4 m_{X_3}^4}(m_{X_2}^2 - \Delta m^2)(m_{X_3}^2 - \Delta m^2)(m_{X_2}\sqrt{\xi_{23}(1-\xi_{12})} \pm m_{X_3}\sqrt{\xi_{12}(1-\xi_{23})})^2 \tag{2.36}$$

Therefore, as long as Δm^2 is positive, the boundary of the true spectrum will be located entirely inside the hypothetical one.

In summary, for the physical phase space of a hypothetical mass spectrum to contain all the signal events of a true spectrum, we must have:

$$m_{X_1}^2 - \tilde{m}_{X_1}^2 = m_{X_2}^2 - \tilde{m}_{X_2}^2 = m_{X_3}^2 - \tilde{m}_{X_3}^2 = m_{X_4}^2 - \tilde{m}_{X_4}^2 = \Delta m^2 \geq 0 \tag{2.37}$$

2.3.1.3 Phase Space density distribution

For a quantitative derivation, note that in this particular case, it is equivalent to set the matrix element square in 2.7 as:

$$\mid \mathcal{M} \mid^2 = \mu^2_{X_1X_21} \frac{\mu^2_{X_2X_32}}{\mid m^2_{X_2} - (\mathbf{p}_2 + \mathbf{p}_3 + \mathbf{p}_{X_4})^2 + im_{X_2}\Gamma_{X_2} \mid^2} \frac{\mu^2_{X_3X_43}}{\mid m^2_{X_3} - (\mathbf{p}_3 + \mathbf{p}_{X_4})^2 + im_{X_3}\Gamma_{X_3} \mid^2} \tag{2.38}$$

With narrow width approximation, we have

$$\frac{1}{\mid m^2_{X_2} - (\mathbf{p}_2 + \mathbf{p}_3 + \mathbf{p}_{X_4})^2 + im_{X_2}\Gamma_{X_2} \mid^2} \approx \delta(m^2_{23} + m^2_{24} + m^2_{34} - K) \frac{\pi}{m_{X_2}\Gamma_{X_2}} \tag{2.39}$$

$$\frac{1}{\mid m^2_{X_3} - (\mathbf{p}_3 + \mathbf{p}_{X_4})^2 + im_{X_3}\Gamma_{X_3} \mid^2} \approx \delta(m^2_{34} - m_{X_3}) \frac{\pi}{m_{X_3}\Gamma_{X_3}} \tag{2.40}$$

where $K = m^2_{X_2} + m^2_2 + m^2_3 + m^2_4$. Therefore, applying 2.11, we get

$$\Gamma_{X_1} = \frac{\mu^2_{X_1X_21}}{16\pi} \frac{\lambda^{1/2}(m_{X_1}, m_{X_2}, 0)}{m^2_{X_1}} = \frac{\mu^2_{X_1X_21}}{16\pi} \frac{m^2_{X_1} - m^2_{X_2}}{m^3_{X_1}} \tag{2.41}$$

and

$$\Gamma_{X_2} = \frac{\mu^2_{X_2X_32}}{16\pi} \frac{\lambda^{1/2}(m_{X_2}, m_{X_3}, 0)}{m^2_{X_2}} = \frac{\mu^2_{X_2X_32}}{16\pi} \frac{m^2_{X_2} - m^2_{X_3}}{m^3_{X_2}} \tag{2.42}$$

and

$$\Gamma_{X_3} = \frac{\mu^2_{X_3X_43}}{16\pi} \frac{\lambda^{1/2}(m_{X_3}, m_{X_4}, 0)}{m^2_{X_3}} = \frac{\mu^2_{X_3X_43}}{16\pi} \frac{m^3_{X_3} - m^2_{X_4}}{m^3_{X_3}} \tag{2.43}$$

Therefore

$$\mid \mathcal{M} \mid^2 \approx \Gamma_{X_1}(4\pi)^6(1 - m^2_{X_2}/m^2_{X_1})^{-1}(1 - m^2_{X_3}/m^2_{X_2})^{-1}(1 - m^2_{X_4}/m^2_{X_3})^{-1} \tag{2.44}$$

Combining everything together, and integrating out m^2_{14}, m^2_{24}, and m^2_{34}, we get:

$$\frac{d\Gamma_{X_1}}{\Gamma_{X_1}} = \frac{1}{4\pi m^2_{X_1}}(1 - m^2_{X_2}/m^2_{X_1})^{-1}(1 - m^2_{X_3}/m^2_{X_2})^{-1}(1 - m^2_{X_4}/m^2_{X_3})^{-1} \frac{\Theta(\Delta_4)}{\sqrt{\Delta_4}} dm^2_{12} dm^2_{13} dm^2_{23} \tag{2.45}$$

Plugging back into 2.45, and applying the definition of phase space distribution, we get

$$d\mathrm{PDF}_{(2+2+2)} = \frac{m_{X_2} m_{X_3}}{4\pi\sqrt{\Delta_4^{(\max)}}} \frac{\Theta(\Delta_4)}{\sqrt{\Delta_4}} dm^2_{12} dm^2_{13} dm^2_{23} \tag{2.46}$$

2.3.2 Δ_4 and its properties

Recall that the boundary of the phase space is determined by $\Delta_4 = 0$

2.3.2.1 Δ_4 distribution of signal

Now we will derive the Δ_4 distribution of the signal events. Starting from 2.46:

$$d\text{PDF}_{(2+2+2)}(m_{12}^2, m_{13}^2, m_{23}^2) = \frac{m_{X_2} m_{X_3}}{32\pi\sqrt{\Delta_4^{(\max)}}} \frac{\Theta(\Delta_4)}{\sqrt{\Delta_4}} dm_{12}^2 dm_{13}^2 dm_{23}^2 \tag{2.46}$$

With

$$\Delta_4 = \Delta_4^{\max}\Big[16\xi_{12}(1-\xi_{12})\xi_{23}(1-\xi_{23}) - 4\big(\frac{m_{X_2}}{m_{X_3}}(\xi_{13}-\xi_{23}+\xi_{12}\xi_{23}) - \frac{m_{X_3}}{m_{X_2}}\xi_{12}(1-\xi_{23})\big)^2\Big] \tag{2.25}$$

Defining $q \equiv \Delta_4/\Delta_4^{\max}$ and changing m_{ij}^2 into ξ_{ij}^2, we can substitute $d\xi_{13}^2$ with $d\Delta_4$:

$$d\text{PDF}_{(2+2+2)}(\xi_{12}, \xi_{23}, q) = \frac{m_{X_2} m_{X_3} (m_{12}^2)^{(\max)} (m_{13}^2)^{(\max)} (m_{23}^2)^{(\max)}}{32\pi\Delta_4^{(\max)}} \frac{\Theta(q)}{\sqrt{q}} \mathcal{J} d\xi_{12} d\xi_{23} dq \tag{2.47}$$

where the prefactor is:

$$\frac{m_{X_2} m_{X_3} (m_{12}^2)^{(\max)} (m_{13}^2)^{(\max)} (m_{23}^2)^{(\max)}}{32\pi\Delta_4^{(\max)}} \tag{2.48}$$

$$= \frac{m_{X_2} m_{X_3} (m_{X_1}^2 - m_{X_2}^2)^2 (m_{X_2}^2 - m_{X_3})^2 (m_{X_3}^2 - m_{X_4})^2 / m_{X_2}^2 m_{X_3}^4}{32\pi (m_{X_1}^2 - m_{X_2}^2)^2 (m_{X_2}^2 - m_{X_3})^2 (m_{X_3}^2 - m_{X_4})^2 / 64 m_{X_2}^2 m_{X_3}^2} \tag{2.49}$$

$$= \frac{2 m_{X_2}}{\pi m_{X_3}} \tag{2.50}$$

with $\mathcal{J}$ the Jacobian:

$$\frac{m_{X_3}}{m_{X_2}} \frac{1}{2\sqrt{16\xi_{12}(1-\xi_{12})\xi_{23}(1-\xi_{23}) - q}} \tag{2.51}$$

where we have combined two solutions and sum them up. Note that only $\mathcal{J}$ depends on ξ_{12} and ξ_{23}, which are going to be integrated out. Therefore, it is sufficient to consider the integral of $\mathcal{J}$ only.

Now do the integration over ξ_{12}. To stay inside the samosa, we must have $16\xi_{12}(1-\xi_{12})\xi_{23}(1-\xi_{23})-q\geq 0$ the upper limit and lower limit are

$$\frac{1}{2}\Big[1\pm\sqrt{1-\frac{q}{4\xi_{23}(1-\xi_{23})}}\Big] \tag{2.52}$$

and the result is

$$\frac{\pi m_{X_3}}{8m_{X_2}\sqrt{\xi_{23}(1-\xi_{23})}} \tag{2.53}$$

Now do the integral for ξ_{23}. Still, we have to let $4\xi_{23}(1-\xi_{23})-q\geq 0$. So the limits are

$$\frac{1}{2}\Big[1\pm\sqrt{1-q}\Big] \tag{2.54}$$

and the result is

$$\frac{\pi m_{X_3}\cos^{-1}(\sqrt{q})}{4m_{X_2}} \tag{2.55}$$

Putting other factors together, we get:

$$d\mathrm{PDF}_{2+2+2}(q)=\frac{\cos^{-1}(\sqrt{q})}{2\sqrt{q}} \tag{2.56}$$

2.3.2.2 Δ_4 contours

From the previous section, it is easy to get the volume enclosed by fixed $\Delta_4>0$:

$$\begin{aligned}
V&=\int dm_{12}^2dm_{13}^2dm_{23}^2\\
&=\int (m_{12}^2)^{(\max)}(m_{13}^2)^{(\max)}(m_{23}^2)^{(\max)}d\xi_{12}d\xi_{13}d\xi_{23}\\
&=\frac{(m_{X_1}^2-m_{X_2}^2)^2(m_{X_2}^2-m_{X_3}^2)^2(m_{X_3}^2-m_{X_4}^2)^2}{m_{X_2}^3m_{X_3}^4}\int \mathcal{J}d\xi_{12}d\xi_{23}dq\\
&=\frac{(m_{X_1}^2-m_{X_2}^2)^2(m_{X_2}^2-m_{X_3}^2)^2(m_{X_3}^2-m_{X_4}^2)^2}{m_{X_2}^3m_{X_3}^4}\int \frac{\pi m_{X_3}\cos^{-1}\sqrt{q}}{4m_{X_2}}dq\\
&=\int \frac{\pi(m_{X_1}^2-m_{X_2}^2)^2(m_{X_2}^2-m_{X_3}^2)^2(m_{X_3}^2-m_{X_4}^2)^2}{4m_{X_2}^3m_{X_3}^3}\cos^{-1}(\sqrt{q})dq\\
&=\frac{16\pi\Delta_4^{\max}}{m_{X_2}m_{X_3}}\int \cos^{-1}\sqrt{q}\;dq
\end{aligned} \tag{2.57}$$

Therefore, for the volume inside the contour q, we have

$$V=\frac{16\pi\Delta_4^{\max}}{m_{X_2}m_{X_3}}\frac{1}{2}\Big[\sqrt{q-q^2}+(1-2q)\cos^{-1}\sqrt{q}\Big] \tag{2.58}$$

Chapter Three: Identifying Phase Space Boundaries with Voronoi Tessellations[1]

3.1 Introduction

Voronoi tessellations [7] are useful in a wide variety of fields, from biology [8] to astronomy [9, 10] to condensed matter physics [11]. In high energy physics, they have been used rather sporadically, e.g., as an optional approach to QCD jet-finding and area determination in FastJet [12] and in the model-independent definition of search regions in SLEUTH [13, 14, 15, 16]. In Ref. [17], some of us pointed out that Voronoi methods can be applied directly to the analysis of data from high energy physics experiments, e.g., when trying to detect the presence of a new physics signal in the data or to perform parameter measurements.

In most Voronoi-based approaches, the goal is to use Voronoi tessellations to identify "neighbors" of data points. The tessellation then automatically provides a number of cell-based attributes for each data point. Ref. [17] argued that using the geometric properties of Voronoi cells, and, in particular, functions of the geometric properties of Voronoi cells and their neighbors, gives valuable additional information and can allow for relatively model-independent searches for targeted "features" in the data. As briefly discussed in Ref. [17], a particularly useful application is the study of kinematic edges when investigating cascade decays in new physics models such as supersymmetry (SUSY) [18].

To understand the importance of edge-finding in multidimensional spaces for SUSY mass measurement, we first note that many extensions of the standard model (SM) are characterized by a $\mathbb{Z}_2$ symmetry under which new physics particles (NPPs) are charged but the SM particles are uncharged. Such a symmetry ensures that the lightest NPP will be stable and hence may constitute the dark matter. With the assumption of such a symmetry, a typical collider event involving NPPs proceeds as

[1]Previously published as D. Debnath, J. S. Gainer, C. Kilic, D. Kim, K. T. Matchev, Y.-P. Yang,"Identifying Phase Space Boundaries with Voronoi Tessellations" on The European Physical Journal **C76** 645. Y.-P. Yang analyzed the variable strength analysis and derived the analytic distribution of $\Delta_4/\Delta_{4,\max}$.

follows:

1. NPPs are pair produced.

2. Each NPP goes through a series of (generally two and three-body) decays called a "decay chain". In each decay, an NPP decays to another, lighter, NPP, and one or more SM particles. The NPPs generally have a small intrinsic width compared with their mass. Hence it is generally a good approximation to view the decay chain as consisting of a series of on-shell decays of NPPs.

3. Eventually the lightest particle charged under the $\mathbb{Z}_2$ is reached. It is stable, and, if a dark matter candidate, uncharged and uncolored. Hence it will escape the detector without being detected.

Popular new physics models within this paradigm include SUSY, where the $\mathbb{Z}_2$ symmetry is called "R-parity"; Universal Extra Dimensions (UED) [19], in which the $\mathbb{Z}_2$ symmetry is called "KK-parity"; and Little Higgs models, in which the $\mathbb{Z}_2$ symmetry is called "T-parity" [20].

As the lightest NPP escapes detection, we are not able to determine directly the masses of the *initial* new physics particles produced in the collision, nor the masses of any *intermediate* particles in the decay, as we would if we were studying a resonance decaying to visible particles. However, we *can* determine the masses of the NPPs by studying the distributions of (functions of) the momenta of observed particles [21].

Much effort has gone into determining the best way to actually perform this mass measurement. The simplest methods involve finding an edge or an endpoint in the one-dimensional distribution of the invariant mass of two (or more) reconstructed objects[2] [22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34]. If one is able to measure enough of these kinematic endpoints, it is then possible to solve for the unknown masses, possibly up to discrete ambiguities [35, 36, 37]. This approach naturally evolved into the so-called "polynomial method" [38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52], where one attempts to solve explicitly for the momenta of the invisible particles in a given event, possibly using additional information from prior measurements of kinematic invariant mass endpoints. Since at hadron colliders the longitudinal momenta of the initial state partons are unknown, much effort went into

[2]From a theorist's point of view, those represent Standard Model (SM) particles visible in the detector.

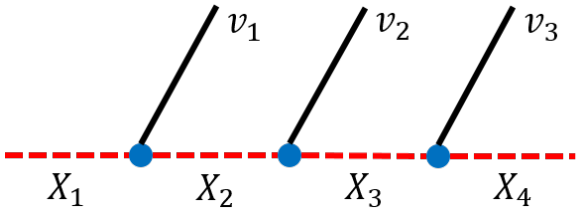


Figure 3.1: Decay process of a heavy resonance X_1 into three visible particles, v_1, v_2 and v_3, along with an invisible particle, X_4, via two on-shell intermediate states, X_2 and X_3. The NPPs, X_i, are denoted by red dashed lines while visible SM particles are denoted by black solid lines.

the development of suitable "transverse variables" [171, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67], which are Lorentz-invariant under longitudinal boosts[3]. In principle, the optimal approach to mass measurement is provided by the so-called "Matrix Element Method" (MEM), [72, 73, 74, 75, 76, 77, 78, 79]. However, its use is often computationally prohibitive, especially when dealing with complicated final states and/or large reducible backgrounds. Many of the approaches described above have been extended in various ways, e.g., the M_{T2} kink method allows the measurement of the mass of the lightest NPP [80, 81, 82, 83, 84, 85, 86, 87], and useful $3+1$-dimensional analogues of the "transverse" invariant mass variables have been suggested [88, 89, 90, 91, 92, 93, 94, 95, 96, 97].

The approach to mass measurement taken here seeks to improve on those described in the existing literature in the following ways:

1. Instead of finding edges or endpoints in the one-dimensional distribution of a single variable, we will attempt to determine the boundary of the signal region in a higher-dimensional phase space. This improves on one-dimensional methods, in increasing greatly the amount of information that can be extracted from the data [1].[4] To be specific, we shall consider the classic SUSY decay chain of three successive two-body decays as shown in Fig. 3.1. From the measured four-momenta, p_i $(i = 1, 2, 3)$, of the three visible particles, v_i, in the decay, one can form three two-body invariant mass combinations, $m_{ij} = \sqrt{(p_i + p_j)^2}$, for

[3]For an alternative approach, see Refs. [68, 69, 70, 71].

[4]At the same time, this approach avoids the drawbacks of some of the more model-dependent, process-specific, and computationally intensive methods, of which the MEM is perhaps the paradigmatic example.

$i < j$. Signal events will then populate the interior of a compact region, $\mathcal{V}_3$, in the three-dimensional phase space, (m_{12}, m_{23}, m_{13}), of invariant masses [98, 99]. The size and the shape of the *two-dimensional* surface boundary of $\mathcal{V}_3$, which we term $\mathcal{S}_2$, contains the complete information about the underlying mass spectrum, $m_{X_i}(i = 1, 2, 3, 4)$, of the NPPs in Fig. 3.1. Therefore, we shall focus on methods for detecting $\mathcal{S}_2$ directly.[5] One can imagine doing this in two ways:

- By defining a kinematic variable which takes the same (constant) value (e.g., zero) everywhere along the phase space boundary.[6] This approach will be discussed below in section 3.3, where we review the relevant variable, Δ_4, introduced in Ref. [1].
- By analyzing the measured density of events in phase space and locating the boundary, $\mathcal{S}_2$, directly using techniques inspired by spatial analyses performed in other fields of science. This will be the main subject of this paper, and will be discussed in sections 3.2.2, 3.3.2, 3.3.3, and 3.4.

2. We build on the idea of Ref. [17] that Voronoi tessellations provide a powerful and model-independent tool for identifying edges (for a brief introduction to Voronoi tessellations, see section 3.2.1 below). While the analysis of Ref. [17] was limited to data in two dimensions, here we extend the procedure to the three-dimensional case and try to delineate the region, $\mathcal{V}_3$, in the phase space of the three variables, (m_{12}, m_{23}, m_{13}). Before tackling a SUSY physics example in section 3.4, we consider several analogous toy examples of increasing complexity in sections 3.2 and 3.3. This helps develop the reader's intuition and motivates some of our analysis choices. Following Ref. [17], in order to select "edge" cells in the Voronoi tessellation of the data, we consider the relative standard deviation[7] (RSD), $\bar{\sigma}_i$, of the volumes of neighboring cells, which is defined as

[5] One can also project the allowed three-dimensional phase space $\mathcal{V}_3$ onto the subspace of two variables, say (m_{12}, m_{13}), and obtain a corresponding *two-dimensional* allowed phase space $\mathcal{V}_2$ whose *one-dimensional* boundary $\mathcal{S}_1$ can be similarly used for mass measurements and disambiguation [98, 36, 37, 63]. With regards to utilizing the full amount of information contained in the data, this approach stands midway between the traditional method of using one-dimensional distributions of single variables and the three-dimensional approach considered here.

[6] Compare to the "singularity coordinate" defined in Ref. [100].

[7] In statistics, the RSD is also known as the coefficient of variation (CV).

follows. Let N_i be the set of neighbors of the i-th Voronoi cell, with volumes, $\{v_j\}$, for $j \in N_i$. The RSD, $\bar{\sigma}_i$, is now defined by

$$\bar{\sigma}_i \equiv \frac{1}{\langle v(N_i) \rangle} \sqrt{\sum_{j \in N_i} \frac{\left(v_j - \langle v(N_i) \rangle\right)^2}{|N_i| - 1}}, \tag{3.1}$$

where we have normalized by the average volume of the set of neighbors, N_i, of the i-th cell

$$\langle v(N_i) \rangle \equiv \frac{1}{|N_i|} \sum_{j \in N_i} v_j. \tag{3.2}$$

The variable defined in eq. (3.1) is a straightforward extension to three dimensions of the "scaled standard deviation" of neighbor areas found to be helpful in Ref. [17].

3. We make crucial use of the recent observation of Ref. [1] that for sufficiently many-body final states there is an enhancement (in fact a singularity) in the phase space density near the boundary, $\mathcal{S}_2$, of the allowed phase space, $\mathcal{V}_3$. Due to the enhancement in the density of signal events near the boundary of phase space, we can alternatively target the boundary points of $\mathcal{V}_3$ as being points in a densely populated region. This motivates us to consider, in addition to $\bar{\sigma}_i$, a second variable related to volume. We choose

$$\bar{v}_i = \frac{v_i}{\langle v(N_i) \rangle}, \tag{3.3}$$

where again we normalize by the average volume (3.2) of the set of neighbors, N_i.

In what follows, we will therefore focus on the two Voronoi-based dimensionless variables (3.1) and (3.3). The former is motivated by the discontinuity in the density of events at the boundary [17], while the latter is motivated by the enhancement in the density of signal events at the boundary [1]. We will find that the judicious combination of these two variables yields a significant increase in sensitivity as compared with either variable in isolation. As a result we find that we are able to identify the boundary, $\mathcal{S}_2$, of the allowed signal phase space, $\mathcal{V}_3$, with a high degree of accuracy, even when the ratio of signal to background events, S/B, is relatively small.

To support these conclusions, as well as to explain in detail to the reader the methods employed, we proceed as follows. In section 3.2 we will provide a brief, but sufficient, review of Voronoi tessellation methods and the use of the geometric properties of Voronoi cells for identifying features in high energy physics data. Section 3.3.1 will review the consideration of events in multi-dimensional phase space, and, especially the observation that, for sufficiently many dimensions, the differential phase space volume is highly peaked near the boundary. Then in sections 3.3.2 and 3.3.3 we study the efficacy of Voronoi methods for finding a densely populated *spherical* boundary in a generalized "phase space", while section 3.4 will examine the application of these methods to an actual benchmark point. We present our conclusions in section 3.5. Throughout our studies, we will use ROC curves to quantify the sensitivity of the variables we define; we briefly review and discuss this approach in Appendix A.

3.2 Voronoi Methods for Finding Boundaries

Voronoi tessellation [101] refers to the procedure, previously proposed by Dirichlet [102] and hinted at by Descartes [103], through which an n-volume containing a set of N_d data points, $\{d_i\}$, is divided into N_d non-overlapping regions, $\{R_i\}$, such that $d_i \in R_i, \forall i$. The boundaries of R_i are chosen such that, for every point in some region R_j, the corresponding data point, d_j, is the nearest data point.

For applications in high energy physics, we consider the data points to be events in a suitably chosen phase space[8]. It is important to make a judicious choice of phase space — on the one hand, it should be of low enough dimensionality to keep the problem tractable in practice, yet the dimensional reduction should not result in the loss of any useful information, e.g., the washing out of interesting features in the underlying phase space distributions. Consider, as an example, the decay chain of Fig. 3.1. In general, the inclusive production of the X_1 resonance will be described by a 9-dimensional phase space, consisting, e.g., of the nine momentum components of the visible particles, v_i. However, three of those degrees of freedom correspond

[8]Here we are following a slightly confusing (but standard) usage and using the term "phase space" to refer to the space of n-tuple values of n observables used to categorize events. This set of observables may not be sufficient to totally specify the kinematics of the event. In section 3.3, "phase space" will have a more precise meaning.

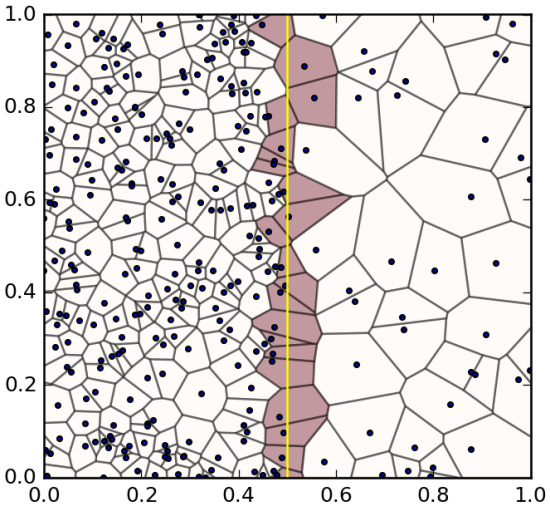


Figure 3.2: $N_{events} = 280$ events distributed according to (3.4) with $\rho = 6$ and the respective Voronoi tessellation. The shaded cells are those crossed by the boundary (vertical yellow line) and are defined to be the "edge cells".

to uninteresting Lorentz boosts of X_1, another three degrees of freedom are angular variables which are sensitive to spins but not the X_i mass spectrum, leaving only the three degrees of freedom relevant to a mass measurement. As already mentioned in the introduction, we can take these three degrees of freedom to be the invariant mass quantities (m_{12}, m_{23}, m_{13}). We shall present the results from our analysis of this physics example in section 3.4, but we first begin with a few toy studies.

3.2.1 Voronoi tessellations in two dimensions

In order to make contact with Ref. [17] and to introduce our notation, we begin by studying several simplified scenarios in two dimensions. In the next section, (3.2.2), we will generalize the approaches taken and the lessons learned here to the case of three dimensions.

3.2.1.1 A linear boundary in two dimensions

In Fig. 3.2, we consider the unit square in the first quadrant ($x \geq 0, y \geq 0$) and simulate $N_{events} = 280$ events (data points) according to the probability distribution

$$f(x, y) = \frac{2}{1+\rho} \left[\rho H(0.5 - x) + H(x - 0.5)\right], \tag{3.4}$$

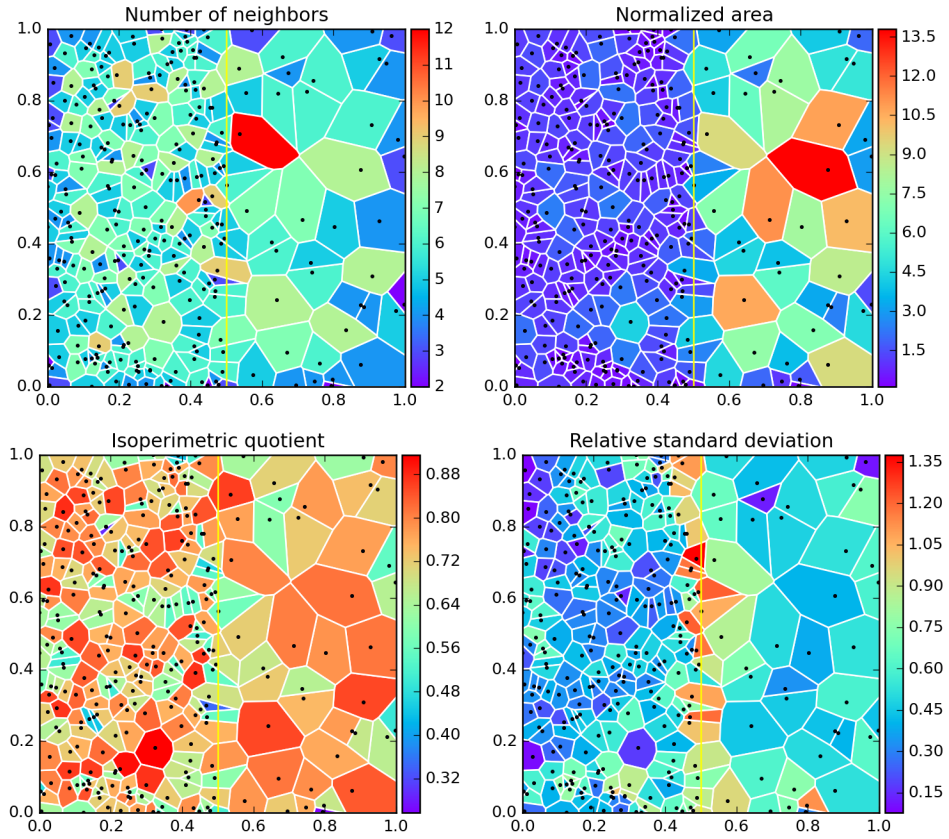


Figure 3.3: The Voronoi tessellation shown in Fig. 3.2, with cells color-coded according to the number of neighboring polygons (upper left), normalized area (3.7) (upper right), isoperimetric quotient (3.8) (lower left), and RSD (3.9) (lower right).

where $H(x)$ is the Heaviside step function and ρ is a constant density ratio, taken in Fig. 3.2 to be $\rho = 6$. The meaning of the distribution (3.4) is very simple: the unit square is divided into two equal halves by the vertical boundary at $x = 0.5$ (the yellow line in Fig. 3.2). Within each half, the density is constant (on average), but the left region is denser by a factor of ρ. This setup produces an edge at $x = 0.5$, where the density changes by a factor of ρ. Our goal will be to detect this edge by tagging the Voronoi cells that are crossed by the boundary line — such cells from now on will be referred to as "edge cells" and in Fig. 3.2 they are shaded in brown. The remaining Voronoi cells away from the edge will be referred to as "bulk" cells, and in Fig. 3.2 they are left white.

The basic idea put forth in Ref. [17] was to study the resulting Voronoi tessellation and identify edge cells from their geometric properties (as well as from the geometric

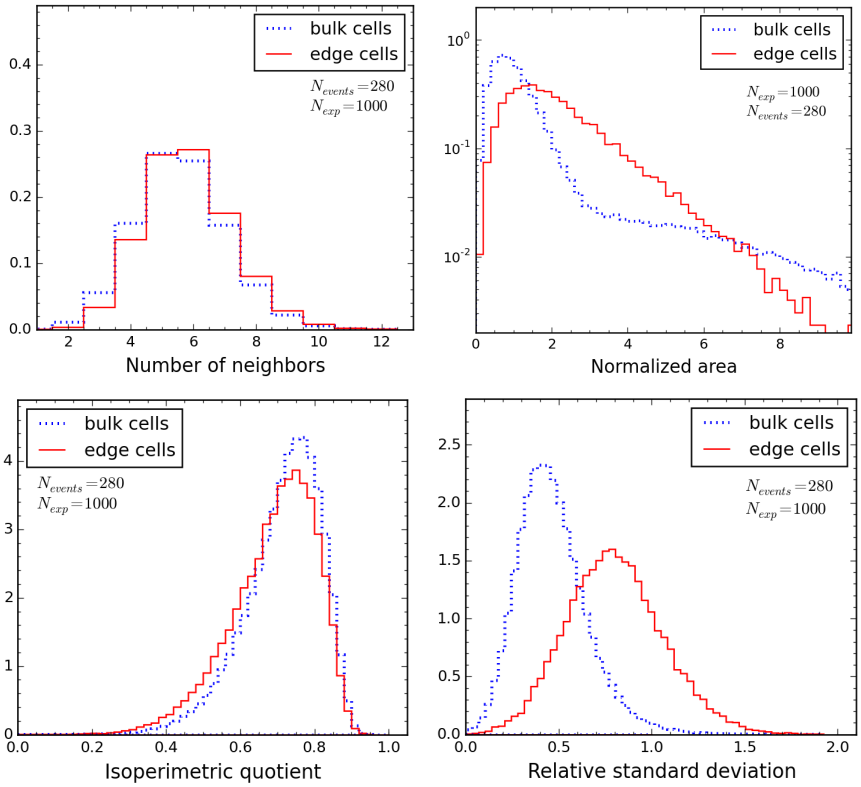


Figure 3.4: Unit-normalized distributions of the four Voronoi cell properties depicted in Fig. 3.3. Blue dotted (red solid) histograms refer to bulk (edge) cells. In order to increase the statistics, we show results from $N_{exp} = 1000$ pseudo-experiments with $N_{events} = 280$ each.

properties of neighboring cells within the immediate vicinity). The Voronoi cells in two dimensions are planar polygons, for which one can investigate the usual geometric properties like number of sides, area, perimeter, etc. Fig. 3.3 shows four possibilities, where the Voronoi cells are color-coded according to the value of the corresponding geometric quantity. Then, in Fig. 3.4, we plot the probability distributions for these geometric quantities separately for the edge cells (red solid lines) and the bulk cells (blue dotted lines). As can be seen in Fig. 3.2, the edge cells, by construction, represent a very small fraction of the total number of Voronoi cells in the tessellation. Thus, in order to increase the statistical precision of the distributions in Fig. 3.4, we generated $N_{exp} = 1000$ pseudo-experiments analogous to the one shown in Fig. 3.2.

In the upper left panels of Figs. 3.3 and 3.4, we study the number of elements, $|N_i|$, in the set of neighbors, N_i. This is equivalent to the number of sides of the i-th

Voronoi polygon. This variable has been studied in the literature [104]: for example, it is known that the Voronoi polygons most commonly have 5 or 6 sides, which is confirmed in Figs. 3.3 and 3.4. There is also a long tail of polygons with many sides, which is conjectured to behave asymptotically as $|N_i|^{-2|N_i|}$ [105]. Indeed, Fig. 3.2 contains an example of a polygon with as many as 12 sides! The upper left panel of Fig. 3.4 demonstrates that, as expected, the $|N_i|$ distributions for bulk and edge cells are rather similar, and are thus not suitable for tagging edge cells [17].

The upper right panels of Figs. 3.3 and 3.4 illustrate a different geometric quantity related to the areas, a_i ($i = 1, 2, \ldots, N_{events}$), of the Voronoi cells. The areas of the Voronoi polygons are meaningful because they provide an estimate of the value of the underlying distribution $f(x, y)$ (3.4) at the corresponding data point (x_i, y_i):

$$f(x_i, y_i) \approx \frac{1}{N_{events} \times a_i}, \tag{3.5}$$

so that $f(x, y)$ is still unit-normalized:

$$\sum_{i=1}^{N_{events}} f(x_i, y_i) \times a_i = \frac{1}{N_{events}} \sum_{i=1}^{N_{events}} \frac{a_i}{a_i} = 1. \tag{3.6}$$

In Figs. 3.3 and 3.4 we choose to normalize the cell areas not locally as in (3.3), but by the *expected* average area in the dense region. Thus, a typical bulk cell to the left of the vertical boundary has normalized area of approximately 1, while a typical bulk cell to the right of the boundary has a normalized area of approximately $\rho = 6$. Note that while we fix the total number of events, N_{events}, the fraction which ends up on one side of the boundary varies — for example, in the single pseudo-experiment depicted in Figs. 3.2 and 3.3, there happen to be 243 events on the left side and 37 events on the right side (to be compared with the expectation of $\rho N_{events}/(\rho + 1) = 240$ events on the left side and $N_{events}/(\rho + 1) = 40$ events on the right side). If the total area is $A = \sum_i a_i$, the expected average size of a bulk cell in the dense region on the left is given by $A(\rho + 1)/(2\rho N_{events})$, hence in the upper right panels of Figs. 3.3 and 3.4 we plot the cell areas, a_i, normalized as

$$\bar{a}_i = \frac{2\rho N_{events}}{\rho + 1} \frac{a_i}{A}. \tag{3.7}$$

The distribution of the Voronoi cell areas when the points have been randomly selected (in a "Poisson process") is not known analytically, and is typically approximated with a three-parameter generalized gamma function, where the parameters are

fitted to results from Monte Carlo simulations [106]. For our purposes, we are not interested in the form of the actual distributions but in the question of whether the distributions for bulk and edge cells show any appreciable differences. As seen in the upper right panel of Fig. 3.4, the area distribution for bulk cells is nearly bimodal; with the normalization (3.7) two peaks are expected near $\bar{a}_i \sim 1$ and $\bar{a}_i \sim \rho = 6$[9]. The area distribution for edge cells, on the other hand, is unimodal, peaking relatively close to $\bar{a}_i \sim 1$. After all, we expect a larger fraction, namely $\rho/(\rho+1)$ of the edge cells, to have their centers on the "dense" side of the boundary and only $1/(\rho+1)$ of the edge cells to have their centers on the "sparse" side of the boundary. A close inspection of Fig. 3.2 confirms this expectation: out of the 21 edge cells, 18 (3) have their centers to the left (to the right) of the vertical boundary, which is consistent with our expectations for $\rho = 6$. In conclusion, it is clear that in this case, the Voronoi cell area by itself is not a very good candidate for an edge-tagging variable [17]. We expect that in the more general situation, where the densities on each side of the boundary are not uniform, this variable will be even more unsuitable.

Having investigated a variable describing the size of the Voronoi polygon, we now examine a variable characterizing the *shape* of the polygon, e.g., the isoperimetric quotient

$$q_i \equiv \frac{4\pi a_i}{p_i^2}, \tag{3.8}$$

where a_i is the area and p_i is the perimeter of the i-th Voronoi polygon. The variable (3.8) is a measure of "roundedness" — it is equal to zero for infinitely thin (pencil-like) polygons, and is equal to 1 in the limit of a perfectly symmetric polygon with infinitely many sides (i.e., a circle). The corresponding results for the isoperimetric quotient are shown in the lower left panels of Figs. 3.3 and 3.4. We observe that the edge cells tend to be slightly more "squashed", but the difference is very minor and not useful in practice.

In a similar vein, one could continue to study other geometric characteristics of a single Voronoi cell, e.g., perimeter, average side length, etc. [107], but, similarly, this is unlikely to lead to any success in identifying edge cells. The reason is that we are trying to detect a discontinuity and therefore we need to study the *relative* properties of cells on both sides of the boundary. One possibility is to compute a

[9] As there are ρ times fewer cells in the low density region, the corresponding peak is ρ times smaller and hence appears to be more of a shoulder than a true peak in the upper right panel of Fig. 3.4. The same considerations hold for the analogous plot in Fig. 3.7.

derivative quantity, e.g., the gradient at each cell location [17]. Another option is to compare the spread in the areas of the neighboring cells, e.g., by computing the RSD, $\bar{\sigma}_i$, of the areas of the cells in N_i (the set of neighbors of the i-th Voronoi polygon) in analogy to (3.1) [17]:

$$\bar{\sigma}_i \equiv \frac{1}{\langle a(N_i)\rangle}\sqrt{\sum_{j\in N_i}\frac{\left(a_j - \langle a(N_i)\rangle\right)^2}{|N_i|-1}}, \tag{3.9}$$

where the normalization now is done by the average *area* of the neighbors

$$\langle a(N_i)\rangle \equiv \frac{1}{|N_i|}\sum_{j\in N_i} a_j. \tag{3.10}$$

The idea is very simple — the neighbors of edge cells are typically quite diverse — some happen to be on the dense side and are therefore relatively small, while others are on the sparse side and are relatively large. As a result, the RSD of neighbor areas for edge cells is expected to be enhanced. On the other hand, for bulk cells, all neighbors are roughly similar, and the RSD of their areas should be small. These expectations are confirmed in the lower right panels of Figs. 3.3 and 3.4. In the temperature plot of Fig. 3.3, the edge cells are clearly distinguished by the different color, and the $\bar{\sigma}_i$ distributions for bulk and edge cells in Fig. 3.4 are visibly displaced from each other. We see that, in agreement with the conclusions from Ref. [17], the RSD, $\bar{\sigma}_i$, is a promising variable for edge detection[10].

3.2.1.2 A circular boundary in two dimensions

Before concluding our discussion in two dimensions, we perform one more toy exercise. In the example of the previous section 3.2.1.1, the boundary was a straight line; in a more realistic situation we will encounter a boundary which is an arbitrary curved line. In anticipation of the physics example discussed in section 3.4, we now consider a two-dimensional example with a curved boundary in the shape of a circle. Instead of the rectangular pattern given by (3.4), we consider the radially symmetric distribution

$$f(\vec{r}) \sim \rho H(1-r) + H(r-1)H(\sqrt{2}-r), \tag{3.11}$$

[10]Ref. [17] also considered a few other variables related to derivatives — the magnitude of the gradient at each data point, the correlation between the directions of the gradients computed at two neighboring cells, the scalar product of the gradients at two neighboring cells, etc. The conclusion, drawn on the basis of ROC curves (see below appendix A) was that the RSD variable, $\bar{\sigma}_i$, was optimal among all those choices.

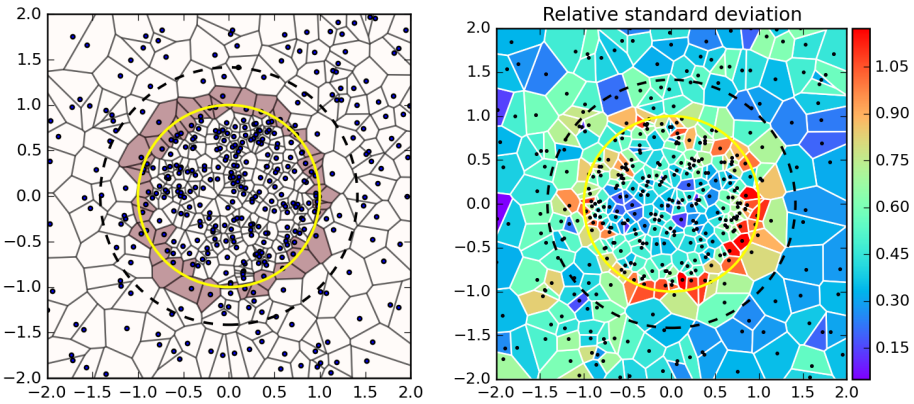


Figure 3.5: The analogue of Fig. 3.2 (left panel) and the analogue of the lower right panel in Fig. 3.3 (right panel), for the radially symmetric distribution (3.11). In order to keep the statistics the same as in Figs. 3.2 and 3.3 we place $N_{events} = 280$ events inside the dashed circle with radius $r = \sqrt{2}$.

where $\vec{r} = (x, y)$ is the position vector in 2 dimensions and $r \equiv |\vec{r}|$ is its magnitude. As in (3.4), the distribution (3.11) describes two regions, the inner region is a unit circle, while the outer region is a hollow disk extending up to $r = \sqrt{2}$ (the circular dashed line in Fig. 3.5). The regions are separated by a circular boundary at $r = 1$, marked with the solid yellow curve in Fig. 3.5. Similarly to the example from section 3.2.1.1, the two regions have equal areas, each region has a constant density, and one region is ρ times denser than the other, see Fig. 3.5. Just as in section 3.2.1.1, we choose $\rho = 6$ and generate $N_{events} = 280$ events according to (3.11); they are distributed so that the ratio of the bulk events on the two sides of the boundary is equal to ρ. Thus, out of the $N_{events} = 280$ events inside the dashed circle with $r = \sqrt{2}$, on average we will have $\rho N_{events}/(\rho + 1) = 240$ events in the dense interior region (the unit circle) and $N_{events}/(\rho + 1) = 40$ events within the sparse exterior hollow disk.[11]

In the left panel of Fig. 3.5, the brown-shaded polygons are by definition the edge cells (those crossed by the yellow circular boundary). The right panel in Fig. 3.5 demonstrates that, once again, the edge cells can be effectively selected by the RSD, $\bar{\sigma}_i$, of the areas of the neighboring cells.

[11] Since our plots are rectangular, in Fig. 3.5 we have extended the exterior region beyond $r = \sqrt{2}$, populating the additional real estate with the same density as the hollow disk. This was done to avoid spurious, but visually distracting empty areas on the plots.

3.2.2 Voronoi tessellations in three dimensions

Since the relevant physics example we treat in section 3.4 is in a three-dimensional phase space, (m_{12}, m_{23}, m_{13}), we shall now generalize our previous discussion to three dimensions. For this purpose, we consider the three-dimensional analogue of (3.11):

$$f(\vec{R}) \sim \rho H(1-R) + H(R-1)H(\sqrt[3]{2}-R), \tag{3.12}$$

where now $\vec{R} = (x, y, z)$ is the position vector in *three* dimensions and $R \equiv |\vec{R}|$. The distribution, (3.12), again describes two regions of constant density, except now the dense region is a three-dimensional spherical core of radius 1. Again, we choose $\rho = 6$ and generate $N_{events} = 4200$ events according to (3.12). The events populate a ball of radius $R = \sqrt[3]{2}$ centered at the origin, $(x, y, z) = (0, 0, 0)$. On average, we expect to have $\rho N_{events}/(\rho+1) = 3600$ events in the core and $N_{events}/(\rho+1) = 600$ events in the outer hollow spherical shell $(1 \leq R \leq \sqrt[3]{2})$.[12]

Figs. 3.6 and 3.7 illustrate this three-dimensional simplified scenario in analogy to Figs. 3.3 and 3.4. Since the Voronoi cells in three dimensions are polyhedra, it is difficult to visualize them on a planar plot. Thus, Fig. 3.6 shows only a slice at a fixed $z = 0$, i.e., an equatorial plane view. The cell boundaries seen in the figure are the intersections of the equatorial plane with the walls of the Voronoi polyhedra. The interiors of those cells are color-coded according to the value of the geometric property (number of faces, volume, etc.) of the corresponding *three-dimensional* polyhedron[13]. For example, the upper left panel in Fig. 3.6 shows that the Voronoi polyhedra typically have a relatively large number of faces (or equivalently, neighbors); the corresponding distribution for bulk cells, shown in the upper left panel in Fig. 3.7, is known to peak at 15 [108]. We also observe that the edge cells are very similar in that regard, i.e., there is no appreciable difference in the number of neighbors as we move across the boundary.

In the upper right panels of Figs. 3.6 and 3.7 we show the corresponding result for the normalized volumes, $\bar{v}_i$, of the Voronoi polyhedra, where, in analogy to (3.7), we scale each volume, v_i, by the expected average volume in the dense core, $\frac{4}{3}\pi(\rho +$

[12] Once again, to avoid misleading voids on the plots, we extend the exterior region beyond $R = \sqrt[3]{2}$, populating this outer region with points having the same density as the hollow sphere.

[13] We caution the reader that the color bars in Fig. 3.6 refer to the three-dimensional Voronoi polyhedra and not the polygons seen in the plots — the latter result from the intersection with the equatorial plane, and, in general, have different properties.

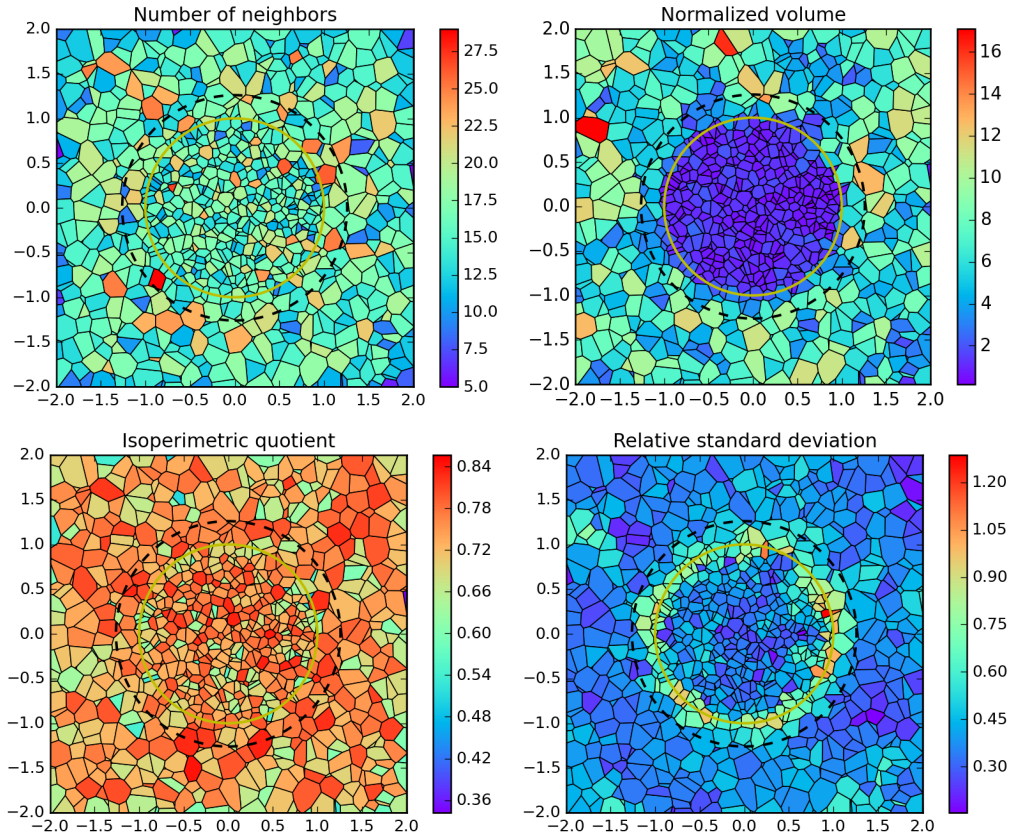


Figure 3.6: Two-dimensional slices at $z = 0$ through phase space for the three-dimensional toy example studied in section 3.2.2. We distribute $N_{events} = 4200$ points according to the three-dimensional probability distribution (3.12) within a sphere of radius $\sqrt[3]{2}$ centered at the origin $(x, y, z) = (0, 0, 0)$. The Voronoi tessellation is done *before* taking the two-dimensional slice, i.e., the cell boundaries seen on these four plots are obtained by intersecting the three-dimensional Voronoi cell boundaries with the plane at $z = 0$. The yellow circle marks the boundary of the dense core. The resulting cells in the two-dimensional slice are color coded by a certain attribute of the corresponding three-dimensional Voronoi cell: number of neighbors (upper left); normalized volume (upper right); isoperimetric ratio (3.13) (lower left), and RSD of the neighboring volumes (3.1) (lower right).

$1)/(\rho N_{events})$. As expected, the distribution for bulk cells is bimodal, while edge cells behave somewhat similarly to the interior bulk cells (as we already saw in the two-dimensional example of section 3.2.1.1).

In the lower left panels of Figs. 3.6 and 3.7 we plot the analogous "isoperimetric

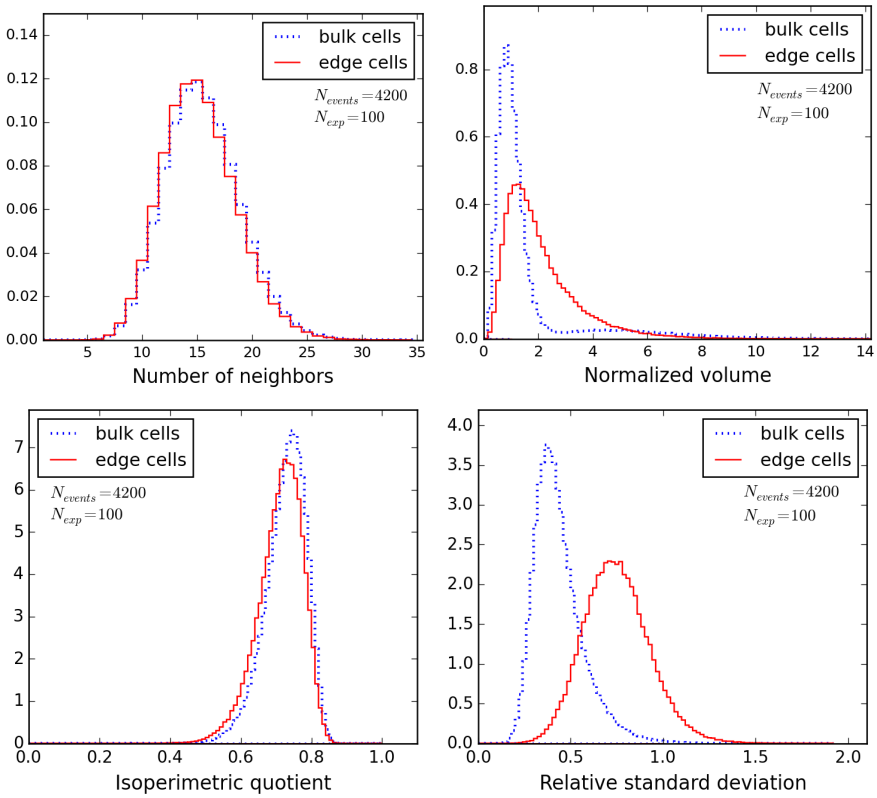


Figure 3.7: The same as Fig. 3.4, but for the three-dimensional toy example depicted in Fig. 3.6. Edge cells are defined to be those Voronoi cells which are crossed by the boundary of the unit sphere ($r = 1$).

quotient" for the three-dimensional case,

$$Q_i \equiv \frac{6\sqrt{\pi}\, v_i}{s_i^{3/2}}, \tag{3.13}$$

where v_i (s_i) is the volume (surface area) of the Voronoi polyhedron and the normalization is chosen so that $Q_i = 1$ for a perfect sphere. Figs. 3.6 and 3.7 show that the shapes of the Voronoi polyhedra, as measured by (3.13), are very similar in the two bulk regions and not much different near the boundary either.

This leaves us with the RSD, $\bar{\sigma}_i$, of the volumes for the set of neighbors, N_i. This quantity was already defined in (3.1) and our results are shown in the lower right panels of Figs. 3.6 and 3.7. We see that $\bar{\sigma}_i$ can efficiently identify edge cells; the circular boundary is clearly seen in the lower right plot of Fig. 3.6. The $\bar{\sigma}_i$ distributions for bulk and edge cells are quite distinct, as shown in Fig. 3.7. Thus we verify that $\bar{\sigma}_i$

remains a promising variable for edge detection beyond the two-dimensional examples studied in Ref. [17].

3.3 Phase Space Considerations

While two and three-body phase space is discussed at length in most standard lectures and textbooks on quantum field theory, a Lorentz-invariant formulation of the general case with an arbitrary number of final state particles is often omitted. This is in part because processes with more than three final state particles can, in almost all circumstances, be analyzed as a sequence of on-shell production and decay stages and in part because the level of formalism required to describe the general case is significantly more involved. Nevertheless, as was shown in Ref. [1], even when a cascade decay proceeds through multiple on-shell stages, treating the phase space in its fully-differential form captures important correlations that cannot be inferred from more traditional one-dimensional observables such as kinematic edges and endpoints. In this context, we briefly review the geometry of four-body phase space,[14] concentrating on the equation describing the boundary of the kinematically available region and on the differential volume element. In section 3.3.2, we shall apply kinematic features obtained from the phase space considerations in section 3.3.1 to our uniform sphere example from section 3.2.2 and use the resulting toy example to study our Voronoi methods for three-dimensional data.

3.3.1 Review of the four-body phase space of a cascade decay

Let us consider the process where a heavy resonance, X, decays into four particles. We first focus on presenting the form of four-body phase space in full generality and will further specialize to the case where the decay proceeds via three-step two-body cascade decays as in Fig. 3.1. Following the argument in Ref. [6], we begin by introducing a 4×4 matrix defined as

$$\mathcal{Z} = \{z_{ij}\} \quad \text{with} \quad z_{ij} = p_i \cdot p_j \,, \tag{3.14}$$

[14]The description of more than four-body phase space can be given in a very similar fashion although there are additional subtleties due to non-linear constraints. We refer the interested reader to Ref. [6].

where the $\{p_i\}$ denote four momenta of final state particles, including the NPP X_4.[15] We then define the characteristic polynomial of $\mathcal{Z}$ as

$$\det\left[\lambda I_{4\times 4} - \mathcal{Z}\right] \equiv \lambda^4 - \left(\sum_{i=1}^{4} \Delta_i \lambda^{4-i}\right) = 0, \tag{3.15}$$

where λ represents the relevant eigenvalues and the Δ_i identify the coefficients of the above polynomial. Specifically, one finds that $\Delta_1 = \mathrm{Tr}[\mathcal{Z}] = \sum_{i=1}^{4} m_i^2$ and $\Delta_4 = -\det[\mathcal{Z}]$. It turns out that the kinematically allowed region is given by $\Delta_{1,2,3,4} > 0$ [6]; the boundary of this region is formed by

$$\Delta_4 = 0, \qquad \Delta_{1,2,3} > 0\,. \tag{3.16}$$

What makes four-body (and higher) phase space particularly interesting is the form of the volume element. In terms of $m_{ij}^2 = (p_i + p_j)^2 = 2z_{ij} + m_i^2 + m_j^2$, the four-body phase space, Π_4, can be written as

$$d\Pi_4 = \left(\prod_{i<j} dm_{ij}^2\right) \frac{8}{(4\pi)^{10} M_X^2 \Delta_4^{1/2}}\, \delta\left(\sum_{i<j} m_{ij}^2 - \left(M_X^2 + 2\sum_{i=1}^{4} m_i^2\right)\right), \tag{3.17}$$

where the normalization has been chosen to reproduce the PDG convention [109] for the well-known expression with non-Lorentz invariant quantities

$$d\Pi_4 = \delta\left(p_X - \sum_{i=1}^{4} p_i\right) \prod_{i=1}^{4} \frac{d^3 p_i}{(2\pi)^3 2E_i}. \tag{3.18}$$

We remark that $d\Pi_4$ is inversely proportional to Δ_4, and, given the fact that $\Delta_4 = 0$ defines the kinematic boundary, as in (3.16), the phase space has a singular structure near $\Delta_4 = 0$. While being an integrable singularity, this implies that events are more likely to be populated close to the boundary rather than far away from it. This observation is ideal for mass measurements which ultimately rely on the determination of this phase space boundary.

Given the generic formalism for the phase space with four particles in the final state, we now specialize to the case where the decay proceeds through the three consecutive two-body decays shown in Fig. 3.1. The X_i's are NPPs represented by red dashed lines, while the v_i's are SM particles represented by black solid lines. For

[15] Note that in this section, Latin indices refer to the final state particles and not data points (Voronoi cells).

simplicity, we assume that all SM particles are mass *less* unless specified otherwise. $X_{1,2,3}$ are assumed to be narrow resonances, while X_4 is collider-stable and invisible.

We point out that the presence of the intermediate particles does *not* affect the enhancement near the boundary of phase space discussed above. Within the narrow width approximation, each internal propagator squared can be replaced by a delta function, whose argument is linear in the z_{ij} or, equivalently, in the m_{ij}^2 variables. Therefore, integrating over those delta functions does *not* introduce any non-trivial Jacobian factors which would ruin the enhancement.

To quantify the enhancement near the boundary for this event topology, we derive the analytic form of the Δ_4 probability distribution and show that it is completely independent of m_{X_i} for the massless limit, i.e., $m_{v_i} = 0$. We start by writing Δ_4 in terms of the experimental observables $m^2_{v_i v_j}$ which are denoted by m_{12}^2, m_{13}^2, and m_{23}^2. These dimensionful variables can be traded for dimensionless, unit-normalized variables ξ_{ij} as

$$m_{ij}^2 \equiv \xi_{ij}\, m_{ij,\max}^2, \tag{3.19}$$

where $0 \leq \xi_{ij} \leq 1$ and the maximal values, $m_{ij,\max}^2$, are given by the well-known kinematic endpoint formulae (see, e.g., [24]):

$$m_{12,\max}^2 = \frac{(m_{X_1}^2 - m_{X_2}^2)(m_{X_2}^2 - m_{X_3}^2)}{m_{X_2}^2}, \tag{3.20}$$

$$m_{13,\max}^2 = \frac{(m_{X_1}^2 - m_{X_2}^2)(m_{X_3}^2 - m_{X_4}^2)}{m_{X_3}^2}, \tag{3.21}$$

$$m_{23,\max}^2 = \frac{(m_{X_2}^2 - m_{X_3}^2)(m_{X_3}^2 - m_{X_4}^2)}{m_{X_3}^2}. \tag{3.22}$$

We also trade the dimension-8 quantity Δ_4 for a dimensionless and unit-normalized quantity q defined as

$$\Delta_4 \equiv q\, \Delta_{4,\max}, \quad 0 \leq q \leq 1. \tag{3.23}$$

Here the maximum value of Δ_4 is given by

$$\Delta_{4,\max} = \left(\frac{(m_{X_1}^2 - m_{X_2}^2)(m_{X_2}^2 - m_{X_3}^2)(m_{X_3}^2 - m_{X_4}^2)}{8 m_{X_2} m_{X_3}} \right)^2. \tag{3.24}$$

As shown in Ref. [1], for any given set of masses, $\{m_{X_i}\}$, in this topology, the proba-

bility of obtaining any given event near the point, $\left\{m_{ij}^2\right\}$, is expressed as

$$dP = \frac{1}{4\pi m_{X_1}^2}\left(1-\frac{m_{X_2}^2}{m_{X_1}^2}\right)^{-1}\left(1-\frac{m_{X_3}^2}{m_{X_2}^2}\right)^{-1}\left(1-\frac{m_{X_4}^2}{m_{X_3}^2}\right)^{-1}\frac{H\left(\Delta_4\right)}{\sqrt{\Delta_4}}dm_{12}^2\,dm_{13}^2\,dm_{23}^2, \tag{3.25}$$

or equivalently, in terms of the dimensionless quantities, ξ_{ij} and q, defined in (3.19) and (3.23), this can be rewritten as

$$dP = \frac{2}{\pi}\frac{m_{X_3}}{m_{X_2}}\frac{H\left(q\right)}{\sqrt{q}}d\xi_{12}\,d\xi_{13}\,d\xi_{23}. \tag{3.26}$$

Here $H(x)$ is the usual Heaviside step function.

Obviously, the expression in eq. (3.26) diverges for $q \to 0$, as expected from the general discussion earlier, and has a non-zero finite value at $q_{\max} = 1$. In order to visualize the enhancement near $q \sim 0$, it is useful to partition the probability density in (3.26) into two components: a flat piece, proportional to 1, and an enhanced piece, containing the $q^{-1/2}$ singularity:

$$\frac{dP}{dV_\xi} \sim \frac{1}{\sqrt{q_{\max}}} + \left(\frac{1}{\sqrt{q}} - \frac{1}{\sqrt{q_{\max}}}\right) = 1 + \left(\frac{1}{\sqrt{q}} - 1\right), \tag{3.27}$$

where $dV_\xi \equiv d\xi_{12}\,d\xi_{23}\,d\xi_{13}$ is a shorthand notation. If events were uniformly distributed over the entire phase space in ξ_{ij}, their probability density would simply be proportional to the first (constant) term in eq. (3.27). Hence, all non-trivial effects in the phase space density distribution are due to the second term (inside the parentheses) in (3.27).

Fig. 3.8 helps us develop some useful intuition about the probability distribution (3.27). The upper left panel shows a scatter plot of physical events in the dimensionless ξ_{ij}-space, generated according to (3.27). We used a mass spectrum of $(m_{X_1}, m_{X_2}, m_{X_3}, m_{X_4}) = (500, 350, 200, 100)$ GeV. The events populate a compact region whose shape has been likened to that of a "samosa" [99]. Since it is difficult to visualize the enhancement near the phase space boundary in this three-dimensional view, in the next three panels of Fig. 3.8 we take a few slices at fixed ξ_{13}: $\xi_{13} = 0.25$ (upper right), $\xi_{13} = 0.5$ (lower left), and $\xi_{13} = 0.75$ (lower right). For each slice at a fixed ξ_{13}, we show all data points whose m_{13}^2 values fall within 0.5 GeV^2 of the nominal value for that slice, i.e., within $\xi_{13} m_{13,max}^2 \pm 0.5$ GeV^2. Then we project those points onto the plane of ξ_{12} vs. ξ_{23} and divide them into two (color-coded) groups. The

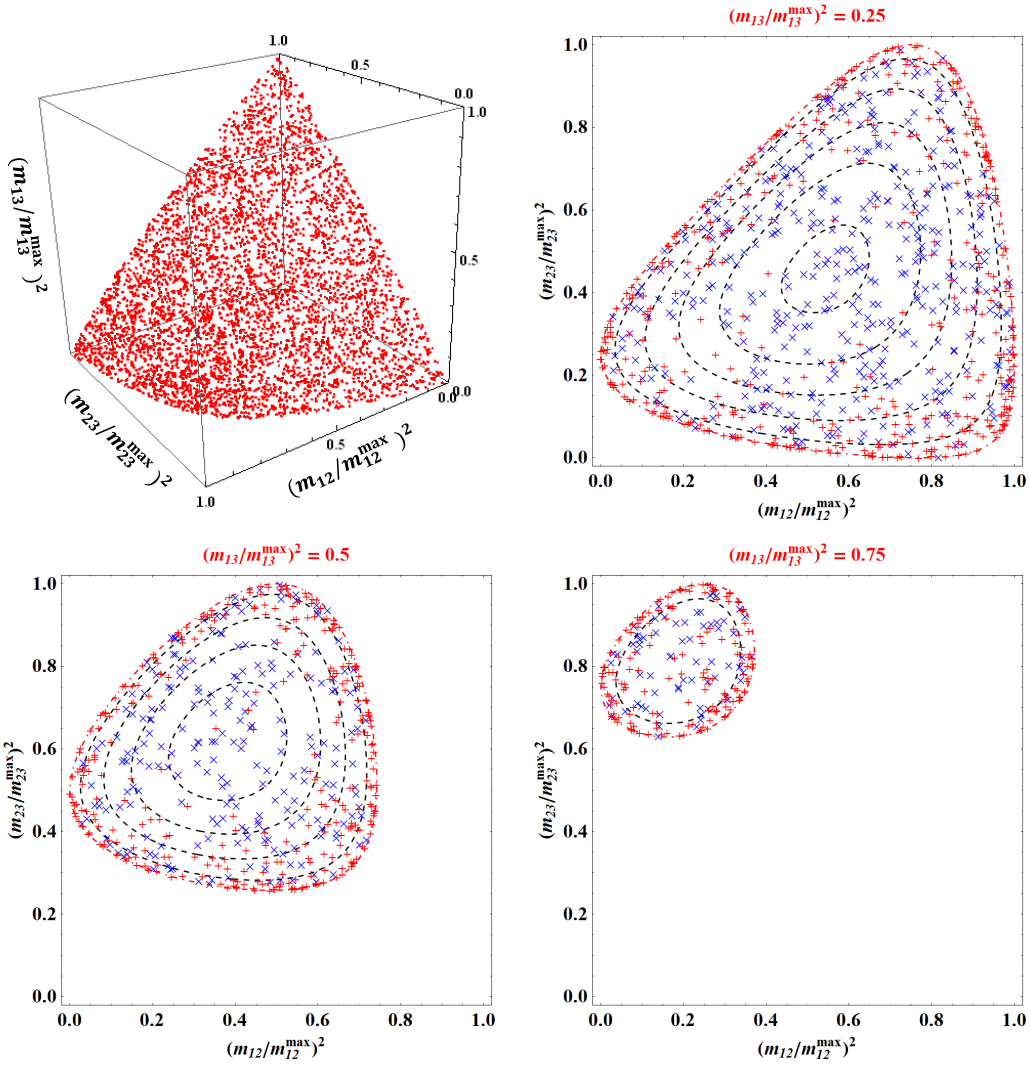


Figure 3.8: The phase space structure implied by eq. (3.25). The data points are generated with the event topology in Fig. 3.1, using a constant matrix element. The mass spectrum is $(m_{X_1}, m_{X_2}, m_{X_3}, m_{X_4}) = (500, 350, 200, 100)$ GeV. A three-dimensional scatter plot (upper left) and three phase space slices at fixed ξ_{13}: $\xi_{13} = 0.25$ (upper right), $\xi_{13} = 0.5$ (lower left), and $\xi_{13} = 0.75$ (lower right). The red dot-dashed (outermost) curve is the contour for $\Delta_4 = 0$, while the black dashed curves correspond to Δ_4 contours for 10%, 30%, 50%, 70%, and 90% of $\Delta_{4,\max}$. The data points which would have emerged via the flat component in eq. (3.27) are represented by blue "$\times$" symbols, whereas the data points from the remaining enhanced component $\sim \frac{1}{\sqrt{q}} - 1$ are represented by red "+" symbols.

data points which would have emerged from the flat piece in (3.27) are denoted with blue "$\times$" symbols, whereas the points arising from the enhanced piece in (3.27) are

identified by red "+" symbols. In addition, we also show several theoretical contours of constant Δ_4 values, starting with the outermost red dot-dashed curve at $\Delta_4 = 0$ (i.e., $q = 0$) representing the phase space boundary. The internal, black dashed curves mark the contours for $\Delta_4 = 0.1\,\Delta_{4,\max}$, $\Delta_4 = 0.3\,\Delta_{4,\max}$, $\Delta_4 = 0.5\,\Delta_{4,\max}$, $\Delta_4 = 0.7\,\Delta_{4,\max}$, and $\Delta_4 = 0.9\,\Delta_{4,\max}$, respectively. Note that some of these contours are absent from the bottom panels because the relevant hyper-surfaces, corresponding to large Δ_4 values do not intersect those slices.

Comparing the densities of red and blue data points, we get an idea about the effect of the enhancement in the vicinity of the phase space boundary. The blue points are more or less uniformly distributed, which is by design. In contrast, the distribution of red points is highly irregular, and their density peaks at the phase space boundary. For a more quantitative understanding, we derive the analytic expression for the probability density function in q and obtain [110]

$$\frac{dP}{dq} = \frac{\arcsin(\sqrt{1-q})}{2\sqrt{q}}. \tag{3.28}$$

As previously advertised, this probability density function is completely independent of all $\{m_{X_i}\}$ and is enhanced near $q \approx 0$. In other words, the fraction of events that lie in a fixed q-interval is *universal*, and it is enhanced near the boundary of the phase space region. For example, roughly 5% of events have $q \leq 10^{-3}$, i.e., less than 0.1% of $\Delta_{4,\max}$.

In Fig. 3.9, we plot the distribution of the q variable from (3.23), taking 20,000 events out of the same event sample as the one used for Fig. 3.8. If we define any phase space point whose q value is less than 5% of $q_{\max}$ as a boundary point, we find that $\sim$ 33% of the events are then categorized as boundary points. The red histogram represents the q distribution with respect to the full data set; the black dashed, vertical line denotes the location of $0.05\,q_{\max}$. The black solid curve shows the theoretical prediction from eq. (3.28). One can easily see that the q distribution (red histogram) is fully consistent with the theory expectation. Indeed, the value of q (or equivalently, Δ_4) is not an experimental observable, since it requires a model assumption (the input of a mass spectrum for X_i). What is needed then is a practical way of tagging the boundary data points with such low values of q by some other means; we employ Voronoi tessellations as an available tool. We Voronoi tessellate our phase space using the full data set. If a given Voronoi cell has vertices on both

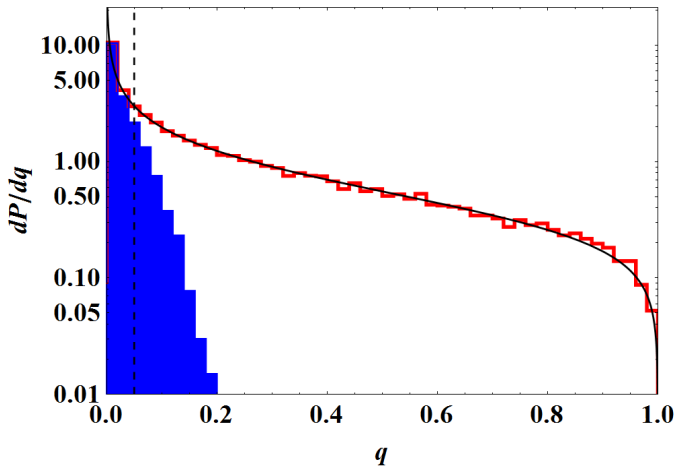


Figure 3.9: Probability density distribution in the q variable (3.23) using the same event sample as in Fig. 3.8. The blue shaded histogram is contributed by the boundary data points which are tagged by the Voronoi tessellation. The black solid curve is the theory prediction for dP/dq in eq. (3.28).

sides of the "samosa" surface defined by $q = 0$ (or equivalently, $\Delta_4 = 0$), then the associated data point is tagged as a boundary point (recall the definition of "edge" cells in Fig. 3.2). The contribution from the boundary points extracted with the above algorithm is shown by the blue shaded histogram in Fig. 3.9, which we find represents $\sim 38\%$ of the events in the sample. As Fig. 3.9 demonstrates, the set of boundary cells which can be tagged by placing a cut on q is essentially the same as the set of boundary cells identified with the Voronoi tessellation. In what follows, therefore, instead of using the variable q, which is experimentally inaccessible, we shall focus on the Voronoi cells belonging to the blue-shaded histogram in Fig. 3.9 and try to develop a tagging method based on their geometric properties, since they *are* experimentally observable.

3.3.2 Density-enhanced sphere boundaries

Inspired by the behavior of the phase space density near the boundary, we deform the density of data points from the sphere example considered previously in section 3.2.2. Performing Voronoi tessellations and studying the properties of the resulting Voronoi cells, we can develop our insight on what is expected from physical examples. Note that Δ_4 vanishes on the phase space boundary and takes its maximum somewhere in

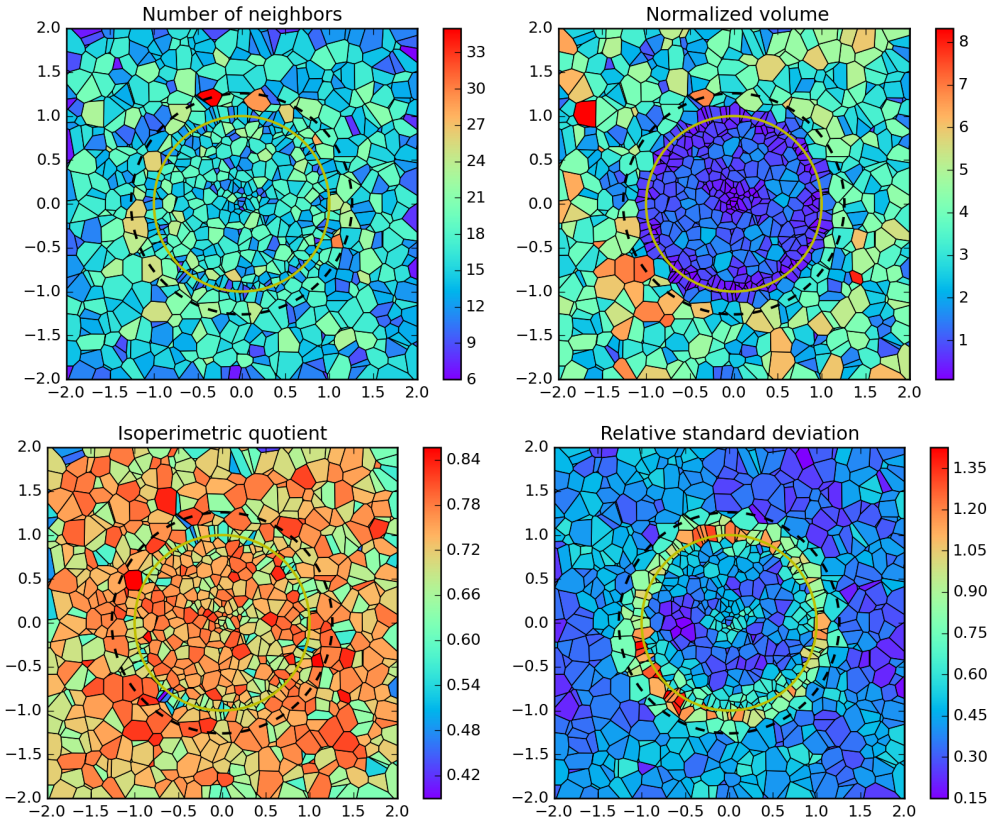


Figure 3.10: The same as Fig. 3.6, but for a toy example in which the dense core has a non-uniform distribution given by (3.30).

the bulk. In other words, the Δ_4 value increases as the distance between the data point of interest and the boundary surface increases (see also contours in Fig. 3.8). Although specifying the value of distance does *not* determine the Δ_4 value, it turns out that there exists a positive correlation between the two quantities [110]. To proceed, we make the simplifying *Ansatz* that the distribution of the data points inside a unit sphere depends only on the radius, R, with an enhancement at $R = 1$. Motivated by the form of the probability density in (3.26), we introduce the following volume density function for the data points inside the unit sphere

$$\frac{dP}{dV} \sim \frac{1}{\sqrt{1-R}}\,. \tag{3.29}$$

Now in analogy to (3.12), we consider the three-dimensional distribution

$$f(\vec{R}) \sim \frac{\rho}{\sqrt{1-R}} H(1-R) + H(R-1)H(\sqrt[3]{2}-R). \tag{3.30}$$

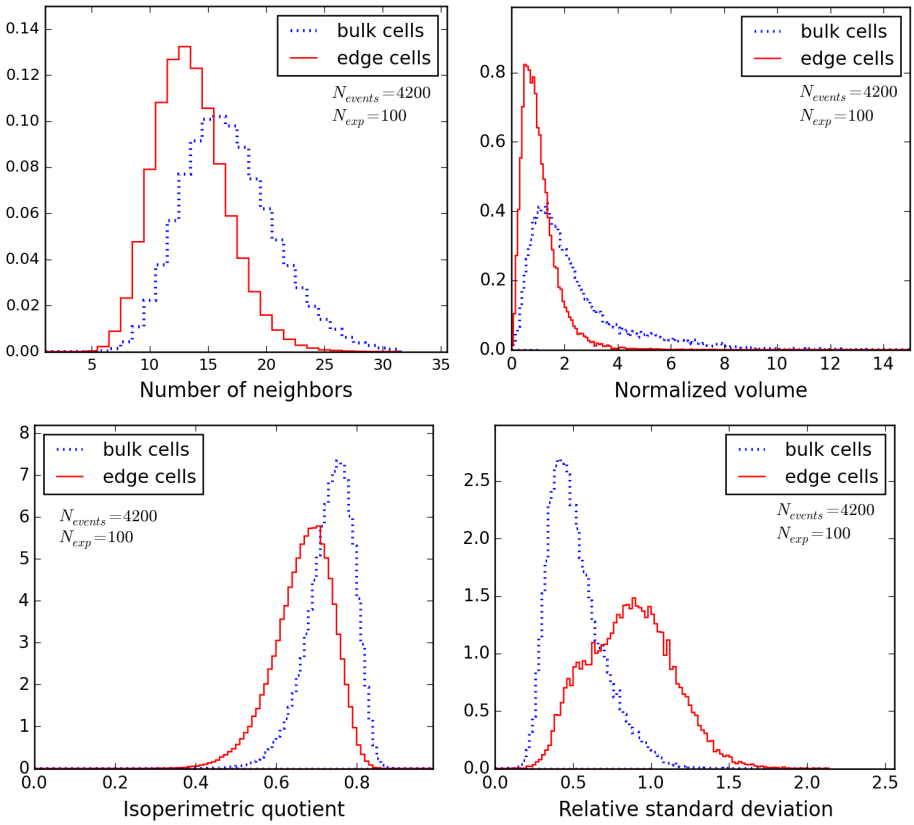


Figure 3.11: The same as Fig. 3.7, but for the example shown in Fig. 3.10.

Following the example from section 3.2.2, we again take the density ratio $\rho = 6$ and generate $N_{events} = 4200$ events according to (3.30). Our results are shown in Figs. 3.10 and 3.11, which are the analogues of Figs. 3.6 and 3.7, respectively.

We see that, in principle, all four variables plotted in Fig. 3.10 show some potential for discriminating edge cells. For example, a careful inspection of the lower left panel of Fig. 3.10 reveals that the edge cells appear somewhat elongated, which results in a lower isoperimetric quotient, as confirmed by the lower left panel in Fig. 3.11. On the other hand, due to the density enhancement near the boundary, we would also expect the edge cells to have smaller normalized volumes. This expectation is also confirmed — in the upper right panels of Figs. 3.10 and 3.11. Finally, the lower right panels of Figs. 3.10 and 3.11 again demonstrate that the RSD of the neighboring areas is a good discriminator, in agreement with our observations from the earlier toy examples. In order to compare the performance of the four variables investigated in Figs. 3.10 and 3.11, we use the concept of a ROC curve, which is reviewed in Appendix A.

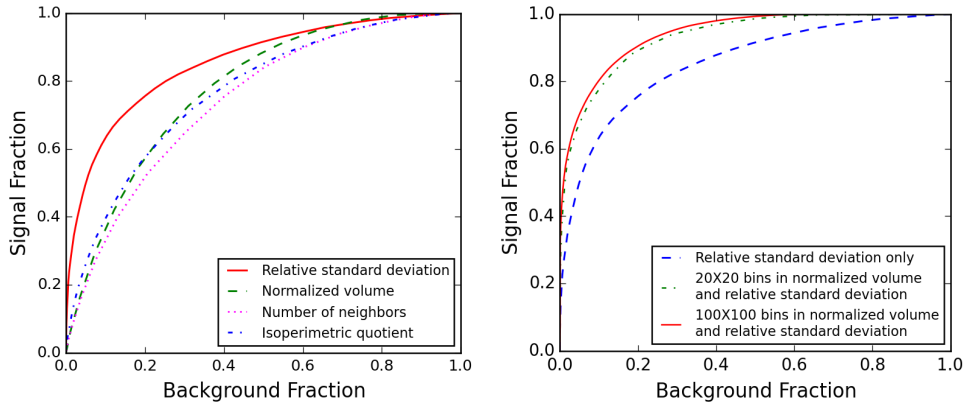


Figure 3.12: ROC curves for the toy example depicted in Fig. 3.10. Left: The four different ROC curves resulting from each of the four variables shown in Fig. 3.11. Right: Improved ROC curves with optimal two-dimensional cuts in the $(\bar{v}, \bar{\sigma})$ plane as illustrated in Fig. 3.13: with 20×20 binning (green dot-dashed) or 100×100 binning (dotted blue). The blue dashed line is the ROC curve based on the $\bar{\sigma}$ variable alone and is identical to the solid red line in the left panel.

3.3.3 Finding density-enhanced sphere boundaries with Voronoi tessellations

We now analyze the example of a sphere with an enhanced density near the boundary considered in section 3.3.2, in terms of ROC curves. In the left panel of Fig. 3.12, we show the ROC curve for each of the four variables depicted in Figs. 3.10 and 3.11: number of neighbors (magenta dotted line), normalized volume (green dashed line), isoperimetric quotient (blue dot-dashed line) and RSD of neighbor areas (red solid line). We observe that the RSD outperforms the other three variables, in agreement with the conclusions from [17] for the two-dimensional case. However, the other three variables also have a certain degree of discriminating power, as seen in Fig. 3.11. The natural question then is how much additional sensitivity can be gained by considering not just one, but *two* variables simultaneously. We studied the correlations between the RSD, $\bar{\sigma}$, and each of the other three variables, and generally find that they are *not* perfectly correlated. (This makes sense intuitively because the RSD is computed from the neighbor set, N_i, while the other three variables are properties of the individual cell.) We concluded that, among the three options, the normalized volume, $\bar{v}$, is the most promising, since it appears least correlated with $\bar{\sigma}$. Therefore, we expect that the sensitivity will improve once we incorporate the normalized volume, $\bar{v}$, in the

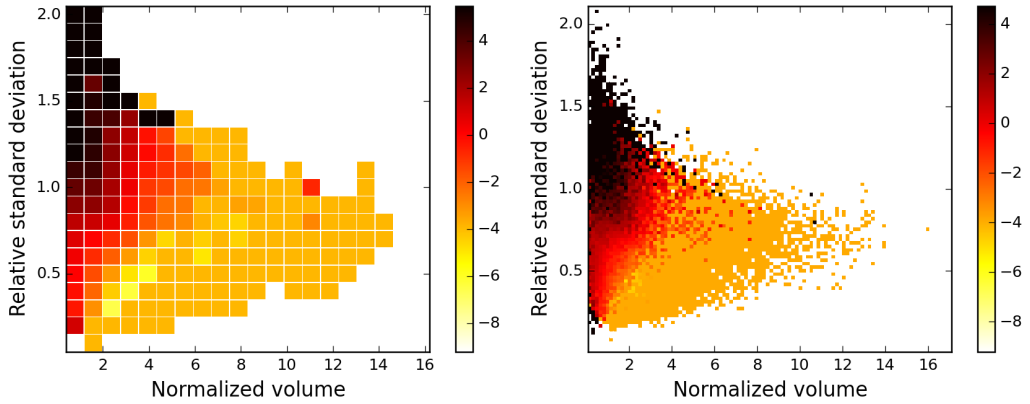


Figure 3.13: Two-dimensional histograms of the expected signal to background ratio in the $(\bar{v}, \bar{\sigma})$ plane: for 20×20 bin (left) and 100×100 bins (right). The ROC curves in the right panel of Fig. 3.12 were built by successively cutting away the bin with the lowest signal-to-background ratio among all remaining bins. (Alternatively, one could start from zero and successively keep adding the bin with the highest signal-to-background ratio among all remaining bins.)

analysis. This expectation is confirmed in the right panel of Fig. 3.12, where we show "improved" ROC curves based on binning in both $\bar{v}$ and $\bar{\sigma}$. The procedure, illustrated in Fig. 3.13, is as follows. We consider the $(\bar{v}, \bar{\sigma})$ plane divided into 20×20 bins (left panel of Fig. 3.13) or 100×100 bins (right panel of Fig. 3.13). We expect the signal, i.e., the boundary Voronoi cells, to populate the bins with small volume and relatively large RSD, while the background, i.e., the bulk cells, are distributed more uniformly throughout the $(\bar{v}, \bar{\sigma})$ plane. In order to build the optimal ROC curve, we need to determine the signal to background ratio, S/B, in each bin, and design the cuts so that we remove successively the bins with the smallest S/B. The bins in Fig. 3.13 are color-coded according to the corresponding value of $\log(S/B)$[16]. Given the finite statistics, there are bins which have no events (neither signal nor background); they are left uncolored. For definiteness, the bins which have some signal events, but no background events, are assigned the same value as the maximal $\log(S/B)$ value among the bins containing both signal and background events. Similarly, the bins which had some background events, but no signal events, were assigned the same value as the minimal $\log(S/B)$ value among the bins containing both signal and background events.

[16] "log" refers to the natural logarithm throughout this work.

Fig. 3.13 shows that, as expected, the bins with the largest S/B (colored in black) are located in the upper left corner of the plot, corresponding to small $\bar{v}$ and large $\bar{\sigma}$. The spread in the cluster of black-colored bins is indicative of the gain in sensitivity due to simultaneous consideration of the two variables, $\bar{v}$ and $\bar{\sigma}$. According to the right panel of Fig. 3.12, the bulk of the gain is already obtained with a 20×20 grid; increasing the number of bins 25 times to a 100×100 grid does not lead to substantial improvement. Therefore, in practice, one might want to consider grids of even smaller size, especially since the true ranking of the bins in terms of S/B depends on the parameter values, e.g., the density enhancement on the boundary and the value of ρ, which are not known *a priori*. This is why when we consider the physics example in the next section, we shall utilize a smaller grid of 15×15 bins in the $(\bar{v}, \bar{\sigma})$ plane (see Fig. 3.16).

3.4 Finding Phase Space Boundaries with Voronoi Tessellations

We now use Voronoi tessellations to find the phase space boundary for SUSY events at the 14 TeV LHC. We consider the $2 + 2 + 2$ topology from Fig. 3.1, where, as usual, a (left-handed) squark $X_1 = \tilde{q}$ undergoes a cascade decay through a heavy neutralino, $X_2 = \tilde{\chi}_2^0$; a slepton, $X_3 = \tilde{\ell}$; and a light neutralino, $X_4 = \tilde{\chi}_1^0$. As in Ref. [1], we consider the production of a squark in association with a neutralino LSP ($\tilde{\chi}_1^0$). Events were generated with MadGraph5 [111]. The mass spectrum that we used was $m_{\tilde{q}} = 350$ GeV, $m_{\tilde{\chi}_2^0} = 300$ GeV, $m_{\tilde{\ell}} = 250$ GeV, and $m_{\tilde{\chi}_1^0} = 200$ GeV. [17]

The particles visible in the detector are: a quark jet $v_1 = j$, a "near" lepton $v_2 = \ell_n$, and a "far" lepton $v_3 = \ell_f$. The relevant phase space is then $(m_{12}, m_{23}, m_{13}) \equiv (m_{j\ell_n}, m_{\ell\ell}, m_{j\ell_f})$. For SUSY signal events, each of these three variables exhibits an

[17] Despite the relatively low squark mass, this study point does not seem to be ruled out by the current LHC data. Since $\tilde{\chi}_1^0$ is mostly Bino, the cross-section for squark-neutralino associated production is suppressed by the left-handed squark hypercharge, and is only ~ 5 fb at 8 TeV. Furthermore, there is no dedicated search for such asymmetric event topologies (squark-LSP production). If we nevertheless test against a standard SUSY search, e.g. a signature with a same-flavour opposite-sign lepton pair, jets and large missing transverse momentum [112], we find a rather low selection efficiency ($\lesssim 1\%$) due to the softness of the visible decay products and the tendency of the two final state neutralinos to be back to back, thus reducing the amount of missing transverses energy.

upper kinematic endpoint. The three endpoint values are given by eqs. (3.20-3.22):

$$m^2_{j\ell_n,max} = 9931 \text{ GeV}^2, \tag{3.31}$$

$$m^2_{\ell\ell,max} = 9900 \text{ GeV}^2, \tag{3.32}$$

$$m^2_{j\ell_f,max} = 11700 \text{ GeV}^2. \tag{3.33}$$

From the three-dimensional point of view, the signal events populate the interior of a compact region in the $(m_{j\ell_n}, m_{\ell\ell}, m_{j\ell_f})$ space, whose boundary is given by the constraint [98, 99]

$$\hat{m}^2_{j\ell_f} = \left[\sqrt{\hat{m}^2_{\ell\ell}\left(1-\hat{m}^2_{j\ell_n}\right)} \pm \frac{m_{\tilde{\ell}}}{m_{\tilde{\chi}^0_2}}\sqrt{\hat{m}^2_{j\ell_n}\left(1-\hat{m}^2_{\ell\ell}\right)}\right]^2, \tag{3.34}$$

which, for convenience, is written in terms of unit-normalized variables (see also (3.19))

$$\hat{m}_{j\ell_n} = \frac{m_{j\ell_n}}{m_{j\ell_n,max}}, \tag{3.35}$$

$$\hat{m}_{\ell\ell} = \frac{m_{\ell\ell}}{m_{\ell\ell,max}}, \tag{3.36}$$

$$\hat{m}_{j\ell_f} = \frac{m_{j\ell_f}}{m_{j\ell_f,max}}. \tag{3.37}$$

Our main goal in this section will be to test the algorithms from the previous sections for tagging the Voronoi cells in the vicinity of the boundary surface (3.34). In addition to the signal events from squark-neutralino associated production with the squark decaying as in Fig. 3.1, we shall also consider a representative number of background events. In order to make contact with the results from the previous sections, in section 3.4.1 we first take the background events to be uniformly distributed in mass-squared phase-space, and we ignore the combinatorial background. Then in section 3.4.2 we study a more realistic case, where the SM background is comprised of dilepton $t\bar{t}$ events and we also account for the combinatorial problem with the two leptons.

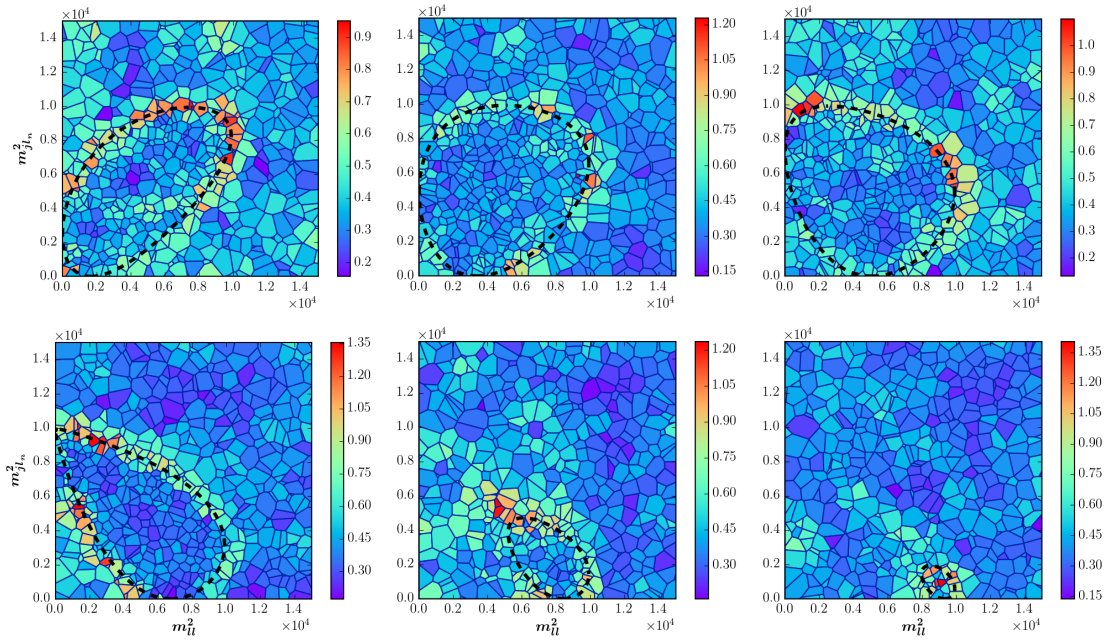


Figure 3.14: Two-dimensional slices of the relevant three-dimensional phase space of the SUSY-like cascade decay in Fig. 3.1. Each slice is in the $(m^2_{\ell\ell}, m^2_{j\ell_n})$ plane at a fixed value of $m^2_{j\ell_f} = 2000$ GeV^2 (upper left panel); $m^2_{j\ell_f} = 4000$ GeV^2 (upper middle panel); $m^2_{j\ell_f} = 6000$ GeV^2 (upper right panel); $m^2_{j\ell_f} = 8000$ GeV^2 (lower left panel); $m^2_{j\ell_f} = 10000$ GeV^2 (lower middle panel); and $m^2_{j\ell_f} = 11000$ GeV^2 (lower right panel). As in Figs. 3.6 and 3.10, the two-dimensional cells seen in the plots result from the intersection of the projective plane with the three-dimensional Voronoi cells, and are color-coded by the value of $\bar{\sigma}_i$ for the corresponding three-dimensional Voronoi cell.

3.4.1 An example with uniform background and no combinatorics

As in the other two three-dimensional examples considered in sections 3.2.2 and 3.3.3, in this section we include "SM physics background" events, which we take to be uniformly distributed everywhere throughout the mass-squared phase-space $(m^2_{j\ell_n}, m^2_{\ell\ell}, m^2_{j\ell_f})$ and normalized so that the density contrast across the boundary (3.34) is equal to $\rho = 4$. Note that in this scenario the interior "bulk" events and the "edge" cells on the surface boundary (3.34) consist of both SUSY signal and SM background events.

As before, we visualize the resulting Voronoi tessellation by presenting two-dimensional

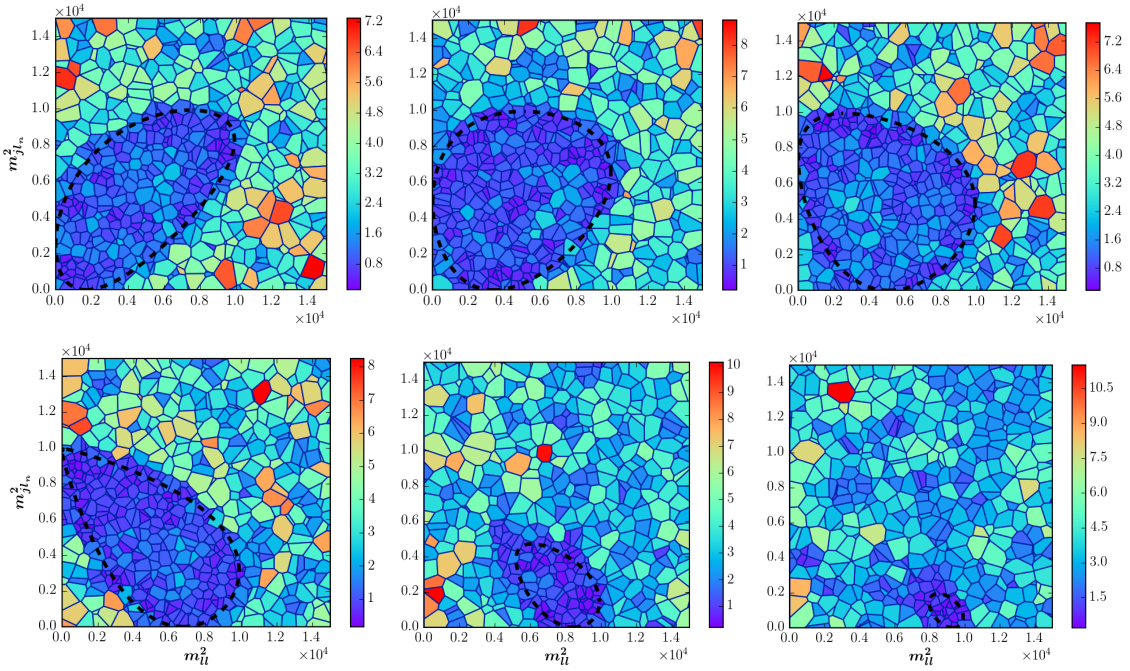


Figure 3.15: The same as Fig. 3.14, but color-coding the cells according to the normalized volume, $\bar{v}_i$, defined in (3.3).

slices of the relevant three-dimensional phase space, in this case[18] $(m^2_{j\ell_n}, m^2_{\ell\ell}, m^2_{j\ell_f})$. In Figs. 3.14 and 3.15 we show six slices in the $(m^2_{\ell\ell}, m^2_{j\ell_n})$ plane at a fixed value of $m^2_{j\ell_f}$ as follows: $m^2_{j\ell_f} = 2000$ GeV^2 (upper left panel); $m^2_{j\ell_f} = 4000$ GeV^2 (upper middle panel); $m^2_{j\ell_f} = 6000$ GeV^2 (upper right panel); $m^2_{j\ell_f} = 8000$ GeV^2 (lower left panel); $m^2_{j\ell_f} = 10000$ GeV^2 (lower middle panel); $m^2_{j\ell_f} = 11000$ GeV^2 (lower right panel). As in Figs. 3.6 and 3.10, the two-dimensional cells seen in the plots result from the intersection of the projective plane with the three-dimensional Voronoi cells and are color-coded by the value of the RSD, $\bar{\sigma}_i$, defined in (3.1) (in Fig. 3.14) or the normalized volume $\bar{v}_i$ defined in (3.3) (in Fig. 3.15) of the corresponding three-dimensional Voronoi cell. Just as in the case of the density-enhanced sphere considered in section 3.3.3, Figs. 3.14 and 3.15 suggest that the edge cells near the phase space boundary (3.34) are characterized both by a large value of $\bar{\sigma}_i$ and a small value of $\bar{v}_i$. Therefore, in designing a selection cut to pick up edge cells, it makes sense to consider both of these two variables at the same time. This is illustrated in the left panel of Fig. 3.16, which is the analogue of Fig. 3.13 for this case. We consider a

[18] It is known that working in terms of the *squared* masses is more convenient and intuitive [98, 99].

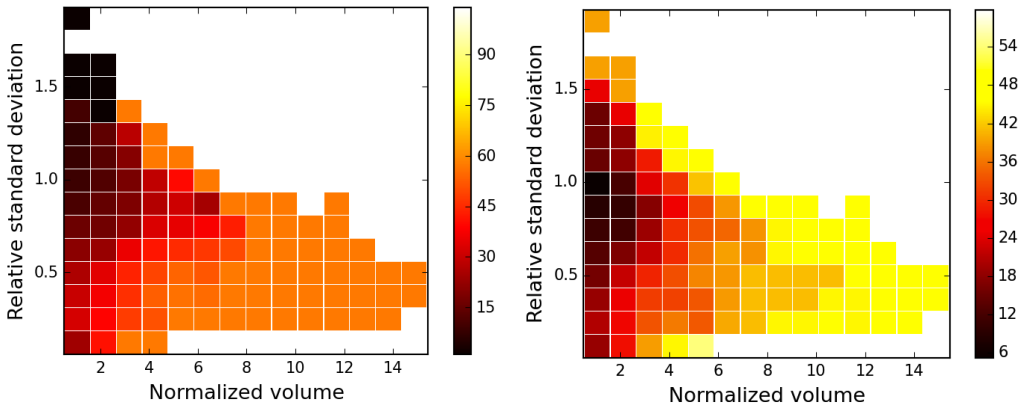


Figure 3.16: The integer ranking of the 15×15 bins in the $(\bar{v}, \bar{\sigma})$ plane according to their signal-to-background ratio. Left: The ranking for the case of $\rho = 4$. Right: the average ranking for the cases of $\rho = 1.2$, $\rho = 1.5$, $\rho = 2.0$, $\rho = 3.0$, and $\rho = 4.0$.

moderately large 15×15 grid in the $(\bar{v}, \bar{\sigma})$ plane and rank the resulting bins according to their signal-to-background ratio[19] as follows. The bin with the highest S/B is assigned rank 1, while the bin with the lowest S/B is assigned rank $15 \times 15 = 225$. In case of a tie between several bins, each bin is assigned the same average rank. Finally, bins with no events at all are ranked at the very bottom.[20] We observe that, similarly to Fig. 3.13, the highest ranked bins in terms of S/B appear at large values of $\bar{\sigma}$ and small values of $\bar{v}$. Using the obtained bin ranking, we can build the corresponding ROC curve, shown with the red solid line in the left panel of Fig. 3.17, which in some sense is the "ideal" ROC curve that could be achieved, if $\bar{\sigma}$ and $\bar{v}$ were the only discriminating variables under consideration.

One can now repeat the same procedure for different values of ρ. We show four more examples in the left panel of Fig. 3.17, with increasingly pessimistic values for the density contrast ρ: 3, 2, 1.5 and 1.2. As expected, the ROC curves become progressively worse, as quantified in the figure. With regards to the bin ranking in each case, we notice the following trend — the highest ranked bins remain at the lowest possible values of $\bar{v}$, but slide down the $\bar{\sigma}$ axis to slightly lower values of the

[19] We remind the reader that for the purpose of the boundary detection analysis performed here, "signal" refers to edge cells, while "background" refers to bulk cells. The interior bulk cells and the edge cells arise from both SUSY signal and SM background events, while the exterior bulk cells are only due to SM background events.

[20] This scheme is analogous to a college football ranking poll among 225 universities, some of which do not have a football program.

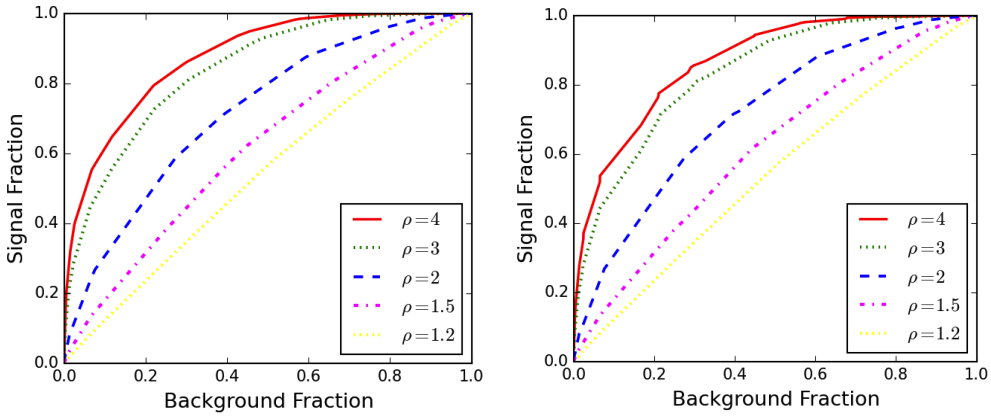


Figure 3.17: The same as the ROC curves shown in the right panel of Fig. 3.12, but for a 15×15 grid. Left: the ranking of the bins in constructing the ROC curve was done with the correct value of ρ (shown on each curve) used in generating the "data". Right: the ranking of the bins was always done according to the "average" ranking shown in the right panel of Fig. 3.16.

RSD, near, or even below $\bar{\sigma} \sim 1$. This is easy to understand intuitively — as ρ is decreased, the number densities on both sides of the surface boundary become more similar, and there is less variation between the sizes of bulk cells on the inside and on the outside. The fact that the bin ranking derived from our Monte Carlo simulations depends on the value of ρ poses an important conceptual problem with this procedure — when the analysis is performed on real data, we will not know the actual value of ρ, and, hence, we will not be certain which particular ordering of bins to use. Nevertheless, the fact that the highest ranked bins are clustered, more or less, in the same location, suggests a possible resolution: we can simply average our results obtained for several different values of ρ and use the resulting average rank for each bin, which is shown in the right panel of Fig. 3.16. The corresponding ROC curves derived with the help of this "average" bin ranking procedure are shown in the right panel of Fig. 3.17. Comparing the two panels of Fig. 3.17, we see that the ROC curves based on the average ranking are only slightly degraded compared to the "ideal" case.

We are now ready to start designing selection cuts for the edge cells. One possibility is to select the cells which fall into the a predetermined number, $N_{top\ bins}$, of the highest ranked bins in the $(\bar{v}, \bar{\sigma})$ plane. If we "cheat", i.e., use the correct value of ρ for the ranking (as in the left panel of Fig. 3.16), we obtain the result

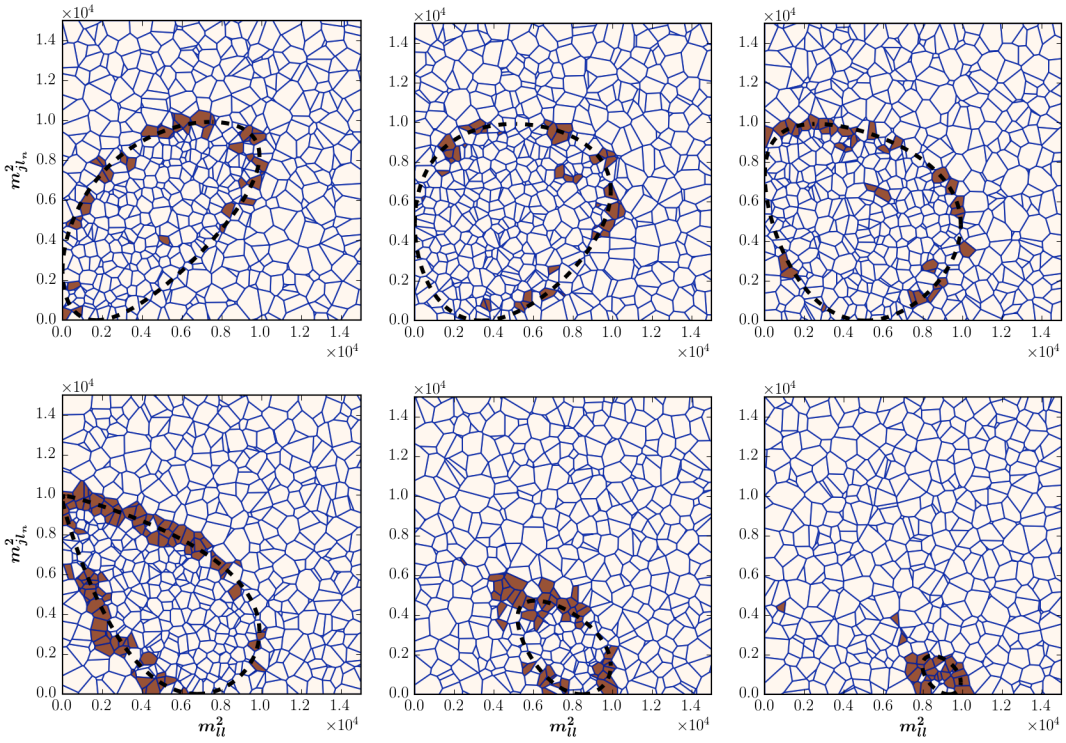


Figure 3.18: The Voronoi cells which pass the two-dimensional selection cut requiring the cell to belong to one of the top 5 bins in terms of signal-to-background ratio for the correct choice of $\rho = 4$ (see the left panel in Fig. 3.16).

shown in Fig. 3.18, where we have chosen $N_{top\ bins} = 5$. In general, the tagged cells are distributed throughout the volume of the three dimensional phase space, so for illustration purposes we again use the same six two-dimensional slices as in Figs. 3.14 and 3.15. We observe that the procedure is pretty efficient in tagging edge cells, and occasionally we tag an isolated bulk cell. Of course, for such a low value of $N_{top\ bins}$, not all edge cells will pass the cut, which will cause the boundary contours (marked with black dashed lines) to appear segmented and incomplete. By increasing the value of $N_{top\ bins}$, we can obviously tag more edge cells and eventually "close" those contours, but at the cost of more mistagged bulk cells.

Since the true value of ρ will be unknown, the plots in Fig. 3.18 are for academic purposes only. The more realistic situation is depicted in the analogous Fig. 3.19, where we again choose $N_{top\ bins} = 5$, only this time we use the average bin ranking from the right panel of Fig. 3.16. The result in Fig. 3.19 is slightly worse than Fig. 3.18 — while we do find a higher rate for mistagging bulk cells (typically in the interior),

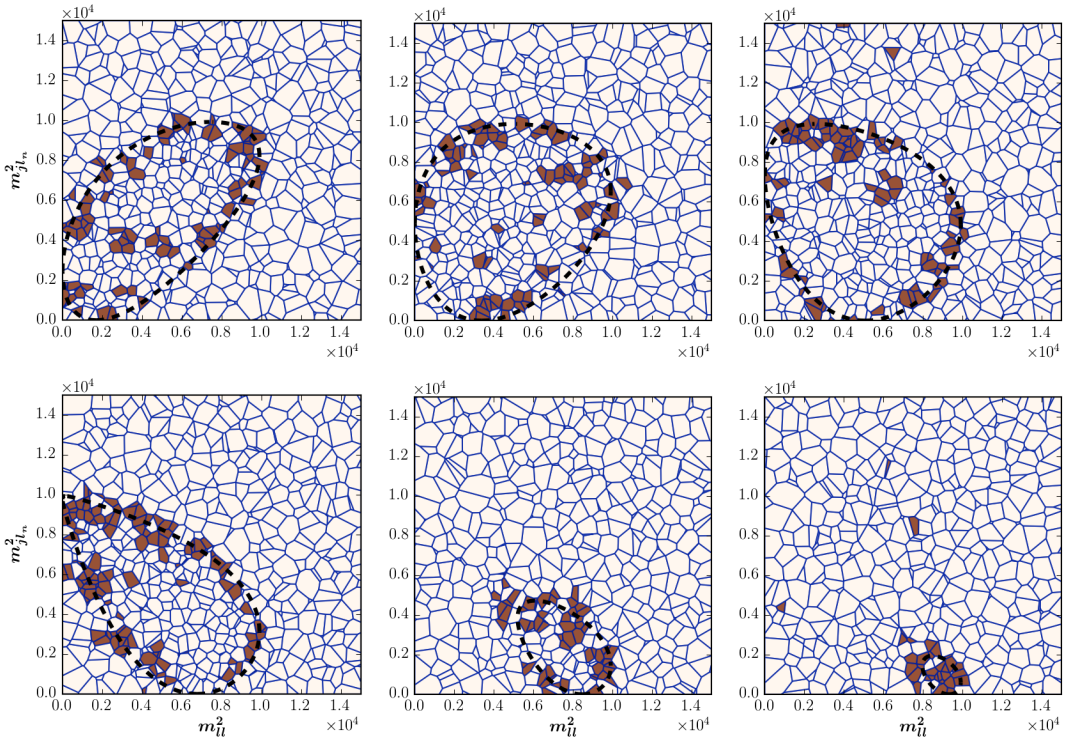


Figure 3.19: The same as Fig. 3.18, but using the *average* ranking of the bins shown in the right panel of Fig. 3.16.

the majority of the tagged cells are very close to the surface boundary, suggesting that this is a promising technique for identifying edge cells.

3.4.2 A realistic example with $t\bar{t}$ background and combinatorics

We now repeat the exercise from the previous section 3.4.1 with two improvements. First, we have to address the combinatorial problem of distinguishing the "near" and "far" lepton. The standard approach in the literature is to trade the original variables $m_{j\ell_n}$ and $m_{j\ell_f}$ for the ordered pair [24, 26, 37]

$$m_{j\ell(high)} \equiv \max\left\{m_{j\ell_n}, m_{j\ell_f}\right\}, \tag{3.38a}$$

$$m_{j\ell(low)} \equiv \min\left\{m_{j\ell_n}, m_{j\ell_f}\right\}. \tag{3.38b}$$

This reordering procedure is pictorially illustrated with the first two rows of plots in Fig. 3.20, where we show scatter plots of signal events for different ranges of the

third invariant mass variable (the dilepton mass): $1,000\ \text{GeV}^2 \leq m^2_{\ell\ell} \leq 3,000\ \text{GeV}^2$ (first column); $3,000\ \text{GeV}^2 \leq m^2_{\ell\ell} \leq 5,000\ \text{GeV}^2$ (second column); $5,000\ \text{GeV}^2 \leq m^2_{\ell\ell} \leq 7,000\ \text{GeV}^2$ (third column); and $7,000\ \text{GeV}^2 \leq m^2_{\ell\ell} \leq 9,000\ \text{GeV}^2$ (fourth column). In the first row, the signal events are plotted in the original plane of $(m^2_{j\ell_n}, m^2_{j\ell_f})$, and the points are color-coded as follows. The points below the diagonal 45° line, which have $m^2_{j\ell_n} \geq m^2_{j\ell_f}$, are colored in red, while the remaining points above the diagonal 45° line with $m^2_{j\ell_n} \leq m^2_{j\ell_f}$, are colored in black. The same data is then plotted in the second row of Fig. 3.20 in the plane of $(m^2_{j\ell(low)}, m^2_{j\ell(high)})$. Notice that the effect of the reordering procedure (3.38) is to leave all the black points in place, while interchanging the x and y coordinates of the red points[21]. After the reordering (3.38), half of the plane on each plot is left blank. In order to avoid such voids, in the third row of Fig. 3.20 we replot the data in the plane of $(m^2_{j\ell(low)}, m^2_{j\ell(high)} - m^2_{j\ell(low)})$, which is fully accessible.

As expected, the scatter plots in the third row of Fig. 3.20 exhibit boundary lines, which we can target with our edge-detecting method. In fact, each plot has two such boundaries where the signal number density sharply changes — one for the red points and another for the black points. At low values of $m_{\ell\ell}$ the "red" ("black") boundary line is an external (internal) boundary, while for high values of $m_{\ell\ell}$ it is the other way around. At intermediate values of $m_{\ell\ell}$ the two boundaries are very close to each other and that is where we expect the edge detection method to perform best.

Having thus taken care of the combinatorial problem, we now also improve our treatment of the background — instead of uniformly distributed background events as in section 3.4.1, we consider dilepton events from $t\bar{t}$ production. The corresponding scatter plots are shown in the fourth (last) row of Fig. 3.20. Since there are 2 b-jets, each background event contributes two entries to the scatter plot. We see that within the relevant range of the plotted variables $m^2_{j\ell(low)}$ and $m^2_{j\ell(high)} - m^2_{j\ell(low)}$, the distribution of the background events is somewhat uniform, with some noticeable clustering near the origin.

We are now in position to apply our Voronoi boundary detection algorithm. The result is shown in Fig. 3.21, where we present nine slices at fixed values of $m^2_{\ell\ell}$ as indicated at the top of each panel. The red (black) dashed line in each plot corresponds to the expected theoretical boundary implied by eq. (3.34) for the set of

[21] One can think of this operation as "origami folding" the scatter plot along the diagonal 45° line [36].

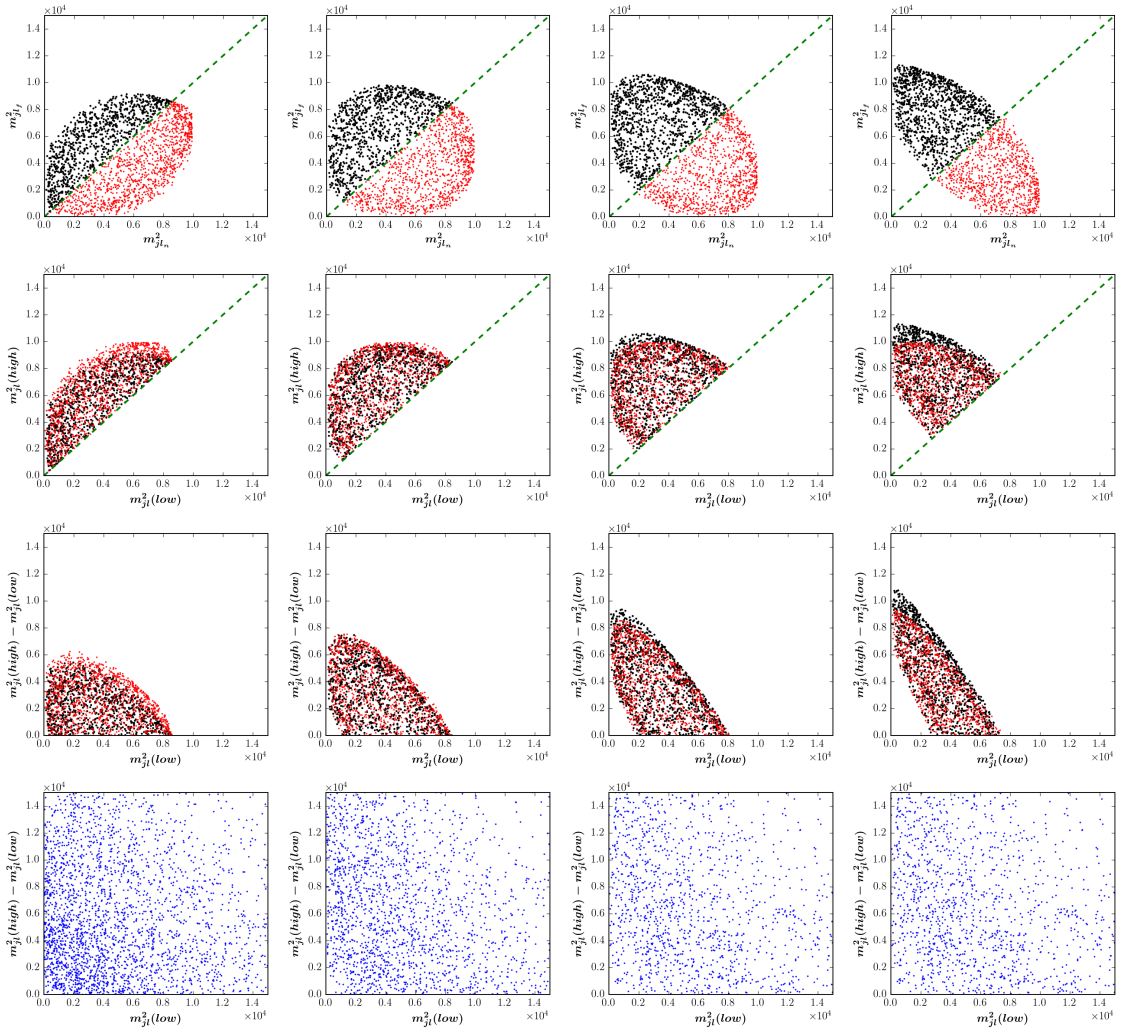


Figure 3.20: Scatter plots of signal events (black and red points, top three rows) and dilepton $t\bar{t}$ events (blue points, bottom row), for different ranges of the dilepton invariant mass squared: $1,000\ \text{GeV}^2 \leq m^2_{\ell\ell} \leq 3,000\ \text{GeV}^2$ (first column); $3,000\ \text{GeV}^2 \leq m^2_{\ell\ell} \leq 5,000\ \text{GeV}^2$ (second column); $5,000\ \text{GeV}^2 \leq m^2_{\ell\ell} \leq 7,000\ \text{GeV}^2$ (third column); and $7,000\ \text{GeV}^2 \leq m^2_{\ell\ell} \leq 9,000\ \text{GeV}^2$ (fourth column). In the first row, the signal events are plotted in the plane of $(m^2_{j\ell_n}, m^2_{j\ell_f})$ and colored red (black) if $m^2_{j\ell_n} \geq m^2_{j\ell_f}$ ($m^2_{j\ell_n} \leq m^2_{j\ell_f}$). The same points are then plotted in the planes of $(m^2_{j\ell(low)}, m^2_{j\ell(high)})$ (second row) and $(m^2_{j\ell(low)}, m^2_{j\ell(high)} - m^2_{j\ell(low)})$ (third row). The background events in the fourth row are plotted in the plane of $(m^2_{j\ell(low)}, m^2_{j\ell(high)} - m^2_{j\ell(low)})$.

points with $m^2_{j\ell_n} \geq m^2_{j\ell_f}$ ($m^2_{j\ell_n} \leq m^2_{j\ell_f}$) (see also the third row in Fig. 3.20). As before, the signal and background events were normalized so that $\rho = 4$. Fig. 3.21

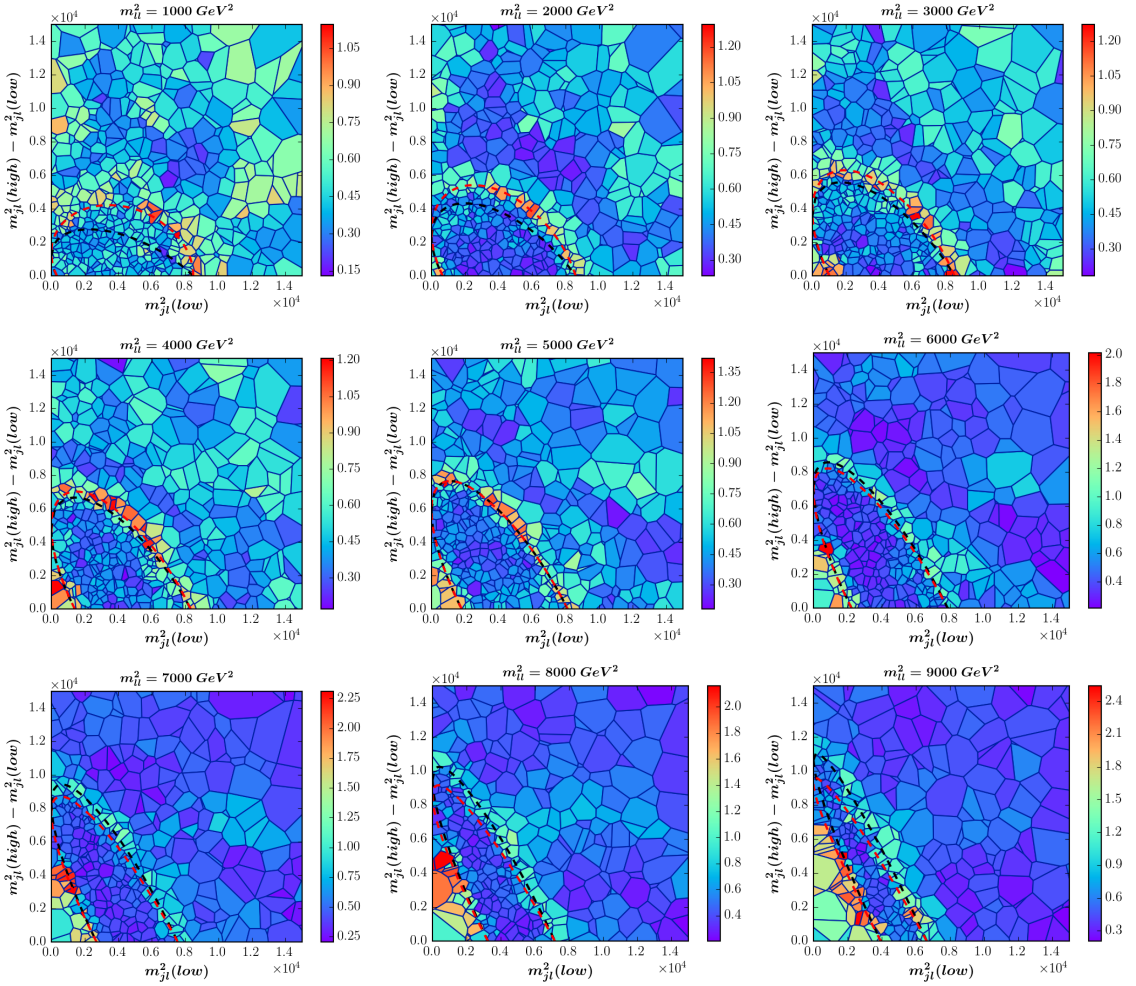


Figure 3.21: The analogue of Fig. 3.14 for the physics example considered in section 3.4.2. We show nine slices at fixed values of $m_{\ell\ell}^2$ as indicated at the top of each panel. The red (black) dashed line in each plot corresponds to the expected theoretical boundary implied by eq. (3.34) for the set of points with $m_{j\ell_n}^2 \geq m_{j\ell_f}^2$ ($m_{j\ell_n}^2 \leq m_{j\ell_f}^2$) (see also the third row in Fig. 3.20).

demonstrates that the Voronoi cells with the highest values of RSD $\bar{\sigma}_i$ (colored in red or orange) are indeed found near the theoretical boundaries (the red or black dashed lines). As anticipated, the method performs well for intermediate values of $m_{\ell\ell}^2 \sim (4,000 - 5,000)$ GeV^2, where the two boundaries resulting from the reordering (3.38) tend to overlap. We also observe that the boundary closer to the origin also seems to be singled out, especially at high values of $m_{\ell\ell}^2$.

3.5 Conclusions

In this paper, we took the first steps towards developing a general method for identifying surface boundaries in 3D phase space distributions from Voronoi tessellations. In the case of a sequential cascade decay like the one exhibited in Fig. 3.1, the surface boundary in the relevant (m_{12}, m_{23}, m_{13}) space is characterized by two properties: 1) it is the location of points where the number density is enhanced, due to the $\Delta_4^{-1/2}$ factor in the phase space distribution (3.25) [1]; 2) it is the location of points where the number density suddenly changes due to the lack of signal events outside the kinematically allowed boundary. These two properties motivate the use of the geometric variables, $\bar{\sigma}_i$ and $\bar{v}_i$, derived from the Voronoi tessellation of the data. (For other options, see [17].) We showed that the edge cells tend to have large values of $\bar{\sigma}_i$ and small values of $\bar{v}_i$, thus we advocated empirically selected cuts in terms of $\bar{v}_i$ and $\bar{\sigma}_i$ for tagging edge cell candidates. We considered several examples of increasing complexity and quantified the efficiency of those selection cuts using the language of ROC curves.

There are several directions in which this line of research can proceed from here.

- **Statistical significance of a set of tagged edge cells.** As we have seen in Figs. 3.18 and 3.19, the method is not perfect, in the sense that it occasionally tags a few bulk cells. Therefore, we need to develop a statistical procedure for determining the statistical significance of a given observed set of tagged edge cell candidates. Such a procedure should involve not just the relative number of tagged cells, but also their correlations, e.g., proximity to each other, connectedness, etc. This is an interesting subject on its own and will be addressed in a future publication [110].

- **Parameter measurements from fitting to a set of tagged edge cells.** Having selected a set of edge cell candidates, one could imagine fitting to the theoretical prediction for the shape of the boundary surface (3.34), obtaining a measurement of the mass spectrum of the new particles X_1, X_2, X_3, X_4. The actual fitting can be done in several different ways, which will also be investigated in [110].

- **Experimental effects.** In order to keep the discussion simple and to the point, in this paper we have ignored the experimental complications arising from finite

particle widths, detector resolution, ISR jet combinatorics, fakes, etc. Our goal was to present the method as a proof of principle, since the Voronoi approach to data analysis is still in its infancy. Once Voronoi-based methods have become more established and mature, it will become worthwhile to perform detailed and more realistic studies beyond parton level and with full detector simulation.

Here we have focused on cases where the number of signal events on the boundary is significant, leading to a "step" function discontinuity in the total density of events as one moves across the boundary. However, there are interesting examples of distributions where the number density is continuous, but exhibits a "kink", i.e., a discontinuity in its derivative (gradient) [113, 114, 115]. In such cases, our methods can still be applied — not directly on the initial data itself, but on a secondary data sample created by taking suitable derivatives. Indeed, while we have identified and studied a promising use of Voronoi tessellations in the analysis of particle physics data, there are many exciting applications yet to be developed. We look forward with anticipation to the future development and adoption of these novel and powerful methods.

Chapter Four: Detecting kinematic boundary surfaces in phase space: particle mass measurements in SUSY-like events[1]

4.1 Introduction

The dark matter problem is currently our best experimental evidence for the existence of new particles and interactions beyond the Standard Model (BSM). A great number of ongoing experiments are trying to discover dark matter particles through direct [118] or indirect detection [119]. In principle, dark matter particles could also be produced in high energy collisions at the Large Hadron Collider (LHC) at CERN, providing a complementary discovery probe in a controlled experimental environment [120].

Since the dark matter particles must be stable on cosmological timescales, in many popular BSM models they carry some conserved quantum number. The simplest choice is a $\mathbb{Z}_2$ parity, which is known as R-parity in models with low energy supersymmetry (SUSY) [18], Kaluza-Klein (KK) parity in models with universal extra dimensions (UED) [19], T-parity in Little Higgs models [20], etc. As a result, the dark matter particles are necessarily produced in pairs: either directly, or in the cascade decays of other, heavier BSM particles [121]. The prototypical such cascade decay is shown in Fig. 4.1, in which a new particle D undergoes a series of two-body decays, terminating in the dark matter candidate A, which is neutral and stable, and thus escapes undetected. Under those circumstances, measuring the set of four masses

$$\{m_D, m_C, m_B, m_A\} \tag{4.1}$$

is a difficult problem, which has been attracting a lot of attention over the last 20 years (for a review, see [21]). The main challenge stems from the fact that the momentum of

[1]Previously published as D. Debnath, J. S. Gainer, C. Kilic, D. Kim, K. T. Matchev, Y.-P. Yang,"Detecting kinematic boundary surfaces in phase space: particle mass measurements in SUSY-like events" on Journal of High Energy Physics 06(2017)092. Y.-P. Yang conducted the Monte Carlo simulation and analysis, as well as the numerical track of "same edge/endpoint" mass spectrum

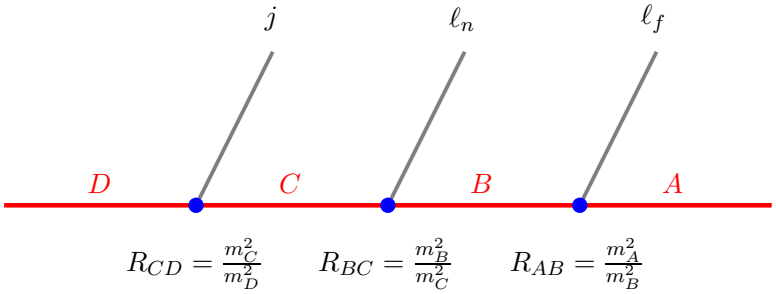


Figure 4.1: The generic decay chain under consideration in this paper: $D \to jC \to j\ell_n B \to j\ell_n\ell_f A$, where A, B, C and D are new BSM particles, while the SM decay products consist of one jet j and two leptons, labelled "near" ℓ_n and "far" ℓ_f. In the SUSY case, D represents a squark $\tilde{q}$, C is a heavier neutralino $\tilde{\chi}_2^0$, B is a charged slepton $\tilde{\ell}$ and A is the lightest neutralino $\tilde{\chi}_1^0$, which escapes undetected. The masses of the BSM particles are denoted by m_D, m_C, m_B and m_A. The corresponding ratios of squared masses R_{CD}, R_{BC} and R_{AB} are introduced for convenience in writing the kinematic endpoint formulas (4.21-4.28) and delineating the relevant regions in the mass parameter space (4.1) (see also eq. (4.20) and Fig. 4.2 below).

particle A is not measured, so that the standard technique of directly reconstructing the new particles as invariant mass resonances does not apply. Instead, one has to somehow infer the new masses (4.1) from the measured kinematic distributions of the *visible* SM decay products.

In the decay chain of Fig. 4.1, the SM decay products are taken to be a quark jet j and two leptons, labelled "near" ℓ_n and "far" ℓ_f. This choice is motivated by the following arguments:

- At a hadron collider like the LHC, strong production dominates, thus particle D is very likely to be colored. At the same time, the dark matter candidate A is neutral, therefore the color must be shed somewhere along the decay chain in the form of a QCD jet. Here we assume that this "color-shedding" occurs in the $D \to C$ transition[2], since one expects the strong decays of particle D to be the dominant ones.

[2]We note that in principle one can test this assumption experimentally, e.g. by constructing suitably defined on-shell constrained M_2 variables corresponding to the competing event topologies [94], or by studying the shapes and the correlations for the invariant mass variables considered below [34]. Such an exercise is useful, but beyond the scope of this paper.

- The presence of leptons among the SM decay products in Fig. 4.1 is theoretically not guaranteed, but is nevertheless experimentally motivated. First, leptonic signatures have significantly lower SM backgrounds and thus represent clean discovery channels. Second, the momentum of a lepton is measured much better than that of a jet, therefore the masses (4.1) will be measured with a better precision in a leptonic channel (as opposed to a purely jetty channel). Finally, if the SM decay products in Fig. 4.1 were all jets, in light of the arising combinatorial problem [122, 123, 124, 125, 126, 127], we would have to resort to sorted invariant mass variables [99, 128], whose kinematic endpoints are less pronounced and thus more difficult to measure over the SM backgrounds.

- From a historical perspective, the best motivation for considering the decay chain of Fig. 4.1 is that it is ubiquitous in SUSY, where D represents a squark $\tilde{q}$, C is a heavier neutralino $\tilde{\chi}_2^0$, B is a charged slepton $\tilde{\ell}$ and A is the lightest neutralino $\tilde{\chi}_1^0$, which escapes the detector and leads to missing transverse energy E_T. In the two most popular frameworks of SUSY breaking, gravity-mediated and gauge-mediated, the combination of (a) specific high scale boundary conditions, and (b) renormalization group evolution of the soft SUSY parameters down to the weak scale, leads to just the right mass hierarchy for the decay chain of Fig. 4.1 to occur. In the late 1990's and early 2000's, this prompted a flurry of activity on the topic of mass determination in such "SUSY-like" missing energy events. Soon afterwards, it was also realized that the decay chain of Fig. 4.1 is not exclusive to supersymmetry, but the same final state signature also appears in other models, e.g. minimal UED [129] and littlest Higgs [130].

To date, a large variety of mass measurement techniques for SUSY-like events have been developed. Roughly speaking, they all can be divided into two categories.

- *Exclusive methods.* In this case, one takes advantage of the presence of two decay chains in the event (they are often assumed identical) *and* the available E_T measurement. Several approaches are then possible. For example, in the so-called "polynomial methods" one attempts to solve explicitly for the momenta of the invisible particles in a given event, possibly using additional information from prior measurements of kinematic endpoints [38, 39, 40, 41, 42, 43,

44, 45, 46, 47, 48, 49, 50, 51, 52].[3] Alternatively, utilizing information from both branches, one could introduce suitable transverse[4] variables whose distributions exhibit an upper kinematic endpoint indicative of the parent particle mass [171, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67]. In the latter case, one still retains a residual dependence on the unknown dark matter particle mass m_A, which must be fixed by some other means, e.g. via the kink method [80, 81, 82, 83, 84, 85, 86, 87] or by performing a sufficient number of independent measurements [57, 61]. While they could be potentially quite sensitive, these exclusive methods are also less robust, since they rely on the correct identification of all objects in the event, and are thus prone to combinatorial ambiguities, the effects from $\not{E}_T$ resolution, initial and final state radiation, underlying event and pileup, etc.

- *Inclusive methods.* In this case, one focuses on the decay chain from Fig. 4.1 itself, disregarding what else is going on in the event. Using only the measured momenta of the visible SM decay products, i.e., the jet and the two leptons, one could form all possible invariant mass combinations[5], namely $m_{\ell\ell}$, $m_{j\ell_n}$, $m_{j\ell_f}$, and $m_{j\ell\ell}$, measure their respective upper kinematic endpoints

$$\left\{ m_{\ell\ell}^{max}, m_{j\ell_n}^{max}, m_{j\ell_f}^{max}, m_{j\ell\ell}^{max} \right\}, \tag{4.2}$$

  and use them to solve for the four input parameters (4.1). As just described, this approach is too naive, as it overlooks the remaining combinatorial problem involving the two leptons ℓ_n and ℓ_f. Since "near" and "far" cannot be distinguished on an event by event basis, the variables $m_{j\ell_n}$ and $m_{j\ell_f}$ are ill defined. This is why it has become customary to redefine the two jet-lepton invariant

[3] For long enough decay chains, the polynomial methods are able to solve for the invisible momenta, even without additional experimental input and without a second decay chain in the event. If the decay chain of Fig. 4.1 contained an additional two-body decay to a visible particle, just 5 events are sufficient for solving the event kinematics [39, 44].

[4] Transversality is not strictly necessary, in fact it may even be beneficial to work with $3+1$-dimensional variants of those variables [31, 32, 33, 88, 89, 90, 91, 92, 93, 95, 96, 97].

[5] In general, one is not limited to Lorentz-invariant variables only, e.g., recently it was suggested to study the peak of the *energy* distribution as a measure of the mass scale [68, 69, 70, 71].

mass combinations as[6]

$$m_{jl(lo)} \equiv \min\left\{m_{jl_n}, m_{jl_f}\right\}, \tag{4.3}$$

$$m_{jl(hi)} \equiv \max\left\{m_{jl_n}, m_{jl_f}\right\}. \tag{4.4}$$

The distributions of the newly defined quantities (4.3) and (4.4) also exhibit upper kinematic endpoints, $m^{max}_{jl(lo)}$ and $m^{max}_{jl(hi)}$, respectively. Then, instead of (4.2), one can use the new well-defined set of measurements

$$\left\{m^{max}_{ll}, m^{max}_{jll}, m^{max}_{jl(lo)}, m^{max}_{jl(hi)}\right\} \tag{4.5}$$

to invert and solve for the input mass parameters (4.1). This procedure constitutes the classic kinematic endpoint method for mass measurements, which has been successfully tested for several SUSY benchmark points [22, 23, 24, 25, 26, 27, 28, 29].

However, despite its robustness and simplicity, the kinematic endpoint method still has a couple of weaknesses. As we show below, taken together, they essentially lead to an almost flat direction in the solution space, thus jeopardizing the uniqueness of the mass determination. The first of these two problems is purely theoretical — it is well known that in certain regions of the parameter space (4.1) the four measurements (4.5) are not independent, but obey the relation [26]

$$\left(m^{max}_{jll}\right)^2 = \left(m^{max}_{jl(hi)}\right)^2 + \left(m^{max}_{ll}\right)^2. \tag{4.6}$$

In practice, this means that the measurements (4.5) fix only three out of the four mass parameters (4.1), leaving one degree of freedom undetermined. In what follows, we shall choose to parametrize this "flat direction" with the mass m_A of the lightest among the four new particles D, C, B and A. One can then use, e.g., the first three of the measurements in (4.5) and solve uniquely for the three heavier masses m_D, m_C, and m_B, leaving m_A as a free parameter. We list the relevant inversion formulas

[6] A more recent alternative approach is to introduce new invariant mass variables which are symmetric functions of $m_{j\ell_n}$ and $m_{j\ell_f}$, thus avoiding the need to distinguish ℓ_n from ℓ_f on an event per event basis [37].

in Appendix B. The obtained one-parameter family of mass spectra

$$\begin{cases} m_D = m_D(m_A; m_{ll}^{max}, m_{jll}^{max}, m_{jl(lo)}^{max}), \\ m_C = m_C(m_A; m_{ll}^{max}, m_{jll}^{max}, m_{jl(lo)}^{max}), \\ m_B = m_B(m_A; m_{ll}^{max}, m_{jll}^{max}, m_{jl(lo)}^{max}), \\ m_A \end{cases} \tag{4.7}$$

will satisfy the three measured kinematic endpoints m_{ll}^{max}, m_{jll}^{max}, and $m_{jl(lo)}^{max}$ by construction. What is more, in parameter space regions where eq. (4.6) holds, the family (4.7) will also obey the fourth measurement of $m_{jl(hi)}^{max}$, so that the four measurements (4.5) will be insufficient to lift the m_A degeneracy in (4.7).

These considerations beg the following two questions, which will be addressed in this paper.

1. In the remaining part of the parameter space, where (4.6) does *not* hold and $m_{jl(hi)}^{max}$ provides an independent fourth measurement, how well is the m_A degeneracy lifted after all? With the explicit examples of Sections 4.3 and 4.4 below, we shall show that although in theory the additional measurement of $m_{jl(hi)}^{max}$ determines the value of m_A, in practice this may be difficult to achieve, since the effect is very small and will be swamped by the experimental resolution.

2. In the region of parameter space in which (4.6) holds, what additional measurement should be used, and how well does it lift the degeneracy? In the existing literature, the standard approach is to consider the constrained[7] distribution $m_{jll(\theta>\frac{\pi}{2})}$, which exhibits a useful *lower* kinematic endpoint $m_{jll(\theta>\frac{\pi}{2})}^{min}$ [131, 132]. In what follows, we shall therefore always supplement the original set of 4 measurements (4.5) with the additional measurement of $m_{jll(\theta>\frac{\pi}{2})}^{min}$ to obtain the extended set

$$\left\{ m_{ll}^{max}, m_{jll}^{max}, m_{jl(lo)}^{max}, m_{jl(hi)}^{max}, m_{jll(\theta>\frac{\pi}{2})}^{min} \right\}, \tag{4.8}$$

[7]The distribution $m_{jll(\theta>\frac{\pi}{2})}$ is nothing but the usual m_{jll} distribution taken over a subset of the original events, namely those which satisfy the additional dilepton mass constraint

$$\frac{m_{ll}^{max}}{\sqrt{2}} < m_{ll} < m_{ll}^{max}\,.$$

In the rest frame of particle B, this cut implies the following restriction on the opening angle θ between the two leptons [132]

$$\theta > \frac{\pi}{2}\,,$$

thus justifying the notation for $m_{jll(\theta>\frac{\pi}{2})}$.

so that in principle there is sufficient information to determine the four unknown masses. Even then, we shall show that the sensitivity of the additional experimental input $m^{min}_{jll(\theta>\frac{\pi}{2})}$ to the previously found flat direction (4.7) is very low. First of all, it is already well appreciated that the measurement of $m^{min}_{jll(\theta>\frac{\pi}{2})}$ is very challenging, since in the vicinity of its lower endpoint, the shape of the signal distribution is concave downward, which makes it difficult to extract the endpoint with simple linear fitting, and one has to use the whole shape of the $m_{jll(\theta>\frac{\pi}{2})}$ distribution [30]. Secondly, as we shall show in the examples below, the variation of the value of $m^{min}_{jll(\theta>\frac{\pi}{2})}$ along the flat direction (4.7) can be numerically quite small, and therefore the sensitivity of the added fifth measurement along the flat direction (4.7) is not that great.

Either way, we see that the known methods for lifting the degeneracy of the flat direction (4.7) will face severe limitations once we take into account the experimental resolution, finite statistics, backgrounds, etc. [35, 36] Thus the first goal of this paper will be to illustrate the severity of the problem, i.e. to quantify the "flatness" of the family of solutions (4.7). For this purpose, we shall reuse the study points from Ref. [36], which at the time were meant to illustrate *discrete* ambiguities, i.e. cases where two distinct points in mass parameter space (4.1) accidentally happen to give mathematically identical values for all five measurements (4.8). Here we shall extend those study points to a family of mass spectra (4.7) which give mathematically identical values for the first three[8] of the measurements (4.8), and numerically *very similar* values for the remaining measurements.

Having identified the problem, the second goal of the paper is to propose a novel solution to it and investigate its viability. Our starting point is the observation that the signal events from the decay chain in Fig. 4.1 populate the interior of a compact region in the $(m_{j\ell_n}, m_{\ell\ell}, m_{j\ell_f})$ space, whose boundary is given by the surface $\mathcal{S}$ defined by the constraint [98, 133, 99]

$$\mathcal{S}: \quad \hat{m}^2_{j\ell_f} = \left[\sqrt{\hat{m}^2_{\ell\ell}\left(1-\hat{m}^2_{j\ell_n}\right)} \pm \frac{m_B}{m_C}\sqrt{\hat{m}^2_{j\ell_n}\left(1-\hat{m}^2_{\ell\ell}\right)}\right]^2, \tag{4.9}$$

which, for convenience, is written in terms of the unit-normalized variables

$$\hat{m}_{j\ell_n} = \frac{m_{j\ell_n}}{m^{max}_{j\ell_n}}, \quad \hat{m}_{\ell\ell} = \frac{m_{\ell\ell}}{m^{max}_{\ell\ell}}, \quad \hat{m}_{j\ell_f} = \frac{m_{j\ell_f}}{m^{max}_{j\ell_f}}. \tag{4.10}$$

[8] And sometimes four, if we are in parts of parameter space where (4.6) holds.

We note that both the shape and the size of the surface $\mathcal{S}$ depend on the input mass spectrum (4.1), i.e., $\mathcal{S}(m_A, m_B, m_C, m_D)$, and this dependence is precisely what we will be targeting with our method to be described below.

In its traditional implementation, the kinematic endpoint method is essentially[9] using the kinematic endpoints (4.2) of the one-dimensional projections of the signal population onto each of the three axes $m_{j\ell_n}$, $m_{\ell\ell}$ and $m_{j\ell_f}$, as well as onto the "radial" direction $m_{j\ell\ell} = \sqrt{m^2_{j\ell_n} + m^2_{\ell\ell} + m^2_{j\ell_f}}$. This approach is suboptimal because it ignores correlations and misses endpoint features along the other possible projections. The only way to guarantee that we are using the full available information in the data is to fit to the three-dimensional boundary (4.9) itself [1, 134], which will be the approach advocated here. As previously observed in [1] (and extended to a broader class of event topologies in [135]), most of the signal events are populated near the phase space boundary (4.9), on which the signal number density ρ_s formally becomes singular. This fact is rather fortuitous, since it implies a relatively sharp change in the local number density as we move across the phase space boundary, even in the presence of SM backgrounds (with some number density ρ_b, which is expected to be a relatively smooth function). Thus, we need to develop a suitable method for identifying regions in phase space where the gradient of the total number density $\rho \equiv \rho_b + \rho_s$ is relatively large, and then fit to them the analytical parametrization (4.9) in order to obtain the best fit values for the four new particle masses (4.1).

The first step of this program was already accomplished in our earlier paper [134], building on the idea originally proposed in [17] for finding "edges" in two-dimensional stochastic distributions of point data. Ref. [17] suggested that interesting features in the data, e.g., edge discontinuities, kinks, and so on, can be identified by analyzing the geometric properties of the Voronoi tessellation [7] of the data.[10] The volume v_i of a given Voronoi cell generated by a data point at some location $\vec{r}_i$ provides an estimate of the functional value of the number density ρ at that location,

$$\rho(\vec{r}_i) \sim \frac{1}{v_i}. \tag{4.11}$$

[9]The fact that one has to use $m_{jl(lo)}$ and $m_{jl(hi)}$ in place of $m_{j\ell_n}$ and $m_{j\ell_f}$ does not change the gist of the argument.

[10]We note the existence of efficient codes for finding Voronoi tessellations in the form of the `qHULL` algorithms [136]. Wrappers that allow the use of these algorithms in many frameworks also exist, and in this work we use a private `Python` code to compute the geometric attributes of the Voronoi cells.

Therefore, in order to obtain an estimate of $|\vec{\nabla}\rho(\vec{r})|$, we can construct variables which compare the properties of the Voronoi cell and its direct neighbors. Among the different options investigated in Refs. [17, 107], the relative standard deviation (RSD), $\bar{\sigma}_i$, of the volumes of neighboring cells, was identified as the most promising tagger of edge cells. The RSD was defined as follows. Let N_i be the set of neighbors of the i-th Voronoi cell C_i, with volumes, $\{v_j\}$, for $j \in N_i$. The RSD, $\bar{\sigma}_i$, is now defined by

$$\bar{\sigma}_i \equiv \frac{1}{\langle v(N_i)\rangle}\sqrt{\sum_{j\in N_i}\frac{\left(v_j - \langle v(N_i)\rangle\right)^2}{|N_i|-1}}, \tag{4.12}$$

where we have normalized by the average volume of the set of neighbors, N_i, of the i-th cell

$$\langle v(N_i)\rangle \equiv \frac{1}{|N_i|}\sum_{j\in N_i} v_j. \tag{4.13}$$

Subsequently, in [134] we showed that this procedure for tagging edge cells can be readily extended to three-dimensional point data, as is the case here. The end result of the method was a set of Voronoi cells which have been tagged as "edge cell candidates" since their values of $\bar{\sigma}_i$ were above the chosen threshold [134]. With the thus obtained set of edge cells in hand, it appears that we are in a perfect position to perform a mass measurement, simply by finding the set of values for $\{m_A, m_B, m_C, m_D\}$ which maximize the overlap between our tagged edge cells and the hypothesized surface $\mathcal{S}$. We have checked that this approach indeed works and gives a reasonable estimate of the true mass spectrum. However, here we prefer to suggest a slightly modified alternative, which accomplishes the same goal, but with somewhat better precision.

The problem with fitting to a subset of the original data set (namely the set of Voronoi cells which happened to pass the $\bar{\sigma}_i$ cut) is that we are still throwing away useful information, e.g., the Voronoi cells which barely failed the cut. In spite of formally failing, those cells are nevertheless still quite likely to be edge cells. Thus, in order to retain the full amount of information in our data, we prefer to abandon this "cut and fit" approach, and instead design a global variable which is calculated over the full data set. The only requirement is that the variable is maximized (or minimized, as the case may be) for the true values of the masses $\{m_A, m_B, m_C, m_D\}$.

In order to motivate such a variable, consider for a moment the case when the function $\rho(\vec{r})$ is known analytically, then let us investigate the (normalized) surface

integral

$$\frac{\int_{\tilde{\mathcal{S}}(\tilde{m}_A,\tilde{m}_B,\tilde{m}_C,\tilde{m}_D)} da\, |\vec{\nabla}\rho(\vec{r})|}{\int_{\tilde{\mathcal{S}}(\tilde{m}_A,\tilde{m}_B,\tilde{m}_C,\tilde{m}_D)} da} \tag{4.14}$$

for some arbitrary trial[11] values $(\tilde{m}_A, \tilde{m}_B, \tilde{m}_C, \tilde{m}_D)$ of the unknown masses (4.1). The meaning of the quantity (4.14) is very simple: it is the average gradient of $\rho(\vec{r})$ over the chosen surface $\tilde{\mathcal{S}}$. We expect the dominant contributions to the integral to come from regions where the gradient is large, and we know that the gradient is largest on the true phase space boundary $\mathcal{S}(m_A, m_B, m_C, m_D)$, defined in terms of the *true* values of the particle masses. However, if our choice for $(\tilde{m}_A, \tilde{m}_B, \tilde{m}_C, \tilde{m}_D)$ is wrong, the integration surface $\tilde{\mathcal{S}}$ will be far from the true phase space boundary $\mathcal{S}$, and those large contributions will be missed. The only way to capture *all* of the large contributions to the integral is to have $\tilde{\mathcal{S}}$ coincide with the true $\mathcal{S}$, and this is only possible if in turn the trial masses are exactly equal to the true particle masses. This suggests a method of mass measurement whereby the true mass spectrum is obtained as the result of an optimization problem involving the quantity (4.14).

Of course, in our case the analytical form of the integrand $|\vec{\nabla}\rho(\vec{r})|$ is unknown, but we can obtain a closely related quantity using the Voronoi tessellation of the data. Following [17, 134], we shall utilize the RSD $\bar{\sigma}_i$ defined in (4.12), which has been shown to be a good indicator of edge cells, and replace the integrand in (4.14) as

$$|\vec{\nabla}\rho(\vec{r})| \longrightarrow g(\vec{r}) \equiv \bar{\sigma}_i \text{ for } \vec{r} \in C_i. \tag{4.15}$$

In other words, the gradient estimator[12] function $g(\vec{r})$ is defined so that it is equal to the RSD $\bar{\sigma}_i$ of the Voronoi cell C_i in which the point $\vec{r}$ happens to be. Eqs. (4.14) and (4.15) suggest that the variable which we should be maximizing is

$$\bar{\Sigma}(\tilde{m}_A, \tilde{m}_B, \tilde{m}_C, \tilde{m}_D) \equiv \frac{\int_{\tilde{\mathcal{S}}(\tilde{m}_A,\tilde{m}_B,\tilde{m}_C,\tilde{m}_D)} da\, g(\vec{r})}{\int_{\tilde{\mathcal{S}}(\tilde{m}_A,\tilde{m}_B,\tilde{m}_C,\tilde{m}_D)} da}. \tag{4.16}$$

It obviously depends on our choice of trial masses $(\tilde{m}_A, \tilde{m}_B, \tilde{m}_C, \tilde{m}_D)$, and as argued above, we expect the maximum of $\bar{\Sigma}$ to occur for the correct choice (m_A, m_B, m_C, m_D),

[11] From here on trial values for the masses will carry a tilde to distinguish from the true values of the masses which will have no tilde. Correspondingly, $\tilde{\mathcal{S}}$ stands for a hypothesized "trial" boundary surface (4.9) obtained with trial values of the mass parameters.

[12] Note that $g(\vec{r})$ is not supposed to be an approximation for $|\vec{\nabla}\rho(\vec{r})|$, the crucial property for us is that the two functions peak in the same location.

i.e.

$$\max_{\tilde{m}_A,\tilde{m}_B,\tilde{m}_C,\tilde{m}_D} \bar{\Sigma}(\tilde{m}_A,\tilde{m}_B,\tilde{m}_C,\tilde{m}_D) \simeq \bar{\Sigma}(m_A,m_B,m_C,m_D). \tag{4.17}$$

This hypothesis will be tested and validated with explicit examples below in Sections 4.3 and 4.4.

The paper is organized as follows. In the next Section 4.2 we shall first review the well known formulas for the one-dimensional kinematic endpoints (4.8) and introduce the corresponding relevant partitioning of the mass parameter space into domain regions. In the next two sections we shall concentrate on the two most troublesome regions, $(3,2)$ and $(3,1)$, where the problematic relationship (4.6) holds. We shall pick one study point in each region, then study how well our conjecture (4.17) is able to determine the true mass spectrum. In principle, (4.17) involves optimization over 4 continuous variables, which is very time consuming (additionally, we have to perform the integration in the numerator of (4.16) by Monte Carlo). This is why for simplicity we choose to illustrate the power of our method with a one-dimensional toy study along the problematic flat direction (4.7). In particular, for each of our two study points we shall assume that the first three kinematic endpoints m_{ll}^{max}, m_{jll}^{max}, and $m_{jl(lo)}^{max}$ are already measured, leaving us only the task of determining the remaining degree of freedom m_A along the flat direction defined in (4.7). Correspondingly, we shall consider the whole family of mass spectra (4.7) which passes through a given study point. This family will eventually take us into the neighboring parameter space regions, including the third potentially problematic region, namely $(2,3)$, in which (4.6) is satisfied. For each family, we shall perform the following investigations

- As a warm up, we shall first illustrate that for each of the three distributions, m_{ll}, m_{jll}, and $m_{jl(lo)}$, the endpoint along the flat direction is the same (as expected by construction).

- We shall then investigate the variation of the kinematic endpoint of the $m_{jl(hi)}$ distribution along the flat direction (4.7). The endpoint value $m_{jl(hi)}^{max}$ is expected to be constant in regions $(3,2)$, $(3,1)$ and $(2,3)$, so the main question will be, how much does it vary in the remaining parameter space regions.

- We shall similarly investigate the variation of the lower kinematic endpoint $m_{jll(\theta>\frac{\pi}{2})}^{min}$ along the flat direction (4.7). Together with the previous item, this will

serve as an illustration of the main weakness of the classic kinematic endpoint method for mass measurements.

- Then we shall illustrate the distortion of the kinematic boundary surface (4.9) along the flat direction (4.7). The size of the distortion will be indicative of the precision with which one can hope to perform the mass measurement (4.17) using the kinematic boundary *surface* in phase space.

- Finally, we shall perform the fitting (4.17) along the flat direction parameterized by $\tilde{m}_A$. We shall show results in two cases: (a) when the background events are distributed uniformly in m^2 phase space, and (b) when the background is coming from dilepton $t\bar{t}$ events.

We shall summarize and conclude in Section 4.5. Appendix B contains the inversion formulas needed to define the flat direction (4.7).

4.2 Endpoint formulas and partitioning of parameter space

4.2.1 Notation and conventions

Following [36], we introduce for convenience some shorthand notation for the mass squared ratios

$$R_{ij} \equiv \frac{m_i^2}{m_j^2} \ , \tag{4.18}$$

where $i, j \in \{A, B, C, D\}$. Note that in (4.18) there are only three independent quantities, which can be taken to be the set $\{R_{AB}, R_{BC}, R_{CD}\}$. To save writing, we will also introduce convenient shorthand notation for the five kinematic endpoints as follows

$$a = (m_{ll}^{max})^2 \, , \quad b = \left(m_{jll}^{max}\right)^2 , \quad c = \left(m_{jl(lo)}^{max}\right)^2 , \quad d = \left(m_{jl(hi)}^{max}\right)^2 , \quad e = \left(m_{jll(\theta > \frac{\pi}{2})}^{min}\right)^2 . \tag{4.19}$$

Note that these represent the kinematic endpoints of the mass *squared* distributions[13].

[13]Contrast to the notation of Ref. [26], which uses a, b, c, d to label the same endpoints, but for the *linear* masses.

In the next two sections we shall use the three endpoint measurements m_{ll}^{max}, m_{jll}^{max}, and $m_{jl(lo)}^{max}$ to fix m_D, m_C and m_B, leaving m_A as a free parameter. Another way to think about this procedure is to note that the parameter space (4.1) can be equivalently parametrized as

$$\{R_{CD}, R_{BC}, R_{AB}, m_A\}\,. \tag{4.20}$$

Then, the endpoint measurements of m_{ll}^{max}, m_{jll}^{max}, and $m_{jl(lo)}^{max}$ can be used to fix the ratios R_{CD}, R_{BC} and R_{AB} (see Appendix B), leaving the overall mass scale undetermined and parametrized by m_A.

4.2.2 Endpoint formulas

The kinematical endpoints are given by the following formulas:

$$a \equiv (m_{ll}^{max})^2 = m_D^2\, R_{CD}\,(1-R_{BC})\,(1-R_{AB}); \tag{4.21}$$

$$b \equiv \left(m_{jll}^{max}\right)^2 = \begin{cases} m_D^2(1-R_{CD})(1-R_{AC}), & \text{for } R_{CD} < R_{AC}, & \text{case } (1,-), \\ m_D^2(1-R_{BC})(1-R_{AB}R_{CD}), & \text{for } R_{BC} < R_{AB}R_{CD}, & \text{case } (2,-), \\ m_D^2(1-R_{AB})(1-R_{BD}), & \text{for } R_{AB} < R_{BD}, & \text{case } (3,-), \\ m_D^2\left(1-\sqrt{R_{AD}}\right)^2, & \text{otherwise}, & \text{case } (4,-); \end{cases} \tag{4.22}$$

$$c \equiv \left(m_{jl(lo)}^{max}\right)^2 = \begin{cases} \left(m_{jl_n}^{max}\right)^2, & \text{for } (2-R_{AB})^{-1} < R_{BC} < 1, & \text{case } (-,1), \\ \left(m_{jl(eq)}^{max}\right)^2, & \text{for } R_{AB} < R_{BC} < (2-R_{AB})^{-1}, & \text{case } (-,2), \\ \left(m_{jl(eq)}^{max}\right)^2, & \text{for } 0 < R_{BC} < R_{AB}, & \text{case } (-,3); \end{cases} \tag{4.23}$$

$$d \equiv \left(m_{jl(hi)}^{max}\right)^2 = \begin{cases} \left(m_{jl_f}^{max}\right)^2, & \text{for } (2-R_{AB})^{-1} < R_{BC} < 1, & \text{case } (-,1), \\ \left(m_{jl_f}^{max}\right)^2, & \text{for } R_{AB} < R_{BC} < (2-R_{AB})^{-1}, & \text{case } (-,2), \\ \left(m_{jl_n}^{max}\right)^2, & \text{for } 0 < R_{BC} < R_{AB}, & \text{case } (-,3); \end{cases} \tag{4.24}$$

where

$$\left(m_{jl_n}^{max}\right)^2 = m_D^2\,(1-R_{CD})\,(1-R_{BC})\,, \tag{4.25}$$

$$\left(m_{jl_f}^{max}\right)^2 = m_D^2\,(1-R_{CD})\,(1-R_{AB})\,, \tag{4.26}$$

$$\left(m_{jl(eq)}^{max}\right)^2 = m_D^2\,(1-R_{CD})\,(1-R_{AB})\,(2-R_{AB})^{-1}\,. \tag{4.27}$$

Finally, the endpoint $m_{jll(\theta>\frac{\pi}{2})}^{min}$ introduced earlier in the Introduction, is given by

$$\begin{aligned} e \;\equiv\; \left(m_{jll(\theta>\frac{\pi}{2})}^{min}\right)^2 &= \frac{1}{4}m_D^2\Bigg\{(1-R_{AB})(1-R_{BC})(1+R_{CD}) \qquad (4.28)\\ &+\; 2\,(1-R_{AC})(1-R_{CD}) - (1-R_{CD})\sqrt{(1+R_{AB})^2(1+R_{BC})^2-16R_{AC}}\Bigg\}. \end{aligned}$$

4.2.3 Partitioning of the mass parameter space

One can see that the formulas (4.22-4.24) are piecewise-defined: they are given in terms of different expressions, depending on the parameter range for R_{CD}, R_{BC} and R_{AB}. This divides the $\{R_{CD}, R_{BC}, R_{AB}\}$ parameter subspace from (4.20) into several distinct regions, illustrated in Fig. 4.2. Following [26], we label those by a pair of integers (N_{jll}, N_{jl}). As already indicated in eqs. (4.22-4.24), the first integer N_{jll} identifies the relevant case for m_{jll}^{max}, while the second integer N_{jl} identifies the corresponding case for $(m_{jl(lo)}^{max}, m_{jl(hi)}^{max})$. One can show that only 9 out of the 12 pairings (N_{jll}, N_{jl}) are physical, and they are all exhibited within the unit square of Fig. 4.2. In what follows, an individual study point within a given region (N_{jll}, N_{jl}) will be marked with corresponding subscripts as $P_{N_{jll}N_{jl}}$.

Using (4.21), (4.22) and (4.24), it is easy to check that the "bad" relation (4.6), which can be equivalently rewritten in the new notation as

$$b = a + d, \tag{4.29}$$

is identically satisfied in regions (3,1), (3,2) and (2,3) of Fig. 4.2. Therefore, as already discussed, in these regions one would necessarily have to rely on the additional information provided by the measurement of the e endpoint (4.28).

Before concluding this rather short preliminary section, we direct the reader's attention to the color-coding in Fig. 4.2, where we have shaded in color six of the

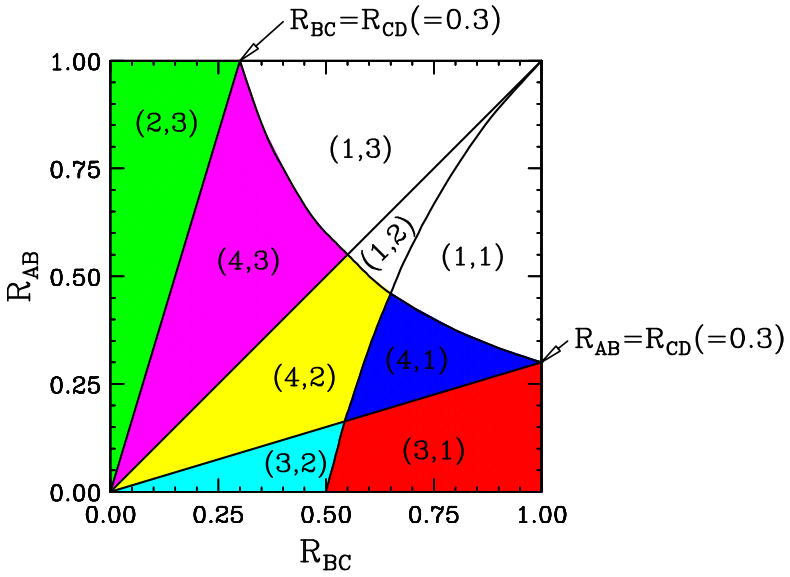


Figure 4.2: A slice through the $\{R_{CD}, R_{BC}, R_{AB}\}$ parameter space at a fixed $R_{CD} = 0.3$. The (R_{BC}, R_{AB}) plane exhibits the nine definition domains (N_{jll}, N_{jl}) of the set of equations (4.22-4.24). For the purposes of this paper, only six of those regions will be in play, and we have color-coded them as follows: region $(3,1)$ in red, region $(4,1)$ in blue, region $(3,2)$ in cyan, region $(4,2)$ in yellow, region $(4,3)$ in magenta, and region $(2,3)$ in green.

parameter space regions: region $(3,1)$ in red, region $(4,1)$ in blue, region $(3,2)$ in cyan, region $(4,2)$ in yellow, region $(4,3)$ in magenta, and region $(2,3)$ in green. It will turn out that the two families of mass spectra considered in the next two sections will visit the six color-shaded regions. For the benefit of the reader, in the remainder of the paper we shall strictly adhere to this color scheme — for example, results obtained for a study point from a particular region will always be plotted with the color of the respective region: study points in region $(3,1)$ are red, study points in region $(4,1)$ are blue, etc.

4.3 A case study in region $(3,1)$

4.3.1 Kinematical properties along the flat direction

In this section we shall study the flat direction (4.7) in mass parameter space which is generated by a study point P_{31} from region $(3,1)$ (the same study point was used

		true branch		auxilliary branch	
Region		$(3,1)$	$(4,1)$	$(4,3)$	$(2,3)$
Study point		P_{31}	P_{41}	P_{43}	P_{23}
m_A (GeV)		236.64	5000.00	2,000.00	100.00
m_B (GeV)		374.16	5126.02	2040.56	124.78
m_C (GeV)		418.33	5168.03	2167.36	272.54
m_D (GeV)		500.00	5256.90	2256.90	362.23
R_{AB}		0.400	0.951	0.960	0.642
R_{BC}		0.800	0.984	0.886	0.210
R_{CD}		0.700	0.966	0.922	0.566
m_{ll}^{max} (GeV)	$\sqrt{a}$	144.91			
m_{jll}^{max} (GeV)	$\sqrt{b}$	256.90			
$m_{jl(lo)}^{max}$ (GeV)	$\sqrt{c}$	122.47			
$m_{jl(hi)}^{max}$ (GeV)	$\sqrt{d}$	212.13	212.12	212.13	212.13
$m_{jll(\theta>\frac{\pi}{2})}^{min}$ (GeV)	$\sqrt{e}$	132.10	129.73	130.79	141.78

Table 4.1: Mass spectrum and expected kinematic endpoints for the study point P_{31} from region $(3,1)$ which was discussed in Ref. [36], together with three additional study points illustrating the different regions from Fig. 4.2 encountered by the parameter space trajectories from Fig. 4.3. By construction, all study points give identical values for the kinematic endpoints m_{ll}^{max}, m_{jll}^{max} and $m_{jl(lo)}^{max}$. Furthermore, in accordance with (4.6), the two study points P_{31} and P_{23} from regions $(3,1)$ and $(2,3)$ have identical values of $m_{jl(hi)}^{max}$. The remaining two study points P_{41} and P_{43}, representing regions $(4,1)$ and $(4,3)$, have essentially the same value for $m_{jl(hi)}^{max}$ as well. The last row lists the predicted values for $m_{jll(\theta>\frac{\pi}{2})}^{min}$, which are slightly different, and allow discriminating between the four endpoints in theory, but not in practice.

in [36] for a slightly different purpose). Table 4.1 lists some relevant information for the study point P_{31}: the input mass spectrum (4.1), the corresponding mass squared ratios (4.18), and the predicted kinematic endpoints (4.8), also reminding the reader of the alternative shorthand notation (4.19). As discussed in the Introduction, starting from the point P_{31}, we can follow a one-dimensional trajectory (4.7) through the parameter space (4.20) so that everywhere along the trajectory the prediction for the three endpoints a, b and c is unchanged (see Fig. 4.6 below). This trajectory is illustrated in Fig. 4.3, where we show its projections onto the three planes (R_{BC}, R_{AB}) (left panel), (R_{AB}, R_{CD}) (middle panel) and (R_{BC}, R_{CD}) (right panel). The lines in

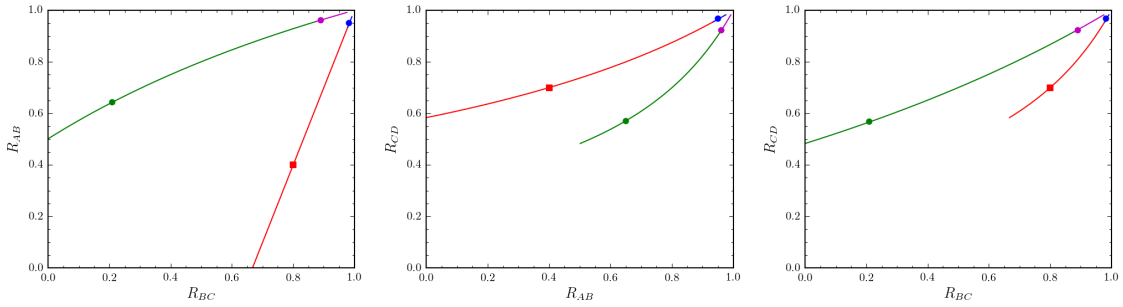


Figure 4.3: The two trajectories in mass parameter space leading to the same endpoints a, b and c. The lines are colored in accordance with the coloring convention for the regions depicted in Fig. 4.2. The red square marks the original study point P_{31} from Table 4.1, while the circles denote the other three study points from Table 4.1: P_{41} in region $(4,1)$ (blue circle), P_{43} in region $(4,3)$ (magenta circle), and P_{23} in region $(2,3)$ (green circle).

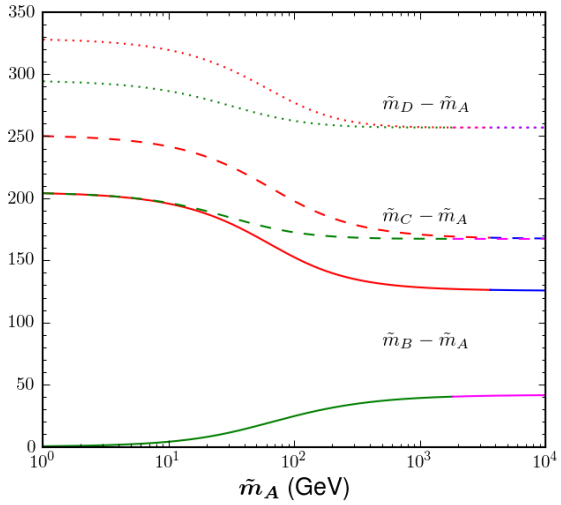


Figure 4.4: Mass spectra along the flat direction specified by the study point P_{31}. As a function of $\tilde{m}_A$, we plot the mass differences $\tilde{m}_B - \tilde{m}_A$ (solid lines), $\tilde{m}_C - \tilde{m}_A$ (dashed lines), and $\tilde{m}_D - \tilde{m}_A$ (dotted lines), which would preserve the values for the three kinematic endpoints a, b and c.

Fig. 4.3 are parametrized by the continuous test mass parameter $\tilde{m}_A$. For any given fixed value of $\tilde{m}_A$, the trajectory in Fig. 4.3 predicts the test values for the other three mass parameters, namely $\tilde{m}_B$, $\tilde{m}_C$ and $\tilde{m}_D$. This is shown more explicitly in Fig. 4.4, where we plot the mass differences $\tilde{m}_B - \tilde{m}_A$ (solid lines), $\tilde{m}_C - \tilde{m}_A$ (dashed lines), and $\tilde{m}_D - \tilde{m}_A$ (dotted lines), as a function of $\tilde{m}_A$.

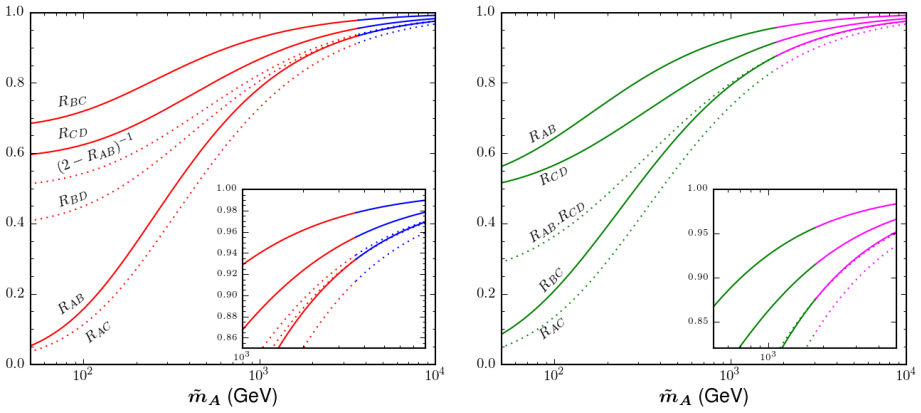


Figure 4.5: The equivalent representation of Fig. 4.4 in terms of the mass squared ratios R_{AB}, R_{BC} and R_{CD} (solid lines). The dotted lines depict various quantities of interest which are used to delineate the regions in Fig. 4.2. The left panel shows the true branch passing through regions $(3,1)$ (red) and $(4,1)$ (blue), while the right panel shows the auxiliary branch through regions $(2,3)$ (green) and $(4,3)$ (magenta). The left insert zooms in on the transition between regions $(3,1)$ and $(4,1)$ near $\tilde{m}_A = 3600$ GeV, while the right insert focuses on the transition between regions $(2,3)$ and $(4,3)$ near $\tilde{m}_A = 1800$ GeV.

All lines in Figs. 4.3 and 4.4 are color-coded using the same color conventions as for the parameter space regions in Fig. 4.2. Initially, as we move away from point P_{31} (marked with the red square in Fig. 4.3), we are still within the red region $(3,1)$, and the trajectory is therefore colored in red and parametrically given by eqs. (B.8-B.10). As the value of $\tilde{m}_A$ is reduced from its nominal value (236.6 GeV) at the point P_{31}, the mass spectrum gets lighter and eventually we reach $\tilde{m}_A = 0$, where (the red portion of) the trajectory terminates at $R_{AB} = 0$, $R_{BC} \simeq 0.67$ and $R_{CD} \simeq 0.58$. If, on the other hand, we start increasing $\tilde{m}_A$ from its nominal P_{31} value, the spectrum gets heavier, and we start approaching the neighboring region $(4,1)$. Eventually, at around $\tilde{m}_A \sim 3600$ GeV, the trajectory crosses into region $(4,1)$ and thus changes its color to blue. This transition is illustrated in the left panel of Fig. 4.5, where we plot the mass squared ratios R_{AB}, R_{BC} and R_{CD} (solid lines), together with some other relevant quantities (dotted lines). In particular, the boundary between regions $(3,1)$ and $(4,1)$ is given by the relation $R_{AB} = R_{BD}$, see (B.2) and (B.29). We can see that crossover more clearly in the insert in the left panel of Fig. 4.5, where the line color changes from red to blue as soon as the R_{BD} (dotted) line crosses the R_{AB} (solid) line.

Once we are in region $(4,1)$, we follow the blue portion of the trajectory in Fig. 4.3, which is parametrically defined by eqs. (B.36-B.38). We choose a representative study point for region $(4,1)$ as well — it is denoted by P_{41} and listed in the third (blue shaded) column of Table 4.1. The corresponding mass spectrum is clearly very heavy, but is nevertheless perfectly consistent with the three measured endpoints a, b and c, as shown in Fig. 4.6. As seen in Fig. 4.3, the blue portion of the mass trajectory appears headed for the point $(R_{AB}, R_{BC}, R_{CD}) = (1,1,1)$, which is indeed reached in the limit of $\tilde{m}_A \to \infty$, without ever entering into the neighboring region $(1,1)$[14].

Fig. 4.3 reveals that the mass family (4.7) through our study point P_{31} includes a segment which starts at $(R_{AB}, R_{BC}, R_{CD}) = (0, 0.67, 0.58)$ and ends at $(R_{AB}, R_{BC}, R_{CD}) = (1,1,1)$, visiting regions $(3,1)$ and $(4,1)$. Since the actual study point P_{31} belongs to this segment, in what follows we shall refer to it as "the true branch". However, Fig. 4.3 also shows that there is an additional disconnected segment of the mass trajectory through the green region $(2,3)$ and the magenta region $(4,3)$. In the following, we shall refer to this additional segment as "the auxiliary branch". Note that this terminology is introduced only for clarity and should not be taken too literally — as far as the measured endpoints a, b and c are concerned, all points on the true and auxiliary branches are on the same footing, since the experimenter would have no way of knowing *a priori* which is the true branch and which is the auxiliary branch. This is why we have to seriously consider points on the auxiliary branch as well. We choose two representative study points, which are listed in the last two columns of Table 4.1: point P_{43} belongs to the magenta region $(4,3)$, while point P_{23} is in the green region $(2,3)$. As shown in Fig. 4.3, the auxiliary branch starts at $(R_{AB}, R_{BC}, R_{CD}) = (0.5, 0, 0.48)$ and asymptotically meets the true branch at the corner point $(R_{AB}, R_{BC}, R_{CD}) = (1,1,1)$. The transition between the two regions $(2,3)$ and $(4,3)$ along the auxiliary branch is illustrated in the right panel of Fig. 4.5. According to (B.21) and (B.54), the boundary between regions $(2,3)$ and $(4,3)$ is defined by the relation $R_{BC} = R_{AB}R_{CD}$. The right panel of Fig. 4.5 confirms this: the color of the auxiliary branch in Figs. 4.3 and 4.4 changes from green to magenta as soon as the dotted line representing the product $R_{AB}R_{CD}$ crosses the solid line for R_{BC}.

To summarize our discussion so far, we have imposed the three endpoint measure-

[14]Note that as the value of R_{CD} increases, the $(1,1)$ region shrinks and for $R_{CD} = 1$ it disappears altogether.

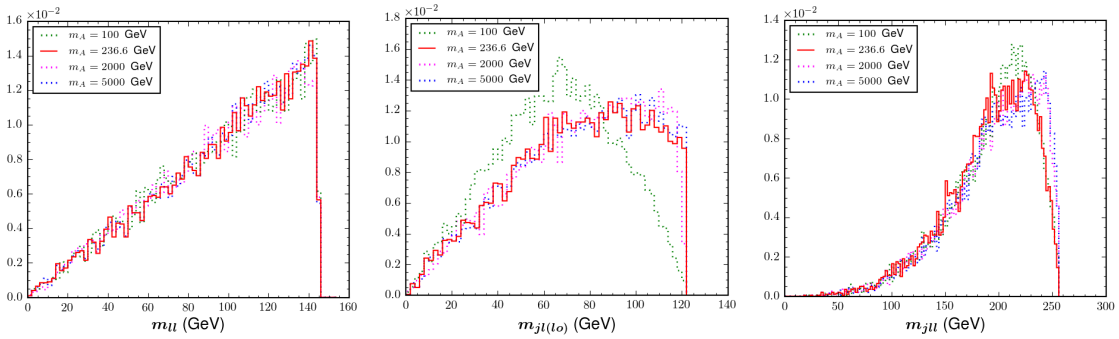


Figure 4.6: Unit-normalized invariant mass distributions for the four study points from Table 4.1: the distribution of $m_{\ell\ell}$ (left panel), $m_{j\ell(lo)}$ (middle panel), and $m_{j\ell\ell}$ (right panel). The lines are color coded according to our conventions from Fig. 4.2 and Table 4.1: red for P_{31}, blue for P_{41}, magenta for P_{43} and green for P_{23}.

ments a, b and c on the four-dimensional parameter space (4.1), reducing it to the one-dimensional parameter curve depicted in Figs. 4.3 and 4.4. The curve consists of two branches which visit four of the colored regions in Fig. 4.2, and we have chosen one study point in each region. The four study points are listed in Table 4.1, and their predicted invariant mass distributions from the ROOT phase space generator [137] are shown in Fig. 4.6: $m_{\ell\ell}$ in the left panel, $m_{j\ell(lo)}$ in the middle panel and $m_{j\ell\ell}$ in the right panel. By construction, for any points along the mass trajectory (4.7), and in particular for the four study points from Table 4.1, these distributions share common kinematic endpoints. Furthermore, as Fig. 4.6 reveals, the shapes of most distributions are also very similar, which makes it difficult to pinpoint our exact location along the mass trajectory (4.7). This is why in the remainder of this section, we shall focus on the question, what additional measurements may allows us to discriminate experimentally points along the two branches in Figs. 4.3 and 4.4, and in particular distinguish between the four study points in Table 4.1.

One obvious possibility is to investigate the remaining kinematic endpoints d and e, which are analyzed in Figs. 4.7 and 4.8, respectively. The left panels show the theoretical predictions for the kinematic endpoints $\sqrt{d} = m^{max}_{j\ell(hi)}$ and $\sqrt{e} = m^{min}_{jll(\theta > \frac{\pi}{2})}$ along the flat direction (4.7) as a function of $\tilde{m}_A$, while the right panels exhibit the corresponding invariant mass distributions for each of our four study points from Table 4.1.

Let us first focus on Fig. 4.7 which illustrates the $\tilde{m}_A$ dependence of the $m_{j\ell(hi)}$

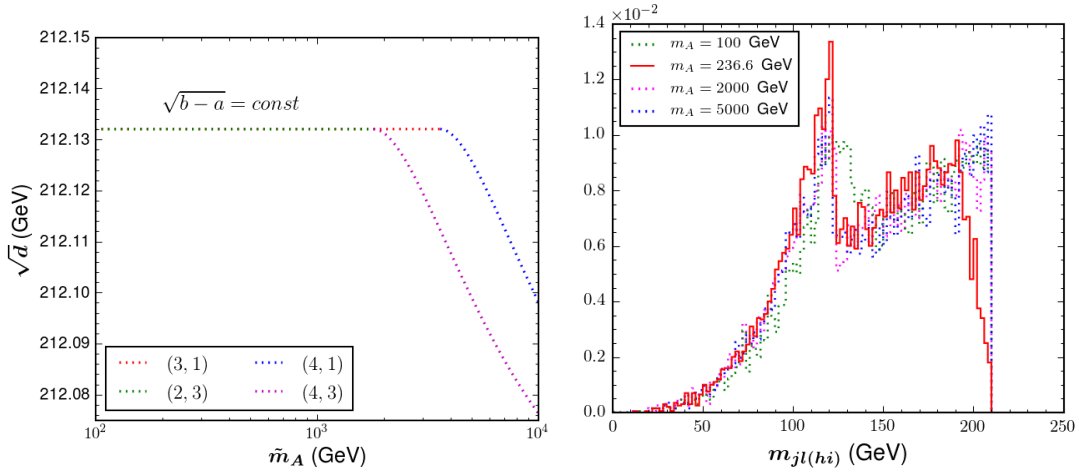


Figure 4.7: Left: The prediction for the kinematic endpoint $\sqrt{d}$ along the flat direction (4.7) generated by P_{31}, as a function of the trial value of the parameter $\tilde{m}_A$. Right: The same as Fig. 4.6, but for the distribution $m_{j\ell(hi)}$.

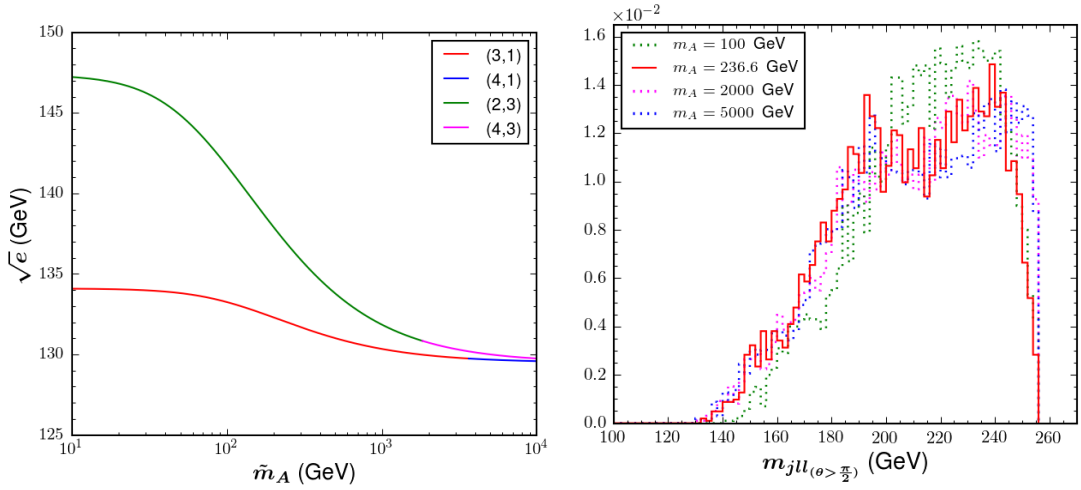


Figure 4.8: The same as Fig. 4.7, but for the endpoint $\sqrt{e}$ and the corresponding distribution $m_{jll(\theta>\frac{\pi}{2})}$.

distribution and its kinematic endpoint $\sqrt{d}$. As we have already discussed, in regions $(3,1)$ and $(2,3)$ the additional measurement of $\sqrt{d}$ is not useful, since it is not independent — the value of d is predicted by the relation (4.29), as confirmed by the left panel in Fig. 4.7, where the red and green dotted lines representing those two regions are perfectly flat and insensitive to $\tilde{m}_A$. However, this still leaves open the possibility that in the remaining two regions, namely $(4,1)$ and $(4,3)$, the measurement of

the d endpoint will be able to lift the degeneracy and determine the value of m_A, since, at least in theory, d is a non-trivial function of $\tilde{m}_A$, see (B.39b) and (B.63b). Unfortunately, Fig. 4.7 demonstrates that this is not the case in practice — the $\tilde{m}_A$ dependence is extremely weak, and the endpoint value for $\sqrt{d}$ only changes by a few tens of MeV as $\tilde{m}_A$ is varied over a range of several TeV! This lack of sensitivity is the reason why we have been referring to the family of mass spectra (4.7) as a "flat direction" in mass parameter space. Clearly, due to the finite experimental resolution, an endpoint measurement with a precision of tens of MeV is not feasible, the anticipated experimental errors at the LHC are significantly higher, on the order of a few GeV [138].

It is instructive to understand this lack of sensitivity analytically, by studying, e.g. the mathematical expression (B.39b) for d which is relevant for region $(4,1)$. Figs. 4.4 and 4.7 already showed that region $(4,1)$ occurs at large values of $\tilde{m}_A$, where the spectrum is relatively heavy — on the order of several TeV. At the same time, the measured parameter inputs into (B.39b), namely the endpoints a, b and c, are all on the order of several hundred GeV. This suggests an expansion in terms of $1/\tilde{m}_A$ as

$$d(a,b,c,\tilde{m}_A) \equiv K_0 + \frac{K_1}{\tilde{m}_A} + \mathcal{O}\left(\frac{1}{\tilde{m}_A^2}\right). \tag{4.30}$$

Using (B.39b), we get the expansion coefficients to be

$$K_0 = \frac{ac}{(a+c)^2}\left(\sqrt{b}+\sqrt{b-a-c}\right)^2, \tag{4.31}$$

$$K_1 = \frac{ac\left[\left(\sqrt{b}+\sqrt{b-a-c}\right)(a^2+ac-2ab+2bc)+(a^2-c^2)\sqrt{b}\right]}{(a+c)^3}. \tag{4.32}$$

Interestingly, the numerical value of K_0 is extremely close to $b-a$:

$$K_0 \equiv \lim_{\tilde{m}_A\to\infty} d = (212.047\text{ GeV})^2 \leftrightarrow b-a = (212.132\text{ GeV})^2. \tag{4.33}$$

Since K_0 is the leading order prediction for d, (4.33) implies that even in region $(4,1)$, the relation (4.29) will still hold to a very good approximation — any deviations from it will be $1/\tilde{m}_A$ suppressed. We can formalize this observation by introducing the value $m_A^{(b)}$ which the parameter $\tilde{m}_A$ takes when the mass trajectory (4.7) crosses the boundary between regions $(3,1)$ and $(4,1)$. Using the continuity of the function

$d(a, b, c, \tilde{m}_A)$, we can write

$$b - a = K_0 + \frac{K_1}{\tilde{m}_A^{(b)}} + \mathcal{O}\left(\frac{1}{\left(\tilde{m}_A^{(b)}\right)^2}\right), \tag{4.34}$$

where the left-hand side is the value of d in region $(3, 1)$ which is given by (B.11), while the right-hand side is the value of d as predicted by the Taylor expansion (4.30) in region $(4, 1)$. Eliminating K_0 from (4.34), we can rewrite the expansion (4.30) in the form

$$d(a, b, c, \tilde{m}_A) \equiv b - a + \frac{K_1}{\tilde{m}_A} - \frac{K_1}{m_A^{(b)}} + \mathcal{O}\left(\frac{1}{\tilde{m}_A^2}\right) \tag{4.35a}$$

$$= b - a - K_1 \frac{\tilde{m}_A - m_A^{(b)}}{\tilde{m}_A m_A^{(b)}} + \mathcal{O}\left(\frac{1}{\tilde{m}_A^2}\right), \tag{4.35b}$$

which manifestly shows that the deviations from the relation (4.29) are $1/m_A$ suppressed. One can check that the sign of the K_1 coefficient (4.32) is positive, then (4.35b) explains why d is a *decreasing* function of $\tilde{m}_A$ in region $(4, 1)$, as observed in the left panel of Fig. 4.7.

Starting from (B.63b), one can repeat the same analysis for the magenta portion of the auxiliary branch which is located in region $(4, 3)$. As the left panel of Fig. 4.7 shows, the conclusions will be the same — the d endpoint is still given approximately by the "bad" relation (4.29), and the corrections to it are tiny and $1/m_A$ suppressed. The right panel in Fig. 4.7 explicitly demonstrates that the variation of the d endpoint along the flat direction is unnoticeable by eye even with perfect resolution, large statistics and no background. The shapes of the $m_{j\ell(hi)}$ distributions are also very similar. As a result, we anticipate that the additional measurement of the d kinematic endpoint and the analysis of the associated $m_{j\ell(hi)}$ distribution will not help much in lifting the degeneracy of the flat direction (4.7).

We now turn to the discussion of the fifth and final kinematic endpoint, e, illustrated in Fig. 4.8. The left panel now shows a more promising result — the variation along the flat direction is much larger than what we saw previously in Fig. 4.7. This is especially noticeable for the auxiliary branch, where the prediction for $\sqrt{e}$ can vary by as much as 17 GeV, suggesting that one might be able to at least rule out some portions of it. At the same time, the variation of $\sqrt{e}$ along the true branch is only

4 GeV, once again making it rather difficult to pinpoint an exact location along the true branch. Unfortunately, these theoretical considerations are dwarfed by the experimental challenges in measuring the e endpoint, as suggested by the right panel of Fig. 4.8. Unlike the other four kinematic endpoints, e is a lower endpoint (a.k.a. "threshold"), which places it in a region where one expects more background. More importantly, the signal distribution is very poorly populated near its lower endpoint - the vast majority of signal events appear sufficiently far away from the threshold, and the measurement will suffer from a large statistical uncertainty. This casts significant doubts on the feasibility of this measurement — in previous studies, the $\sqrt{e}$ endpoint was either the measurement with the largest experimental error from the fit (on the order of 10 GeV [24]), or one could not obtain a measurement for it at all [26]. One could hope to improve on the precision by utilizing shape information [139], but this introduces additional systematic uncertainty, since the background shape and the shape distortion due to cuts has to be modeled with Monte Carlo.

Being mindful of the challenges involved with the measurement of the e endpoint, in this paper we shall look for an alternative method for lifting the degeneracy along the flat direction. Our proposal is to study the shape of the kinematic boundary (4.9), which is a two-dimensional surface in the three-dimensional space of observables

$$\left\{m^2_{j\ell(lo)}, m^2_{j\ell(hi)}, m^2_{\ell\ell}\right\}. \tag{4.36}$$

As a proof of principle, we first illustrate the change in the shape of the surface (4.9) as we move along the flat direction. Our results are shown in Fig. 4.9 (for the true branch) and in Fig. 4.10 (for the auxiliary branch). Following [134], we visualize the surface (4.9) by showing a series of two-dimensional slices in the $(m^2_{j\ell(lo)}, m^2_{j\ell(hi)} - m^2_{j\ell(lo)})$ plane, where the slight modification of the "y-axis" was done in order to avoid wasted space on the plots due to the unphysical areas with $m_{j\ell(lo)} > m_{j\ell(hi)}$. Each slice is taken at a fixed value of $m^2_{\ell\ell}$, starting from a very low value (10 GeV^2) and going up all the way until the kinematic endpoint $(m^{max}_{\ell\ell})^2 = 20,976$ GeV^2. The red solid lines in Fig. 4.9 correspond to the nominal case of the study point P_{31}. In each panel, the signal events will be populating the areas delineated by these red solid lines. As pointed out in [1], the density of signal events is enhanced near the phase space boundary, i.e. signal events will cluster close to the solid red lines; this property can be incorporated into the algorithm for detecting the surface boundary [134]. It is worth noting that in general, each panel contains two signal populations, which arise

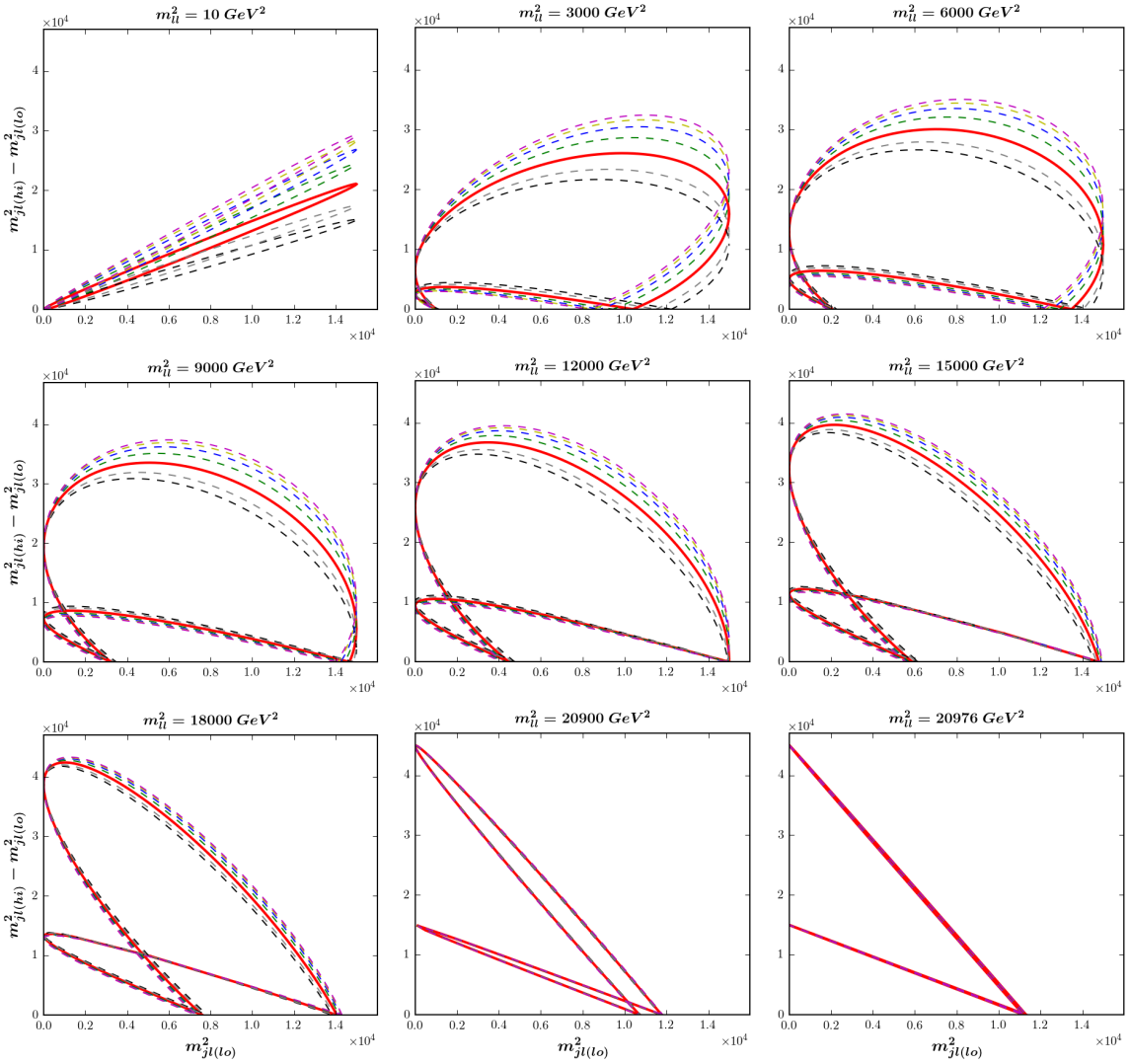


Figure 4.9: Signal kinematic boundaries in the $(m^2_{j\ell(lo)}, m^2_{j\ell(hi)} - m^2_{j\ell(lo)})$ plane, at nine fixed values of $m^2_{\ell\ell}$. Results are shown for several points along the true branch in regions $(3,1)$ and $(4,1)$. The red solid line represents the case of the P_{31} study point with $\tilde{m}_A = 236.6$ GeV, while the dashed lines correspond to other values of $\tilde{m}_A$ along the true branch: $\tilde{m}_A = 0$ (black), $\tilde{m}_A = 100$ GeV (gray), $\tilde{m}_A = 500$ GeV (green), $\tilde{m}_A = 1000$ GeV (blue), $\tilde{m}_A = 2000$ GeV (yellow) and $\tilde{m}_A = 5000$ GeV (magenta).

from the reordering (4.3-4.4) [36]. As we vary the value of $m^2_{\ell\ell}$, the shape of the red solid lines changes in accordance with eq. (4.9), which follows from simple phase space considerations. However, the main purpose of Figs. 4.9 and 4.10 is to check how much the shape is modified relative to the nominal case of P_{31} when we vary the value of

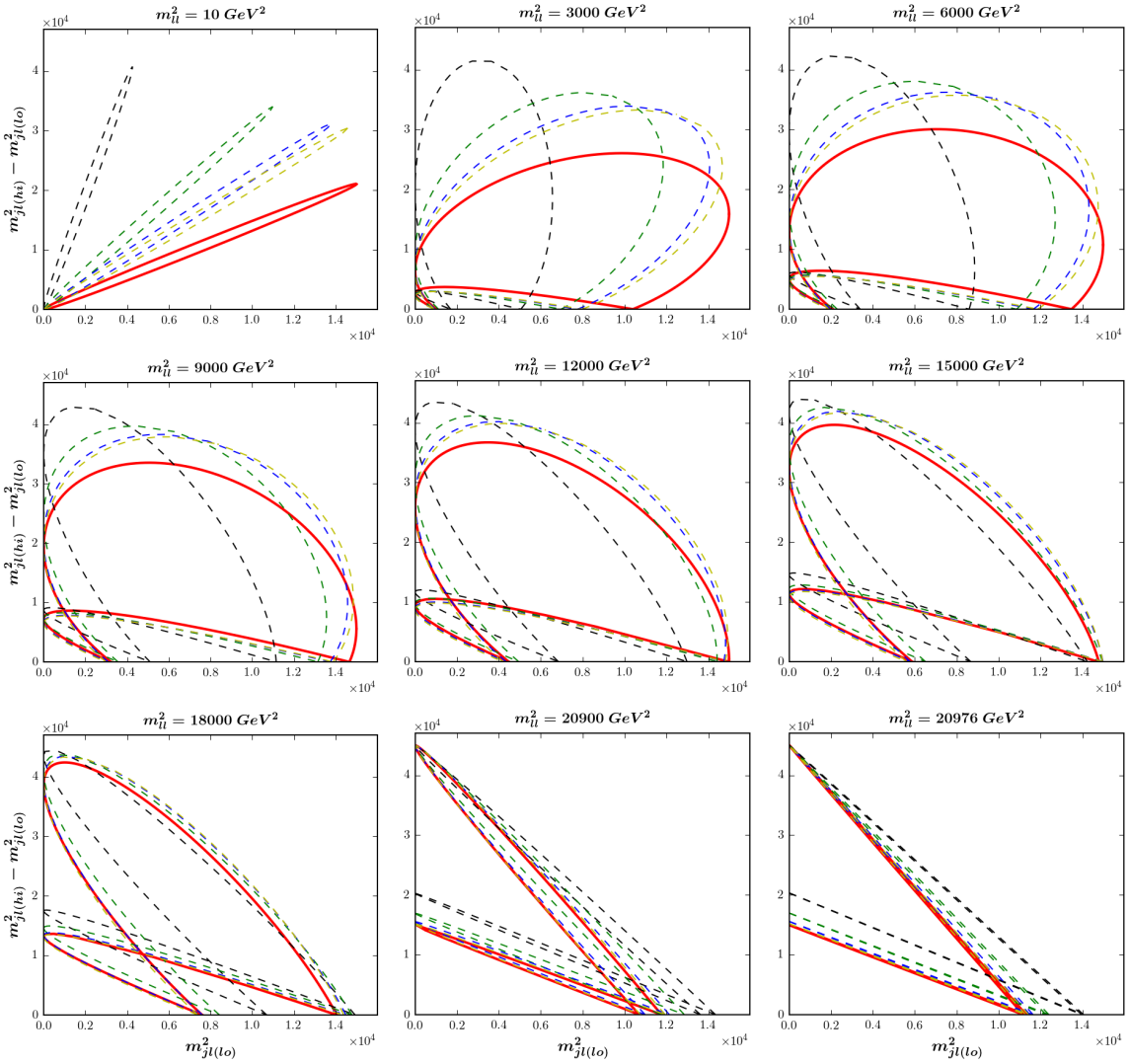


Figure 4.10: The same as Fig. 4.9, but for the auxiliary branch going through regions $(2,3)$ and $(4,3)$. The dashed lines represent points with $\tilde{m}_A = 100$ GeV (black), $\tilde{m}_A = 500$ GeV (green), $\tilde{m}_A = 2000$ GeV (blue) and $\tilde{m}_A = 6000$ GeV (yellow). For reference, we also show the case of the true mass spectrum for point P_{31} (red solid lines), although P_{31} does not belong to the auxiliary branch.

$\tilde{m}_A$ along the flat direction (4.7). The dashed lines in Fig. 4.9 show results for several representative values of $\tilde{m}_A$ along the true branch: $\tilde{m}_A = 0$ (black), $\tilde{m}_A = 100$ GeV (gray), $\tilde{m}_A = 500$ GeV (green), $\tilde{m}_A = 1000$ GeV (blue), $\tilde{m}_A = 2000$ GeV (yellow) and $\tilde{m}_A = 5000$ GeV (magenta). We observe noticeable shape variations, especially at low to intermediate values of $m^2_{\ell\ell}$, which bodes well for our intended purpose of

measuring the value of m_A. Fig. 4.9 aids in visualizing why sensitivity is lost when performing one-dimensional projections. Consider, for example the variable $m_{j\ell(lo)}$. The top two rows of Fig. 4.9 show that as $\tilde{m}_A$ is varied along the flat direction, the boundary contours are being stretched vertically, which does not have any effect on the $m_{j\ell(lo)}$ endpoint. Later on, when the events are projected vertically on the $m_{j\ell(lo)}$ axis to obtain the $m_{j\ell(lo)}$ distribution seen in the middle panel of Fig. 4.6, the effects from this vertical stretching tend to be washed out and the resulting $m_{j\ell(lo)}$ distributions have very similar shapes.

Fig. 4.10 shows the analogous results for the auxiliary branch. Once again, the red solid lines represent the study point P_{31}, while the dashed lines correspond to four values of $\tilde{m}_A$: $\tilde{m}_A = 100$ GeV (black), $\tilde{m}_A = 500$ GeV (green), $\tilde{m}_A = 2000$ GeV (blue) and $\tilde{m}_A = 6000$ GeV (yellow). This time the shape variation along the flat direction is much more significant compared to what we saw in Fig. 4.9. This observation agrees with our expectation based on Fig. 4.8 that points on the auxiliary branch behave quite differently from our nominal study point P_{31}, especially at low $\tilde{m}_A$.

4.3.2 A toy study with uniformly distributed background

In the remainder of this section we shall illustrate our proposed method for mass measurement with two exercises. In each case, we shall assume that the standard set of one-dimensional kinematic endpoints (4.2) has already been well measured and used to reduce the relevant mass parameter space (4.1) to the flat direction (4.7) parametrized by the test mass $\tilde{m}_A$ for the lightest new particle A. This is done only for simplicity — in principle, our method would also work without any prior information from endpoint measurements, but by using those, we are reducing the 4-dimensional optimization problem in (4.17) to the much simpler one-dimensional optimization problem

$$\max_{\tilde{m}_A} \bar{\Sigma}\left(\tilde{m}_A, \tilde{m}_B(\tilde{m}_A), \tilde{m}_C(\tilde{m}_A), \tilde{m}_D(\tilde{m}_A)\right) \simeq \bar{\Sigma}(m_A, m_B, m_C, m_D), \tag{4.37}$$

where $\tilde{m}_B(\tilde{m}_A)$, $\tilde{m}_C(\tilde{m}_A)$, and $\tilde{m}_D(\tilde{m}_A)$ are the masses of particles B, C and D along the flat direction. Our main emphasis here is on demonstrating the advantages of our method relative to the method of kinematic endpoints. In Section 4.3.1 we already showed that while the method of kinematic endpoints does a good job in reducing the

unknown mass parameter space (4.1) to the flat direction (4.7), it does a poor job of lifting the degeneracy along the flat direction. Thus, if we can show that our method can perform the remaining mass measurement along the flat direction, we will have accomplished our goal.

In order to make contact with our previous studies in [134], we begin with a simple toy exercise where in addition to the signal events from the cascade decay in Fig. 4.1, we also consider a certain number of background events, which we take to be uniformly distributed in the mass squared space of observables (4.36). While the assumption of uniform background density is unrealistic, such an exercise is nevertheless worth studying for several reasons. First, our method is completely general and applies in any situation where we have a decay of the type shown in Fig. 4.1, while to correctly identify the relevant backgrounds, we must be a lot more specific — we need to fix the signature, the type of production mechanism (which determines what else is in the event), the cuts, etc. In order to retain generality, we choose to avoid specifying those details and instead we generate background events by pure Monte Carlo according to a flat hypothesis. Second, as shown in [134], a uniform background distribution is actually a pretty good approximation to more realistic backgrounds resulting, e.g., from dilepton $t\bar{t}$ events (compare to the results in Section 4.3.3 below). Finally, our method is attempting to detect a discontinuity in the measured event density caused by a signal kinematic boundary, so the exact shape of the background distribution is not that important, as long as it is smooth and without any sharp kinematic features.

In order to detect the exact location of the kinematic boundary, we shall be computing the quantity $\bar{\Sigma}$ defined in (4.16) along the flat direction (4.7), i.e.

$$\bar{\Sigma}(\tilde{m}_A) \equiv \bar{\Sigma}\left(\tilde{m}_A, \tilde{m}_B(\tilde{m}_A), \tilde{m}_C(\tilde{m}_A), \tilde{m}_D(\tilde{m}_A)\right). \tag{4.38}$$

We shall perform several versions of the exercise, with varying levels of signal-to-background. For this purpose, we vary the ratio of signal to background events inside the true "samosa" surface $\mathcal{S}(m_A, m_B, m_C, m_D)$:

$$S/B \equiv \frac{\int_{V_\mathcal{S}} \rho_s \, dV}{\int_{V_\mathcal{S}} \rho_b \, dV}, \tag{4.39}$$

where $V_\mathcal{S}$ is the volume inside the samosa $\mathcal{S}(m_A, m_B, m_C, m_D)$, while ρ_s and ρ_b are the signal and background event densities from Section 4.1, respectively. In this exercise, we shall fix the overall normalization by choosing $N_B = 1000$ background events inside $\mathcal{S}$.

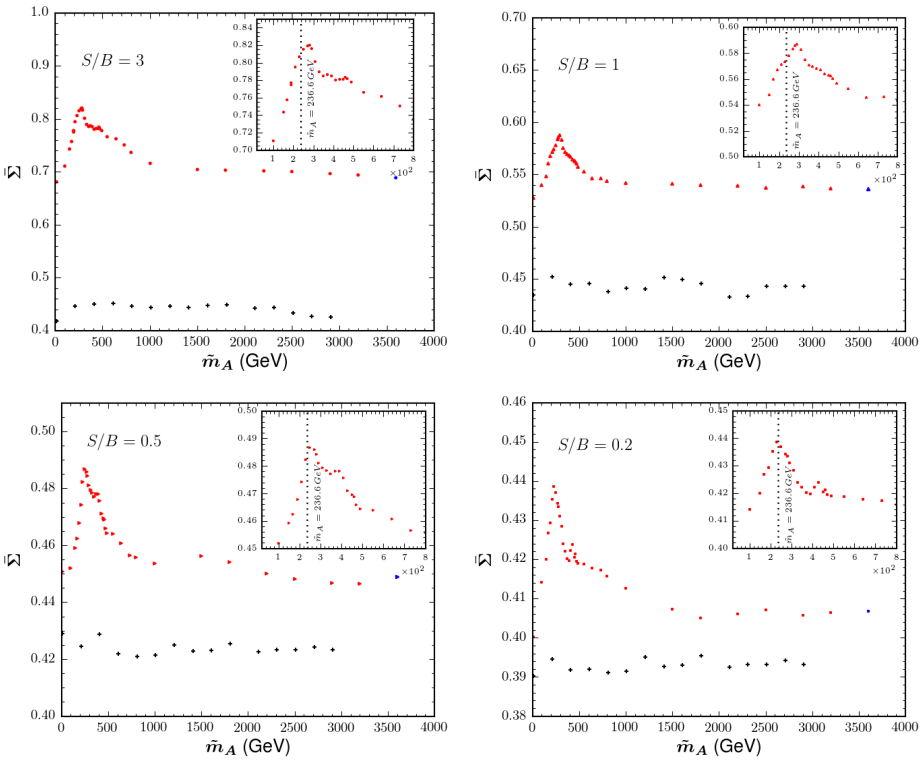


Figure 4.11: The quantity $\bar{\Sigma}(\tilde{m}_A)$ defined in (4.38) as a function of $\tilde{m}_A$ for different values of the signal to background ratio S/B defined in (4.39): $S/B = 3$ (upper left panel), $S/B = 1$ (upper right panel), $S/B = 0.5$ (lower left panel) and $S/B = 0.2$ (lower right panel). The colored symbols correspond to the true branch with the color conventions from Fig. 4.2, while the black crosses indicate points on the auxiliary branch. The insert on each panel zooms in on the region near the peak value for $\bar{\Sigma}(\tilde{m}_A)$.

Our main result is shown in Fig. 4.11, which plots the quantity $\bar{\Sigma}(\tilde{m}_A)$ along the flat direction, for several different choices of S/B: $S/B = 3$ (upper left panel), $S/B = 1$ (upper right panel), $S/B = 0.5$ (lower left panel) and $S/B = 0.2$ (lower right panel). Each panel contains two sets of points: the colored symbols represent points on the true branch, while the black crosses indicate points on the auxiliary branch[15].

There are several important lessons from Fig. 4.11:

- *Viability of the method.* We see that in each panel, the maximum of $\bar{\Sigma}$ is

[15]Recall from Fig. 4.4 that for any given choice of $\tilde{m}_A$, there is one point on the true branch and a corresponding point on the auxiliary branch.

obtained for a value of $\tilde{m}_A$ which is close to the true value $m_A = 236.6$ GeV. This validates our conjecture[16], eq. (4.17), and proves the viability of our method.

- *Precision of the method.* Of course, we did not recover *exactly* the input value for m_A, but in each case, came relatively close. Each panel of Fig. 4.11 contains an insert which zooms in on the region near the peak, which is sampled more finely. For the different values of $S/B = \{3.0, 1.0, 0.5, 0.2\}$, the $\bar{\Sigma}$ maxima are obtained at $\tilde{m}_A = \{280, 290, 250, 230\}$ GeV, correspondingly. Since the measurement is not perfect, it may be instructive to compare the theoretical boundary for the input study point P_{31} to the boundary surface found by the fit. This is illustrated in Fig. 4.12, where in analogy to Figs. 4.9 and 4.10 we show two-dimensional slices at fixed $m^2_{\ell\ell}$ of the Voronoi tessellation of the data for the case of $S/B = 3$. The Voronoi cells are color coded by their value of $\bar{\sigma}_i$ defined in (4.12). As in Figs. 4.9 and 4.10, the red solid line in each panel is the expected signal boundary for the nominal case of point P_{31}. We notice that the cells with the highest values of $\bar{\sigma}_i$ are indeed clustered near the nominal boundary, in agreement with the results from Refs. [134, 17]. On the other hand, the boundary delineated by the black dashed lines in Fig. 4.12 corresponds to the best fit value of $\tilde{m}_A = 280$ GeV, which was found in the upper left panel in Fig. 4.11. The difference between the solid red and black dashed contours in Fig. 4.12 is essentially a measure of the resolution of our method.

- *Elimination of the auxiliary branch and the large $\tilde{m}_A$ tail of the true branch.* One very positive piece of news from Fig. 4.11 is that the whole auxiliary branch has very low values for $\bar{\Sigma}$ which makes it easy to rule it out — one can see that no point on the auxiliary branch was ever in contention for the top spot. Similar comments, albeit to a lesser extent, also apply to the long tail along the true branch at large $\tilde{m}_A$. In particular, region $(4, 1)$ seems to be ruled out, as well as the large $\tilde{m}_A$ portion of region $(3, 1)$. In effect, the range of possible values for $\tilde{m}_A$ along the flat direction has been significantly narrowed down to a small interval within a few tens of GeV of the true value m_A.

- *The adverse effect of the background.* Comparing the different panels in Fig. 4.11, we see that as we make S/B smaller, the difference between the true and aux-

[16]Strictly speaking, Fig. 4.11 tests only the one-dimensional version (4.37).

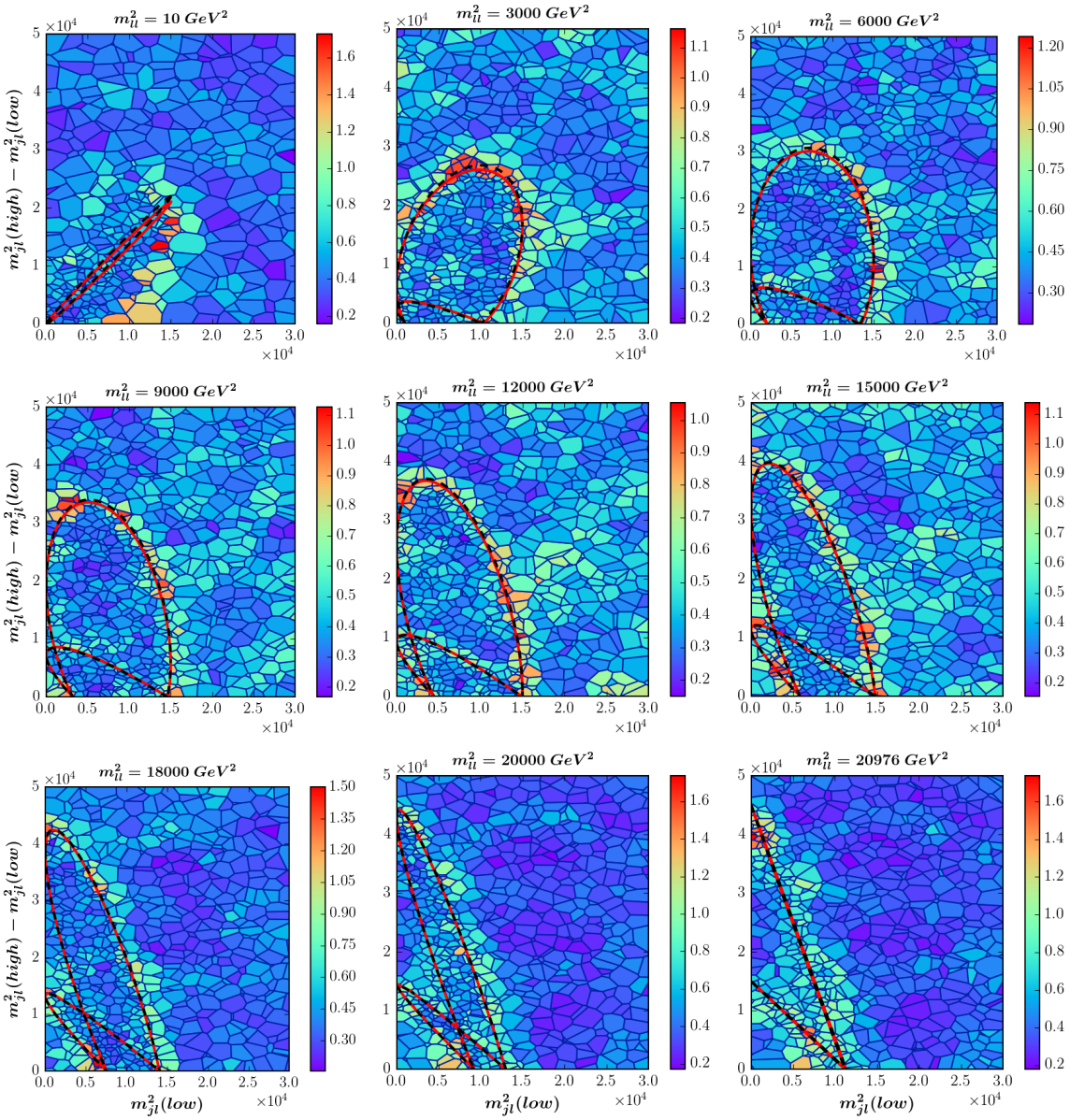


Figure 4.12: Two-dimensional views at fixed $m_{\ell\ell}^2$ of the Voronoi tessellation of the data for the case of $S/B = 3$. The red solid line is the expected signal boundary for the nominal case of point P_{31}, i.e., with the true value $\tilde{m}_A = m_A = 236.6$ GeV. The black dashed line corresponds to the mass spectrum with $\tilde{m}_A = 280$ GeV, which was found to maximize the quantity $\bar{\Sigma}$ in the top left panel of Fig. 4.11.

iliary branch is reduced, but the auxiliary branch is still disfavored. As for the true branch, the peak near m_A still persists, even in the case when the data is dominated by background events. This is not surprising, since the background

distribution is relatively smooth, so that in the background-dominated regions of phase space there aren't too many Voronoi cells with large values of $\bar{\sigma}_i$, which could adversely affect the fit.

4.3.3 A study with $t\bar{t}$ dilepton background events

We are now in position to repeat the exercise from Section 4.3.2, with signal events from $D + A$ associated production and background taken from dilepton $t\bar{t}$ events, which represent the main background to the signature from Fig. 4.1 of a jet plus two opposite sign, same flavor leptons (the electroweak backgrounds involving leptonic Z decays can be suppressed with a Z mass veto). Events were generated at parton level for LHC at 14 TeV with MadGraph5 [140] version 2.1.1 with the default PDF set `cteq6l1`. For signal we used the SUSY version of the cascade decay in Fig. 4.1, and considered the associate production of a squark $\tilde{q}$ with the lightest neutralino $\tilde{\chi}_1^0$, namely $pp \to \tilde{q}\tilde{\chi}_1^0$ [141, 142]. Since each $t\bar{t}$ background event contains two jets, there is a two-fold ambiguity in the jet selection. We will use both possible pairings, so that each background event will contribute two entries to our data. Of course, we do not know *a priori* how many of those entries will end up inside the nominal boundary surface $\mathcal{S}(m_A, m_B, m_C, m_D)$, which is why we have to use a slightly different normalization from Sec 4.3.2. We shall fix the number of signal events to $N_S = 3000$, and then we shall consider several values[17] for the number of dilepton $t\bar{t}$ events: $N_B = \{3000, 4000, 5000, 6000\}$. From Monte Carlo we then find that these choices correspond to $S/B = \{1.52, 1.14, 0.91, 0.76\}$ inside the $\mathcal{S}$ boundary, see (4.39).

Our main result is shown in Fig. 4.13, which is the analogue of Fig. 4.11 for this case. Once again, we find that the function $\bar{\Sigma}(\tilde{m}_A)$ is maximized in the vicinity of $\tilde{m}_A = m_A = 236.6$ GeV. The fitting procedure is illustrated in Fig. 4.14, which is the analogue of Fig. 4.12. The red solid lines show the boundary contours for the nominal value of $m_A = 236.6$ GeV, while the black dashed lines are for the best fit value of $\tilde{m}_A = 280.0$ GeV, which was found in the top left panel of Fig. 4.13. Fig. 4.13 shows that once again, our procedure has disfavored the whole auxiliary branch and narrowed down the range of viable values of $\tilde{m}_A$ to a few tens of GeV around the nominal value m_A.

[17] The anticipated signal-to-background ratio is model-dependent. In this sense, SUSY may not be the best case for discovery, since other scenarios, e.g., UED [143, 144, 145], have higher signal cross-sections.

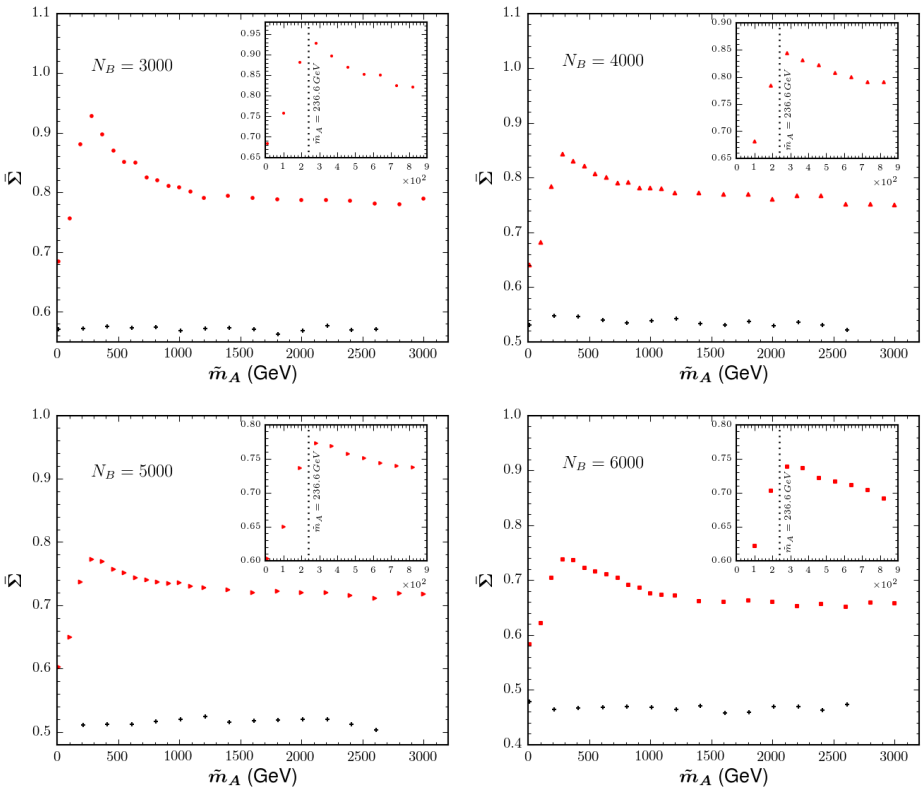


Figure 4.13: The analogue of Fig. 4.11 for the exercise with $t\bar{t}$ background events considered in Section 4.3.3. Results are shown for $N_S = 3000$ signal events and several choices for the number of background events: $N_B = 3000$ (upper left panel), $N_B = 4000$ (upper right panel), $N_B = 5000$ (lower left panel) and $N_B = 6000$ (lower right panel).

4.4 A case study in region $(3, 2)$

In this section, we shall repeat the analysis from Section 4.3, only this time our nominal study point, from now on labelled as P_{32}, will be chosen within the cyan region $(3, 2)$ of Fig. 4.2. Recall that the problematic relation (4.29) was satisfied in three of the colored regions in Fig. 4.2, namely $(2, 3)$, $(3, 1)$ and $(3, 2)$. The former two regions, $(2, 3)$ and $(3, 1)$, were already visited by the mass trajectory studied in Section 4.3, thus here for completeness we will also illustrate the case of region $(3, 2)$.

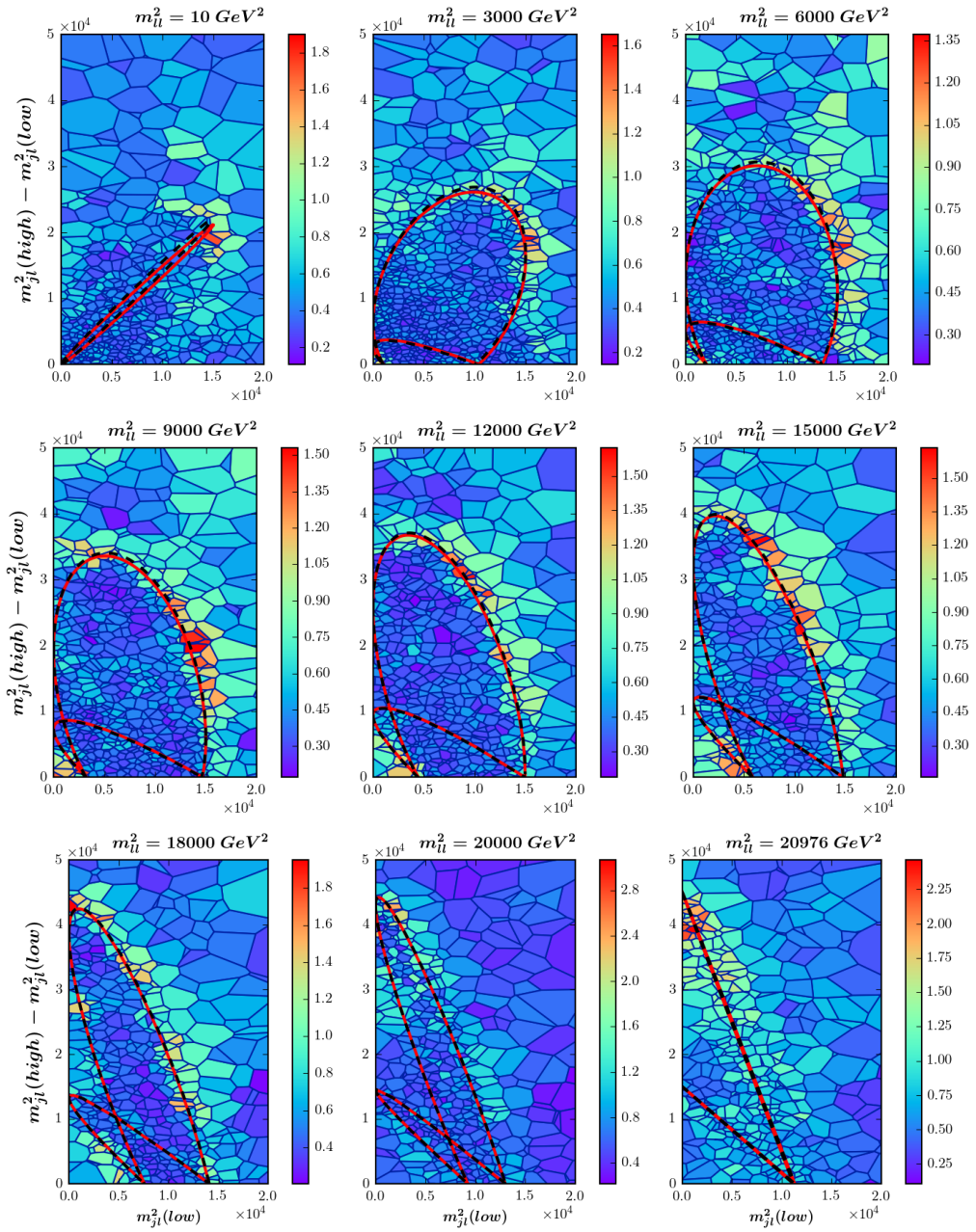


Figure 4.14: The analogue of Fig. 4.12 for the exercise with $t\bar{t}$ background events considered in Section 4.3.3. The Voronoi tessellation was done for the case of $N_B = 3000$. The red solid line is the phase space boundary for the nominal value $m_A = 236.6$ GeV, while the black dashed line corresponds to the best fit value $\tilde{m}_A = 280.0$ GeV found in the the top left panel of Fig. 4.13.

		true branch		mirror branch		
Region		$(3,2)$	$(3,1)$	$(4,2)$	$(4,3)$	$(2,3)$
Study point		P_{32}	P_{31}	P_{42}	P_{43}	P_{23}
m_A (GeV)		126.49	5000.00	90.00	150.00	500.00
m_B (GeV)		282.84	5207.42	194.61	250.96	609.04
m_C (GeV)		447.21	5324.17	399.99	460.80	815.72
m_D (GeV)		500.00	5372.07	458.78	518.78	869.36
R_{AB}		0.200	0.922	0.214	0.357	0.674
R_{BC}		0.400	0.957	0.237	0.297	0.557
R_{CD}		0.800	0.982	0.760	0.789	0.880
m_{ll}^{max} (GeV)	$\sqrt{a}$	309.84				
m_{jll}^{max} (GeV)	$\sqrt{b}$	368.78				
$m_{jl(lo)}^{max}$ (GeV)	$\sqrt{c}$	149.07				
$m_{jl(hi)}^{max}$ (GeV)	$\sqrt{d}$	200.00	200.00	198.23	199.87	200.00
$m_{jll(\theta>\frac{\pi}{2})}^{min}$ (GeV)	$\sqrt{e}$	247.94	237.47	253.72	250.99	243.81

Table 4.2: The same as Table 4.1, except now the starting point is a point (P_{32}) from region $(3,2)$.

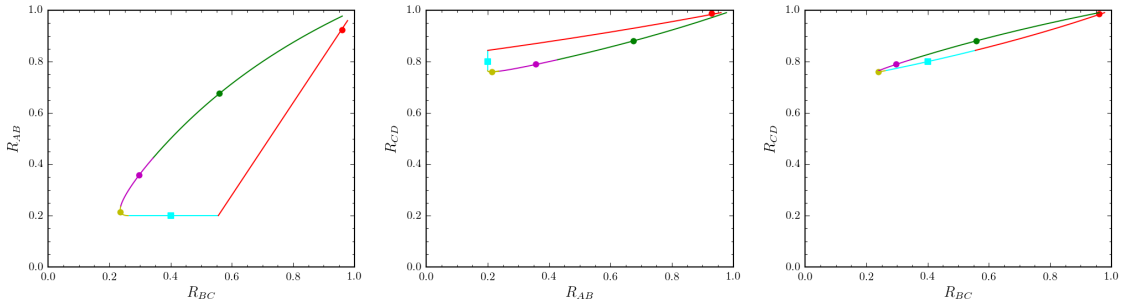


Figure 4.15: The same as Fig. 4.3, but for the flat direction generated by point P_{32} from Table 4.2.

4.4.1 Kinematical properties along the flat direction

The mass spectrum for the study point P_{32} and the corresponding mass squared ratios and kinematic endpoints are shown in the cyan-shaded column of Table 4.2. Point P_{32} was used previously in Ref. [36] as an example of a *discrete* two-fold ambiguity, while here it serves to define a flat direction (4.7) in mass parameter space. This flat direction is illustrated in Figs. 4.15 and 4.16, which are the analogues of Figs. 4.3

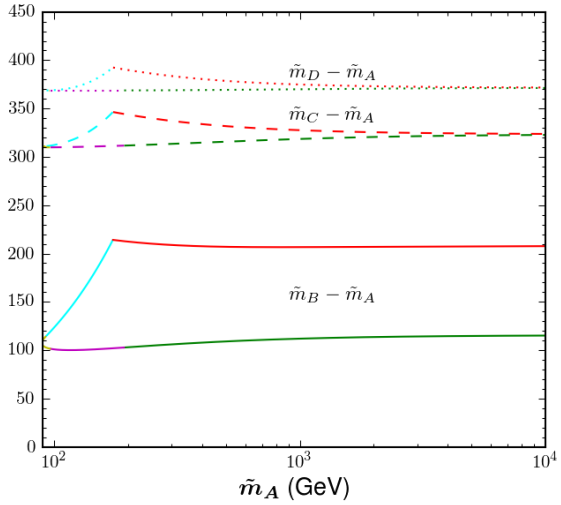


Figure 4.16: The analogue of Fig. 4.4, but for the flat direction defined in Fig. 4.15.

and 4.4, respectively. According to Figs. 4.15 and 4.16, the mass trajectory now goes through five of the six colored regions in Fig. 4.2: there is a true branch through the red region $(3,1)$, the cyan region $(3,2)$ and the the yellow region $(4,2)$, as well as an auxiliary branch through the yellow region $(4,2)$, the magenta region $(4,3)$ and the green region $(2,3)$. As in Section 4.3, we choose one representative study point in each of these regions. The four additional study points, P_{31}, P_{42}, P_{43} and P_{23}, are also listed in Table 4.2, and their columns are shaded with the color of their respective regions in Fig. 4.2.

The flat direction depicted in Fig. 4.15 is again parametrized by the trial value $\tilde{m}_A$ for the mass of the lightest new particle A. However, as seen in Fig. 4.16, this time the allowed range for $\tilde{m}_A$ does not extend all the way to $\tilde{m}_A = 0$, and instead the true and auxiliary branch meet inside the yellow region $(4,2)$ around at the lowest value $\tilde{m}_A \sim 89$ GeV.

The transitions between two neighboring regions along the flat direction can be understood from Fig. 4.17, which plots the mass squared ratios R_{AB}, R_{BC} and R_{CD} (solid lines) and several other quantities (dotted lines) which are relevant for defining the regions from Fig. 4.2, as a function of the mass trajectory parameter $\tilde{m}_A$. For example, relations (B.3) and (B.13) imply that the boundary between the cyan region $(3,2)$ and the red region $(3,1)$ is given by $R_{BC} = (2 - R_{AB})^{-1}$. Indeed, the lines in Figs. 4.16 and 4.17 change color from cyan to red when the R_{BC} curve crosses the

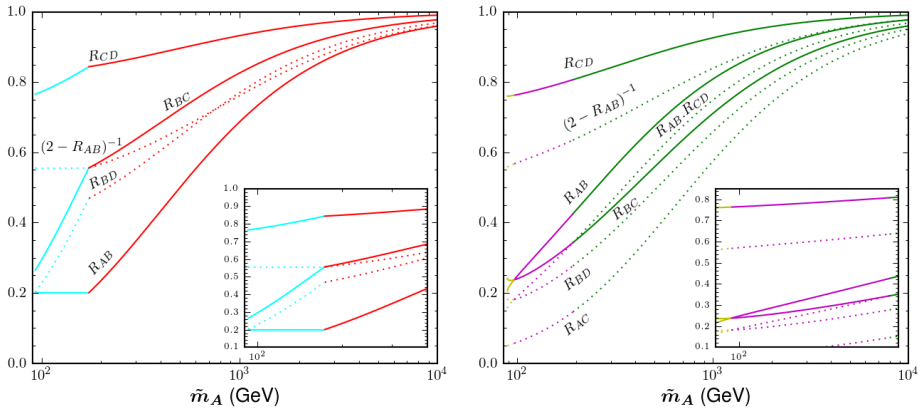


Figure 4.17: The equivalent representation of Fig. 4.16 in terms of the mass squared ratios R_{AB}, R_{BC} and R_{CD} (solid lines). The dotted lines depict various quantities of interest which are used to delineate the regions in Fig. 4.2. The left panel shows the true branch passing through regions $(3,2)$ (cyan) and $(3,1)$ (red), while the right panel shows the auxiliary branch through regions $(4,2)$ (yellow), $(4,3)$ (magenta) and $(2,3)$ (green). The left insert zooms in on the transition between regions $(3,2)$ and $(3,1)$ near $\tilde{m}_A = 173$ GeV, while the right insert focuses on the transitions between regions $(4,2)$ and $(4,3)$ near $\tilde{m}_A = 97$ GeV and between regions $(4,3)$ and $(2,3)$ near $\tilde{m}_A = 193$ GeV.

dotted line representing the function $(2 - R_{AB})^{-1}$ near $\tilde{m}_A = 173$ GeV. Similarly, it follows from (B.21) and (B.54) that the boundary between the green region $(2,3)$ and the magenta region $(4,3)$ is given by $R_{BC} = R_{AB}R_{CD}$. The right panel of Fig. 4.17 confirms that the line color changes from magenta to green when the solid line for R_{BC} is intersected by the dotted line for $R_{AB}R_{CD}$. Finally, according to (B.42) and (B.53), the transition between the yellow region $(4,2)$ and the magenta region $(4,3)$ occurs at $R_{AB} = R_{BC}$, and this is borne out by Fig. 4.17 as well.

By construction, all five study points from Table 4.2 predict identical values for the three kinematic endpoints a, b and c. This is demonstrated in Fig. 4.18, which is the analogue of Fig. 4.6, but for the five study points from Table 4.2. As before, the distributions in Fig. 4.18 are color-coded according to our color conventions from Fig. 4.2. The nominal input study point P_{32} is represented by the solid line, while the dotted lines mark the other four study points. Given that the five study points look very similar on Fig. 4.18, we now focus on the remaining two distributions, $m_{j\ell(hi)}$ and $m_{jll(\theta>\frac{\pi}{2})}$, which are investigated in Figs. 4.19 and 4.20. The left panels show the predictions for the kinematic endpoint $\sqrt{d} = m^{max}_{j\ell(hi)}$ and the threshold

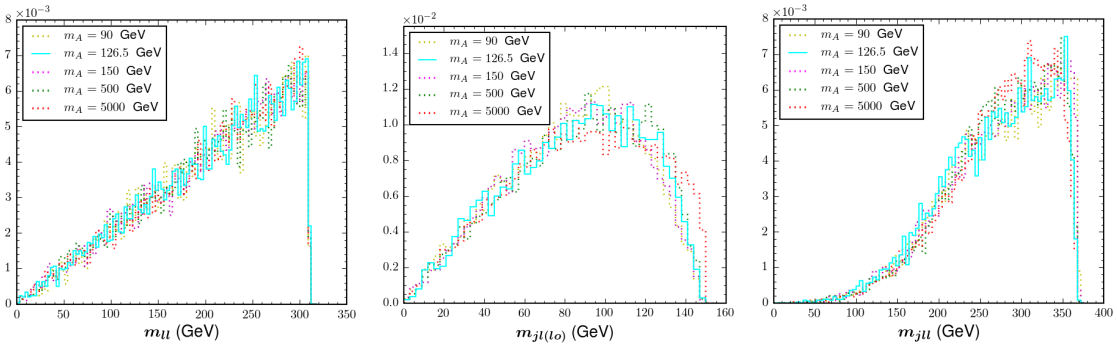


Figure 4.18: The analogue of Fig. 4.6, but for the five study points exhibited in Table 4.2. The distributions are color-coded according to our color conventions for the regions in Fig. 4.2.

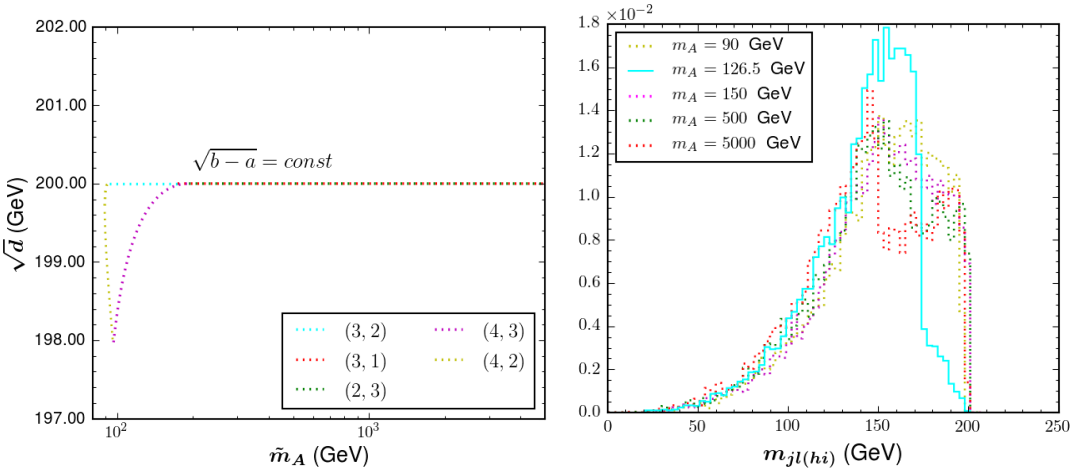


Figure 4.19: The analogue of Fig. 4.19, but for the flat direction defined in Fig. 4.15 (left panel) and for the five study points from Table 4.2 (right panel).

$\sqrt{e} = m^{min}_{j\ell\ell(\theta > \frac{\pi}{2})}$, respectively, while the right panels plot the corresponding kinematic distributions for the five study points from Table 4.2.

By now, we should not be surprised by the extreme flatness of the curves exhibited in the left panel of Fig. 4.19. The mass trajectory from Fig. 4.15 passes through all three of the regions where the endpoint d is not an independent quantity, but is fixed by the relation (4.29) and is therefore strictly independent of $\tilde{m}_A$. In the remaining two regions, $(4, 2)$ and $(4, 3)$, Fig. 4.19 shows a maximal deviation of only 2 GeV from the prediction $\sqrt{d} = \sqrt{b-a}$ of (4.29). Taken together, the left panels of Figs. 4.7 and 4.19 justify our terminology of the mass trajectory (4.7) as a "flat direction" in

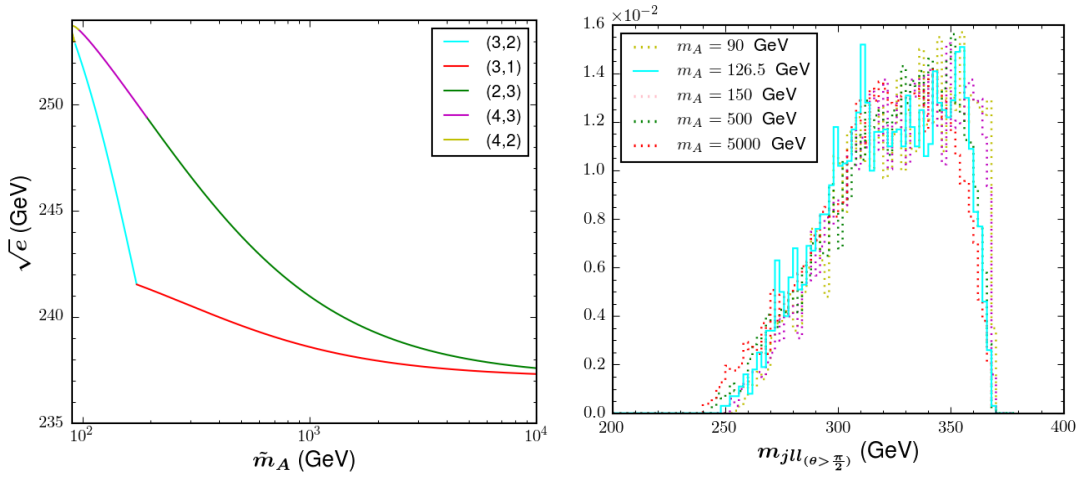


Figure 4.20: The same as Fig. 4.19, but for the kinematic endpoint $\sqrt{e}$ and the corresponding $m_{jll(\theta>\frac{\pi}{2})}$ distribution.

mass parameter space.

On the other hand, the left panel of Fig. 4.20 shows a much more significant variation of the kinematic threshold variable $\sqrt{e}$ along the flat direction. The total variation is on the order of 15 GeV, which is of the same order as our previous result in Fig. 4.8. However, as we already discussed in Section 4.3, the measurement of $\sqrt{e}$ presents significant experimental challenges, as one can deduce from the very minor apparent variation of the $m_{j\ell\ell(\theta>\frac{\pi}{2})}$ distributions shown in the right panel of Fig. 4.20. This motivates searching for alternative methods for lifting the degeneracy along the flat direction.

As already discussed in Section 4.3, one such method is to track the deformation of the shape of the kinematic boundary (4.9) along the flat direction. The effect is illustrated in Figs. 4.21 and 4.22, which are the analogues of Figs. 4.9 and 4.10 for the example of a flat direction considered in this section. Once again, the solid red lines in each panel indicate the kinematic boundaries for the nominal study point P_{32} with $\tilde{m}_A = 126.5$ GeV, while the dashed lines are drawn for several other values of $\tilde{m}_A$, chosen so that they illustrate the typical range of shape fluctuations. Along the true branch, in Fig. 4.21 we plot contours for $\tilde{m}_A = 100$ (black), $\tilde{m}_A = 173$ GeV (green), $\tilde{m}_A = 500$ GeV (blue), $\tilde{m}_A = 2000$ GeV (yellow) and $\tilde{m}_A = 4000$ GeV (gray). Even though we are confined to the true branch only, when we compare Fig. 4.21 to its analogue, Fig. 4.9, we observe a much larger variation in the shape of

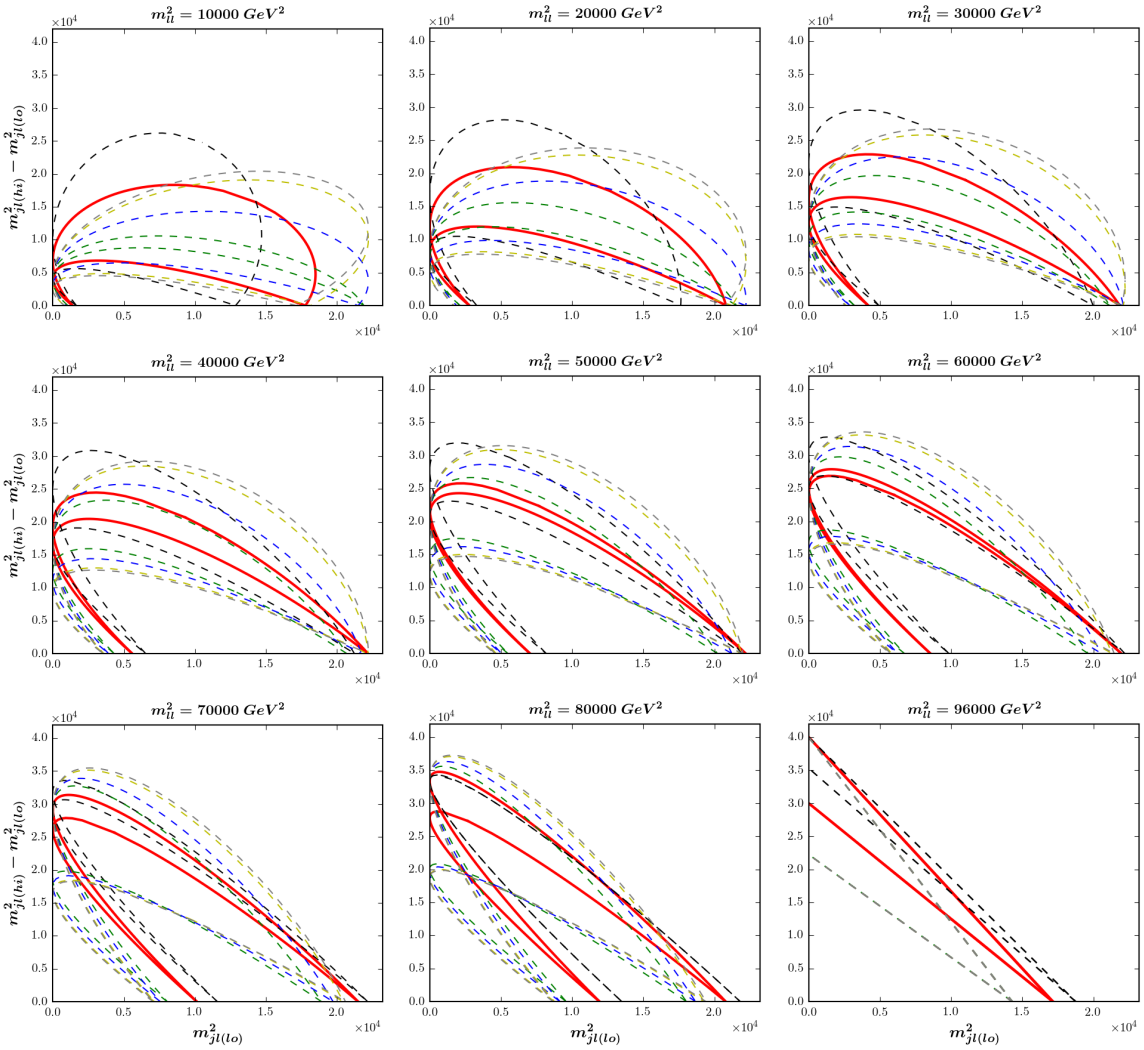


Figure 4.21: The same as Fig. 4.9, but for the true branch in Fig. 4.15. The red solid line represents the case of the P_{32} study point with $\tilde{m}_A = 126.5$ GeV, while the dashed lines correspond to other values of $\tilde{m}_A$ along the true branch: $\tilde{m}_A = 100$ (black), $\tilde{m}_A = 173$ GeV (green), $\tilde{m}_A = 500$ GeV (blue), $\tilde{m}_A = 2000$ GeV (yellow) and $\tilde{m}_A = 4000$ GeV (gray).

the kinematic boundary in the present case, which promises good prospects for the mass measurement exercise to follow. The results shown in Fig. 4.22 for the auxiliary branch are also quite good. This should not come as a surprise, since the exercise in Section 4.3 already indicated that the auxiliary branch has a different kinematic behavior, as reflected in the shape of the phase space boundary.

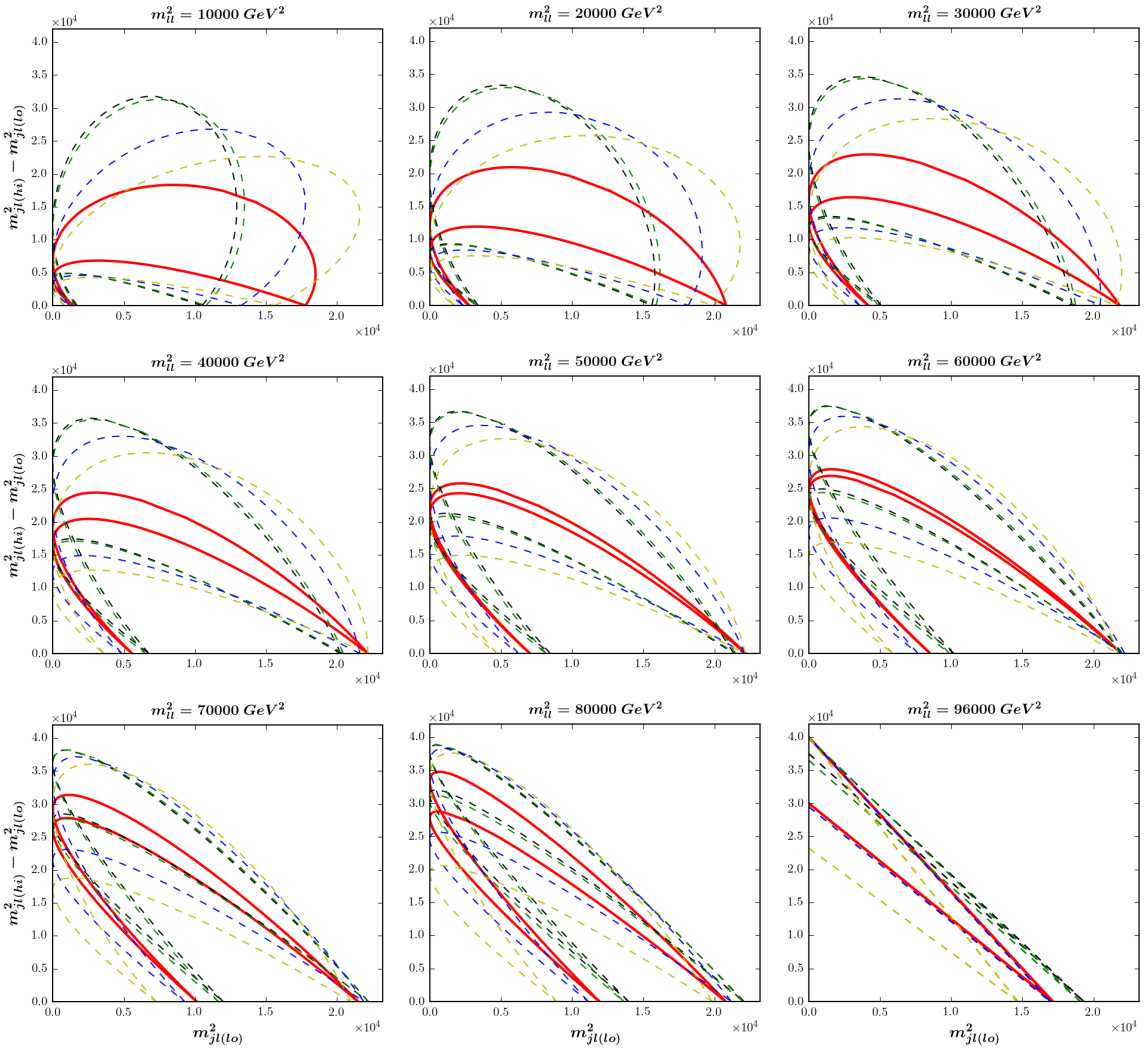


Figure 4.22: The same as Fig. 4.21, but for the auxiliary branch in regions $(4,2)$, $(4,3)$ and $(2,3)$. The dashed lines represent points with $\tilde{m}_A = 90$ GeV (black), $\tilde{m}_A = 150$ GeV (green), $\tilde{m}_A = 500$ GeV (blue) and $\tilde{m}_A = 5000$ GeV (yellow). For reference, we also show the case of the true mass spectrum for point P_{32} (red solid lines), although P_{32} itself does not belong to the auxiliary branch.

4.4.2 A toy study with uniformly distributed background

In the remainder of Section 4.4 we shall repeat the two exercises from Sections 4.4.2 and 4.4.3, only this time we shall use P_{32} as our input study point, and perform the measurement along the corresponding flat direction described in Figs. 4.15 and 4.16.

First we consider the case of uniformly distributed (in mass squared) background

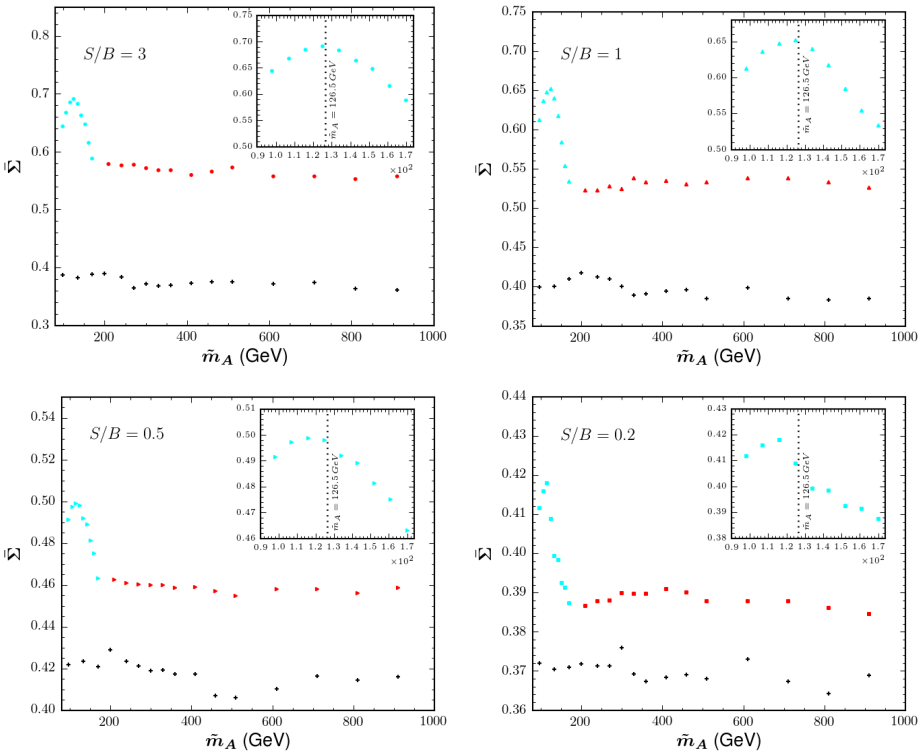


Figure 4.23: The same as Fig. 4.11, but now taking point P_{32} as input and measuring along the flat direction depicted in Fig. 4.15.

events, and proceed to evaluate the quantity $\bar{\Sigma}$ along the flat direction. As before, we fix $N_B = 1000$ and then vary the signal-to-background ratio inside the boundary surface $\mathcal{S}$. Fig. 4.23 shows our results for the same choices of S/B as in Fig. 4.11: $S/B = 3$ (upper left panel), $S/B = 1$ (upper right panel), $S/B = 0.5$ (lower left panel) and $S/B = 0.2$ (lower right panel). We find that the function $\bar{\Sigma}(\tilde{m}_A)$ once again peaks in the vicinity of the true value $m_A = 126.5$ GeV. Specifically, for $S/B = \{3.0, 1.0, 0.5, 0.2\}$, the maxima are found at $\tilde{m}_A = \{125, 125, 116, 116\}$ GeV, to be contrasted with the true value of $m_A = 126.5$ GeV. In all four cases, the auxiliary branch is disfavored, as it always gives low values for $\bar{\Sigma}$, while the true branch is restricted to a very narrow region near the true mass spectrum.

Fig. 4.24 provides a consistency check on our fitting procedure, similarly to Fig. 4.12. We show two-dimensional views at fixed $m^2_{\ell\ell}$ of the Voronoi tessellation of the data for the case of $S/B = 3$. The red solid line is the expected signal boundary for the nominal case of point P_{32}, i.e., with the true value $\tilde{m}_A = m_A = 126.5$ GeV. The black

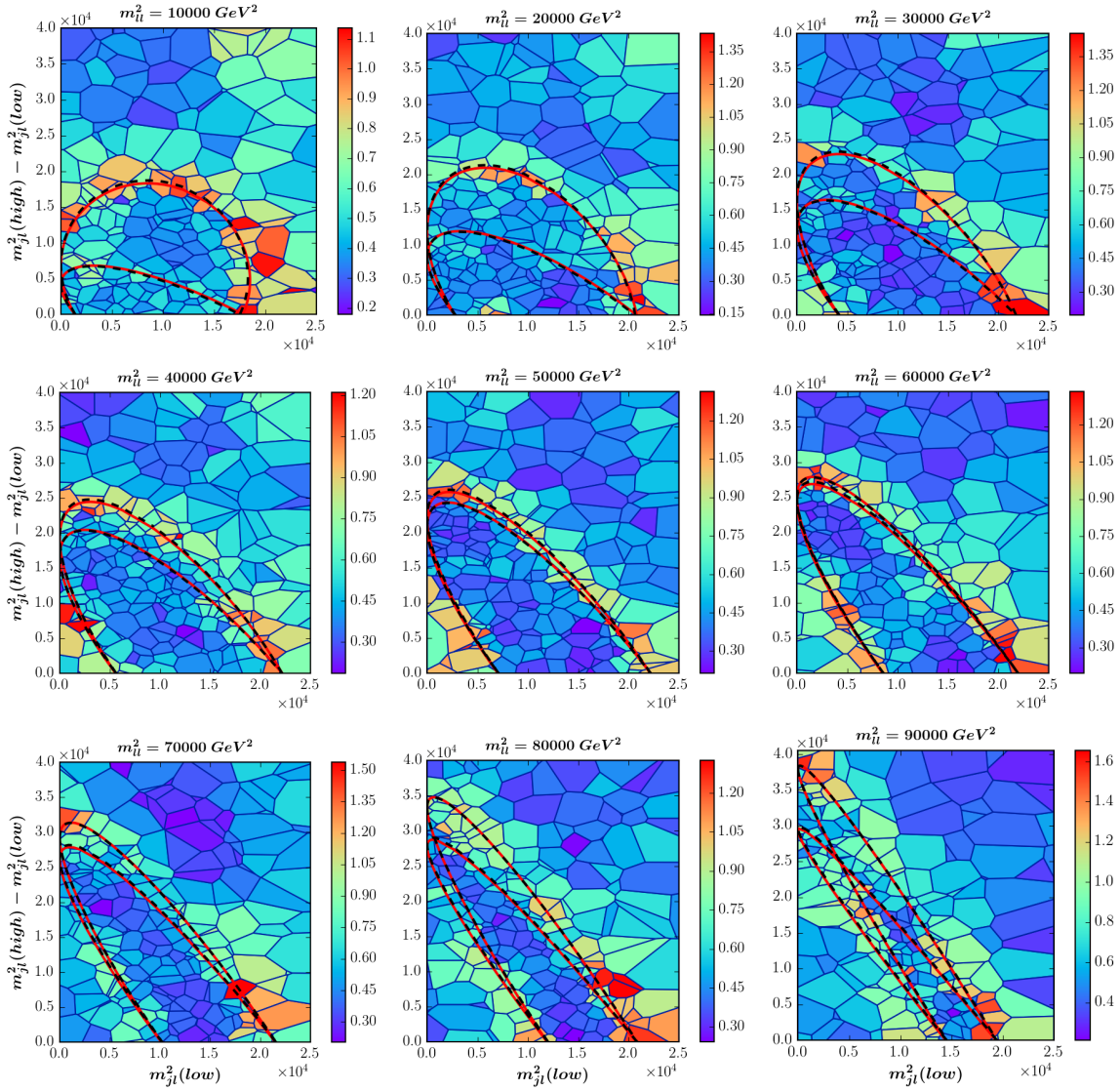


Figure 4.24: The same as Fig. 4.12, but for the exercise performed in Section 4.4.2, with point P_{32} as input (solid red lines). The black dashed line corresponds to the mass spectrum with $\tilde{m}_A = 125$ GeV, which was found to maximize the quantity $\bar{\Sigma}$ in the top left panel of Fig. 4.23.

dashed line then corresponds to the best fit, i.e., a mass spectrum with $\tilde{m}_A = 126$ GeV, which was found to maximize the quantity $\bar{\Sigma}$ in the top left panel of Fig. 4.23.

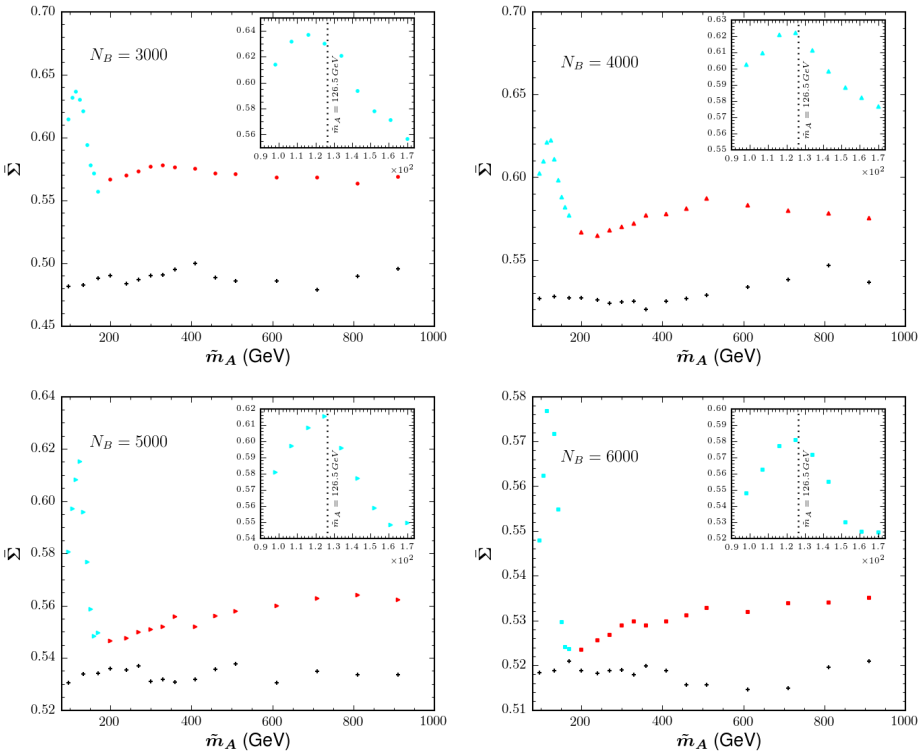


Figure 4.25: The same as Fig. 4.13, but using study point P_{32} as input.

4.4.3 A study with $t\bar{t}$ dilepton background events

Our final task will be to repeat the P_{32} exercise with dilepton $t\bar{t}$ events as was done in Section 4.3.3. As before, we fix the number of signal events $N_S = 3000$ and then consider several values for the number of background events: $N_B = \{3000, 4000, 5000, 6000\}$. In each case, we compute the function $\bar{\Sigma}(\tilde{m}_A)$ along the flat direction of Fig. 4.15. The results are shown in Fig. 4.25, which has the same qualitative behavior as Fig. 4.23. The $\bar{\Sigma}$ values for the auxiliary branch tend to be low, and the branch is disfavored. The global peak of $\bar{\Sigma}(\tilde{m}_A)$ is again found in the vicinity of the right answer (for $N_B = \{3000, 4000, 5000, 6000\}$, the peak is at $\tilde{m}_A = \{116, 125, 125, 125\}$ GeV), and the large $\tilde{m}_A$ tail of the true branch is also disfavored. One final consistency check is provided by Fig. 4.26, which shows a comparison of the kinematic boundaries for the nominal study point P_{32} with $m_A = 126.5$ GeV (red solid lines), and the boundaries for the best fit value $\tilde{m}_A = 125$ GeV (black dashed lines).

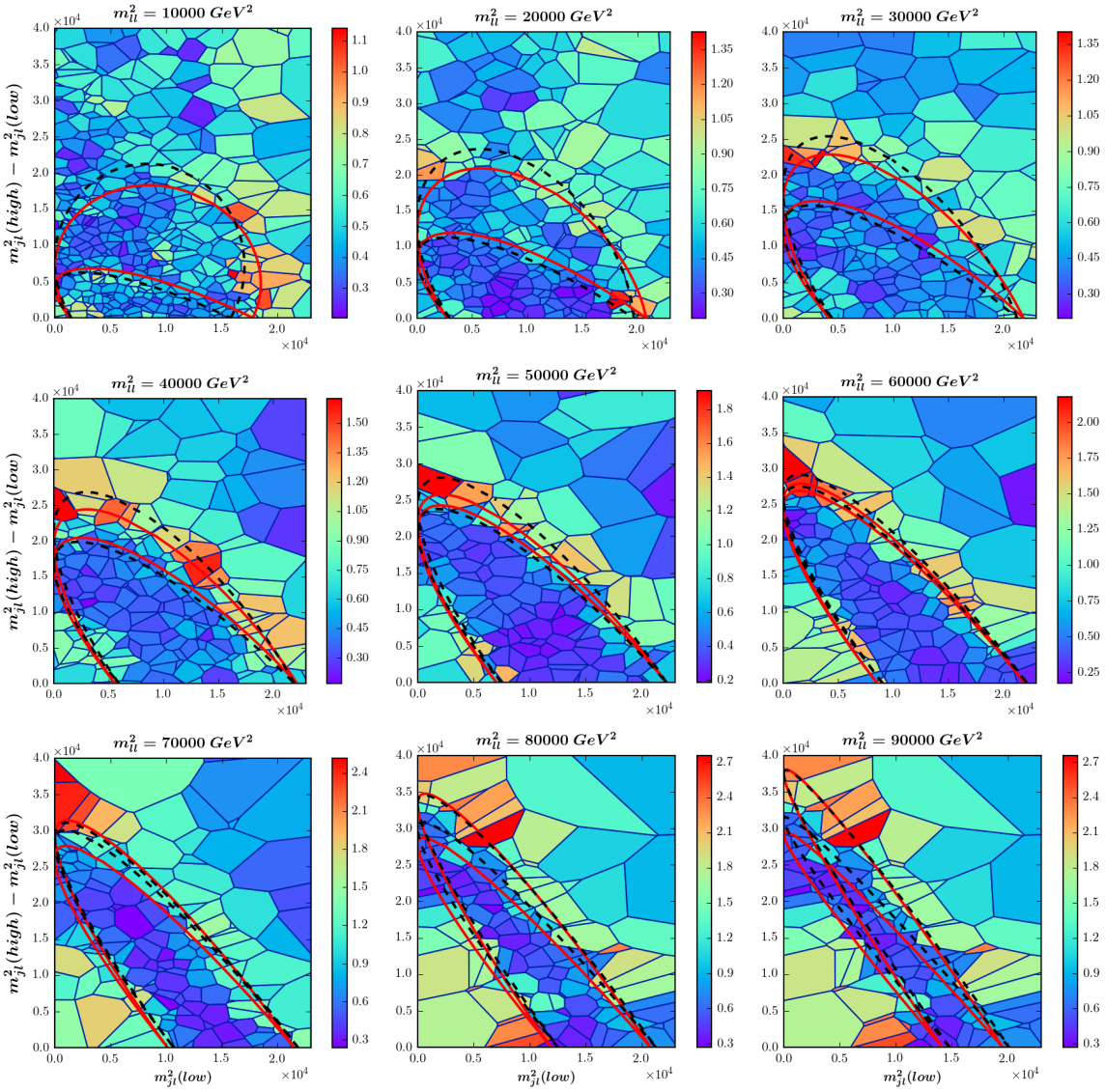


Figure 4.26: The same as Fig. 4.14, but for the exercise performed in Section 4.4.3, using study point P_{32} as input. The red solid line is the phase space boundary for the nominal value $m_A = 126.5$ GeV, while the black dashed line corresponds to the best fit value $\tilde{m}_A = 116$ GeV found in the the top left panel of Fig. 4.25.

4.4.4 A detector level study

In this paper, we introduced the new Voronoi-based method for mass measurement as a proof of principle, and showed that at the parton level it does reasonably well in the two examples considered so far in Sections 4.3 and 4.4. Before concluding, we would

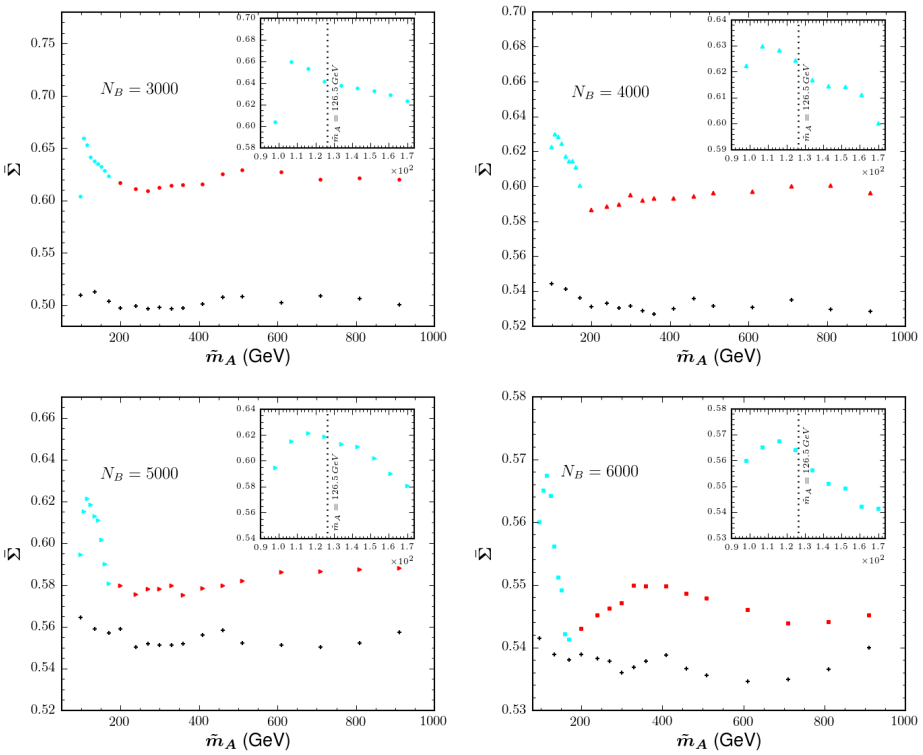


Figure 4.27: The same as Fig. 4.25, but accounting for the detector resolution as described in the text.

like to also test the method in the presence of detector effects (this subsection) and combinatorics (see Sec. 4.4.5 below). For this purpose, we first repeat the exercise from Section 4.4.3, only this time we account for the finite detector resolution by smearing the jet energies with the typical hadronic calorimeter resolution

$$\frac{\sigma}{E} = \left(\frac{1}{\sqrt{E}}\right) \tag{4.40}$$

and electromagnetic calorimeter resolution

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{0.0363}{\sqrt{E}}\right)^2 + \left(\frac{0.124}{E}\right)^2 + 0.0026^2 \tag{4.41}$$

in CMS [146], with the energy measured in GeV. Smearing of the muon momenta is done according to the "Full System" values in Fig. 1.5 of [146]. The result of the fitting exercise is shown in Fig. 4.27. We see that the peak structure in the vicinity of the correct mass value (126.5 GeV) is preserved, but somewhat degraded due to the detector resolution effects.

4.4.5 D-pair production and combinatorics effects

Our proposed mass measurement method uses a single decay chain like the one depicted in Fig. 4.1. In that sense, the method is inclusive and model-independent, since it does not depend on what else is going on in the event. In particular, the method is equally applicable when particle D is produced singly, in pairs, or in association with another object. Nevertheless, a well-motivated and widely studied class of models are the SUSY-like dark matter scenarios in which all particles A, B, C and D carry negative parity under the additional Z_2 symmetry. In that case, D has to be produced in association with another negative parity object. If A is a neutral dark matter candidate, then D must carry color, therefore D pair-production is strong and may dominate the inclusive cross-section for D production.

The presence of a second D decay chain in the event can either be a blessing or a curse. If both D particles decay the same way, as in Fig. 4.1, we can attempt to double our statistics by considering the second decay chain as well. However, this comes at a cost, since now we have to face the combinatorial problem of associating the different reconstructed final state objects to one of the two decay chains. First, there is a two-fold ambiguity of associating each jet to the correct side, and furthermore, there is an additional two-fold ambiguity in the case when all leptons are the same flavor. The simplest approach would be to consider all possible combinations and use the resulting data set for building the Voronoi tessellation, then proceeding with the fitting of the boundary surface as before. The result from this exercise is shown in Fig. 4.28 for the case of 100 (left panel) and 500 (right panel) signal events. We see that despite the pollution from wrong combinatorics, the peak in $\bar{\Sigma}$ is still very well visible, and found in the right location.

Our final study is reserved for the most challenging example so far — a case with a severe combinatorics problem, in the presence of SM ($t\bar{t}$) background events. For signal, let us again consider D-pair production, only this time let *all three* of the decay products of the second D particle be QCD jets. Thus, each one of our signal events has 2 leptons and 4 jets, and picking the correct jet becomes a difficult task. Now, instead of using all possible combinations, we design a preselection cut in order to improve our chances of capturing the correct jet pairing. For this purpose, we consider the four possible jet-lepton-lepton combinations and compute the corresponding three-body invariant masses. Next, we rank-order these four values [99, 128] and eliminate

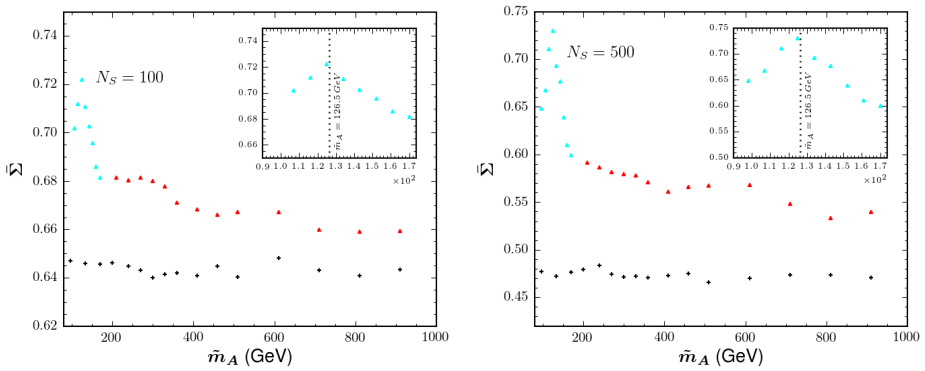


Figure 4.28: The same as Fig. 4.25, but for signal events where particles D are pair-produced and decay as in Fig. 4.1. The left (right) panel is made with 100 (500) signal events.

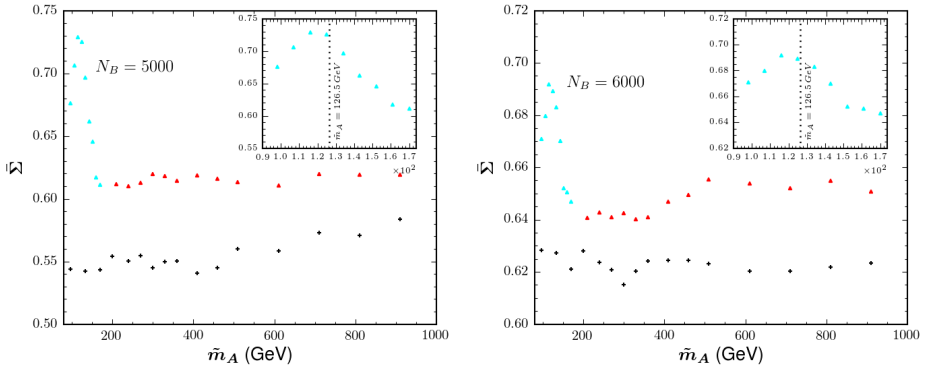


Figure 4.29: The same as the lower two plots in Fig. 4.25, but for signal events where particles D are pair-produced and one of them decays as in Fig. 4.1, while the other decays to 3 jets and particle A.

from further consideration the two jets which correspond to the two largest jet-lepton-lepton invariant masses, since those jets are very likely to come from the decay chain opposite the two leptons. The remaining two jets still cause a two-fold ambiguity, which we handle as in Fig. 4.28: by simply plotting both combinations. The end result from the analysis is shown in Fig. 4.29. As in the example from Sec. 4.4.3, here we also include a certain number of $t\bar{t}$ background events: 5000 in the left panel and 6000 in the right panel. The fact that the $\bar{\Sigma}$ peak is again obtained in the correct location indicates that our method can be viable in the presence of combinatorial background due to pair production of particles D.

4.5 Conclusions

In this paper, we reconsidered the classic endpoint method for particle mass determination in SUSY-like decay chains like the one shown in Fig. 4.1. Our main points are:

- We have identified a "flat direction" in the mass parameter space (4.1), along which mass differences can be measured relatively well, but the overall mass scale remains poorly constrained. (The analytical formulas parametrizing this flat direction can be found in Appendix B.) We quantified the problem with examples of specific study points, P_{31} and P_{32}, considered in Sections 4.3 and 4.4, respectively.

- We then proposed a new method for mass measurements in general, and for extracting the mass scale along the flat direction, in particular. The method takes advantage of the changes in the *shape* of the two-dimensional kinematic boundary surface within the fully differential three-dimensional space of observables, as one moves along the flat direction. We have tested our Voronoi-based algorithm [134, 17] for detecting the boundary surface and demonstrated that it can be usefully applied in order to lift the degeneracy along the flat direction. This approach represents the natural extension of the one-dimensional kinematic endpoint method to the relevant three dimensions of invariant mass phase space.

- We introduced a new variable, $\bar{\Sigma}$, which is the average RSD per unit area, calculated over the hypothesized kinematic boundary. We showed that the location of the $\bar{\Sigma}$ maximum correlates very well with the true values of the new particle masses, see Figs. 4.11, 4.13, 4.23, 4.25, and 4.27.

The work reported here can be extended in several directions. First of all, the method can be readily generalized to longer decay chains with more visible particles, where the boundary enhancement is even more pronounced [135], and therefore, the detection of the boundary surface should be in principle easier. One could also try to apply Voronoi-based boundary detection algorithms for the *discovery* of new physics. It is also interesting to develop a general and universal method for estimating the statistical significance of the local peaks found in Figs. 4.11, 4.13, 4.23 and 4.25, and

hence the statistical precision of our mass measurement. These, along with many other interesting questions, will be investigated in future studies [110].

Chapter Five: Enhancing the discovery prospects for SUSY-like decays with a forgotten kinematic variable[1]

5.1 Introduction

The possible existence of particles beyond the Standard Model (SM) at the TeV scale is theoretically motivated both by naturalness considerations for the electroweak scale [147], and by the so-called WIMP (weakly interacting massive particle) miracle for obtaining the correct dark matter relic abundance [148]. Nevertheless, as we approach the end of Run II of the Large Hadron Collider (LHC), we have as yet no conclusive evidence of new particles beyond the SM (BSM) [149]. This requires us to pause, rethink and perhaps re-optimize our search strategies, in preparation for what may lie ahead. In particular, we should be mindful of the following challenges:

- *The signal may be buried under a large SM background.* Of course, one obvious possibility for why partner particles may so far have evaded detection is that they are simply too heavy and therefore have small production cross sections. If that is the case, then discovery could be waiting around the corner, provided that the signatures of the new particles are distinctive. For instance, significant mass gaps in the spectrum of the new particles will result in high p_T leptons and jets in the final state and a sizable missing transverse energy, E_T. Therefore, while the signal cross section may be low, signal over background can still be large and reaching discovery sensitivity will simply be a question of collecting sufficient statistics. This scenario is rather uninteresting to us, and instead in this paper we focus on the alternative — that the new particles are being produced in sizable numbers, but their signatures are

[1]Previously published as D. Debnath, J. S. Gainer, C. Kilic, D. Kim, K. T. Matchev, Y.-P. Yang,"Enhancing the discovery prospects for SUSY-like decays with a forgotten kinematic variable" on on Journal of High Energy Physics 05(2019)08. Y.-P. Yang derived the analytic distribution of $\Delta_4/\Delta_{4,\max}$, conducted Monte Carlo simulation and mass spectrum scan, and analyzed the results with receiver operating characteristic curve

plagued by large SM backgrounds, so the name of the game is whether we can identify selection criteria which have the best potential for discriminating against the background. This attitude is supported by the flurry of theoretical activity in recent years in designing models which "hide" the new physics from the LHC. One of the standard methods for doing so is to arrange for a "compressed" mass spectrum with a mass degeneracy of the relevant particles, such as supersymmetric (SUSY) partners, so that the resulting decay products are too soft to be triggered upon and tagged in the experimental analysis [150, 151, 152, 153, 154, 155, 156, 157, 158, 159], or a "stealth" mass spectrum, where the new physics signature becomes identical to the SM background, since the additional particles are too soft to make any appreciable difference [160, 161, 95, 162, 163, 164, 165]. Our aim will be to highlight a kinematic variable that, either by itself or in conjunction with more conventional variables, can more effectively select signal over background when the signal spectrum is compressed and when signal events contain multi-stage cascade decays.

- *Exclusive searches may be reducing the signal statistics to unobservable levels.* When searching for new physics, one has to find the right balance between inclusive and exclusive searches. Inclusive searches are more robust since they have fewer theoretical assumptions about the event topology and have a higher signal efficiency. On the flip side, they tend to suffer from larger SM backgrounds. In contrast, exclusive searches have the potential to reach higher sensitivity when the correct assumptions are made about the features of signal events, since those features can then be used to reduce backgrounds, but at the cost of relying on the assumptions about event topology that may prove to be incorrect.

  In our study we will remain much more inclusive than in experimental searches that model the topology of the *e*ntire event, and instead we will only operate on the assumption that the event contains (at least) one SUSY-like cascade decay proceeding through a sequence of two-body decays and with an invisible particle at the end of the decay chain. We will make no assumptions about whether the particle at the beginning of the cascade is singly or pair-produced, and if the latter, what the "other side" of the event looks like. Because of this, we will not make direct use of$\not\!\!E_T$, or any other transverse variables. Adopting

a benchmark final state with three visible and one invisible final state particles [see Fig. 5.1(d)], we will focus our attention on fully Lorentz-invariant kinematic variables.

- *Uncertainties in background modeling.* A required component of any new physics search is the prediction of the expected SM background. Depending on the final state, this may turn out to be a difficult task, plagued by large systematics. Ideally one would like to use data-driven background estimates, and not rely on theoretical input or Monte Carlo. The classic technique for such searches is the "bump hunting" method with sideband subtraction. Fig. 5.1(a-c) shows examples of simpler decay chains for which this method is easily applied. Fig. 5.1(a) depicts a visibly decaying resonance, here to two visible particles v_1 and v_2. In this case, the relevant kinematic variable is the invariant mass $m_{v_1 v_2}$ of the decay products — it exhibits a Breit-Wigner peak at the mass m_{X_1} of the new resonance. Since the $m_{v_1 v_2}$ distribution for the SM background is expected to be smooth, one can interpolate from the sidebands and obtain a reliable prediction for the background under the peak. This tried-and-true method has been used successfully many times in the past, including most recently for the discovery of the Higgs boson in the diphoton channel [166, 167].

  However, the method runs into a complication if one of the final state particles is invisible in the detector, e.g. particle χ in Fig. 5.1(b). Nevertheless, the procedure still goes through, only this time one has to use a suitable kinematic variable which retains the "bump" feature for the signal, namely the *transverse* invariant mass m_{T,X_1} [168, 169, 170]. The downside of the transverse mass variable m_T (and the related mass variables m_{T2} [171], m_2 [90, 92, 94], etc.) is that its definition uses theE_T measurement, which forces a departure from inclusivity, and also suffers from the systematics of all possible detector effects. For decay chains containing more than one visible particle, one can remain more inclusive by working only with Lorentz-invariant variables constructed from the momenta of these particles. For the two-stage decay chain in Fig. 5.1(c), the only such kinematic variable is the invariant mass $m_{v_1 v_2}$, whose distribution does have a distinctive feature [172]. While these cases have all been studied in great detail in the past, there has not been a comparable effort to design optimized variables for a longer decay chain, such as in Fig. 5.1(d). We will

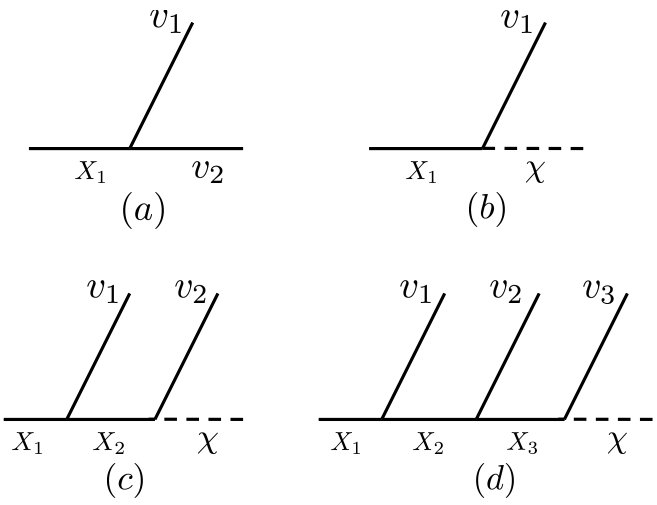


Figure 5.1: Benchmark decay topologies which allow for inclusive searches for the production of a new heavy resonance X_1. Here v_1, v_2 and v_3 are SM particles which are reconstructed in the detector (either directly, or through their respective visible decay products), while χ is a potential dark matter candidate which is invisible in the detector. X_2 and X_3 are additional BSM particles with masses $m_{X_1} > m_{X_2} > m_{X_3} > m_\chi$.

therefore adopt this decay topology as our benchmark in this paper. Our main goal will be to identify and study a kinematic variable for this decay topology that is robust to a certain amount of uncertainty in the modeling of the relevant backgrounds.

Based on the arguments above, an obvious choice of kinematic variables to consider are the pair-wise[2] invariant masses of the visible decay products, $m_{v_1v_2}$, $m_{v_2v_3}$, and $m_{v_1v_3}$, or some combination of those. For plotting convenience, in what follows we shall actually use the *squares* of those variables and denote them as

$$m_{12}^2 \equiv m_{v_1v_2}^2, \qquad m_{23}^2 \equiv m_{v_2v_3}^2, \qquad m_{13}^2 \equiv m_{v_1v_3}^2. \tag{5.1}$$

[2]The invariant mass variable $m_{v_1v_2v_3}$ of all three visible particles is not an independent quantity, since

$$m_{v_1v_2v_3}^2 = m_{v_1v_2}^2 + m_{v_2v_3}^2 + m_{v_1v_3}^2 - m_{v_1}^2 - m_{v_2}^2 - m_{v_3}^2.$$

The variables (5.1) are in principle good candidates for the analysis, not only because they are Lorentz invariant, but also because their distributions exhibit interesting kinematic features (edges and endpoints) which are traditionally used for determining the masses of the new particles X_1, X_2, X_3 and χ [22, 23, 173, 24, 26, 27, 28, 29, 36, 37].

However, as discussed in refs. [98, 174, 1, 99, 175, 135], the multidimensional phase space $\{m_{12}^2, m_{23}^2, m_{13}^2\}$ in this case in fact contains more information than is captured by edge-and-endpoint variables alone. As we will be describing in more detail in section 5.2, the vicinity of the endpoints corresponds only to a fraction of the full boundary of the kinematically available phase space. This boundary is defined via the condition[3] $\Delta_4 = 0$ where the variable Δ_4 will be introduced and defined in section 5.2 below. For now we simply remark that the location of this boundary contains the complete information about the spectrum in the cascade decay [98, 174]. A determination of this boundary (using Voronoi tessellations [17, 134]) has already been shown to result in an improvement in the measurement of the new physics mass spectrum [175].[4] More importantly, the phase space volume element has an enhancement near the boundary, even in the case of a compressed spectrum [1]. This suggests that Δ_4 may be an effective discovery variable, especially in difficult scenarios of compressed spectra. The main goal of this paper will be to investigate the suitability of the Δ_4 variable as an analysis variable, either on its own, or when paired with the edge-and-endpoint variables[5].

In order to demonstrate the basic idea, we adopt a specific realization of our benchmark decay topology from Fig. 5.1(d), by specifying a final state on which we will base our analysis (see Fig. 5.2). In particular, we will take X_1 and X_3 to be charged particles, while X_2 and χ are neutral. We also take the neutral particles to be flavor singlets. The SM particles produced in the second and third stages of the cascade are therefore oppositely charged, and have the same flavor, whereas the charge and flavor assignments of the SM particle produced in the first stage of the cascade are uncorrelated with the other two. Furthermore, in order to concentrate on what can be achieved using phase space techniques for discovery, we will aim to minimize possible complications due to challenging collider objects, so we choose

[3] Alternative parametrizations of the kinematic boundary can be found in [98, 174, 99].

[4] For a related qualitative discussion, see page 573 in [174].

[5] Note that Δ_4 is only defined for the phase space of four or more final state particles, and therefore cannot be used for the topologies in Fig. 5.1(a)-(c).

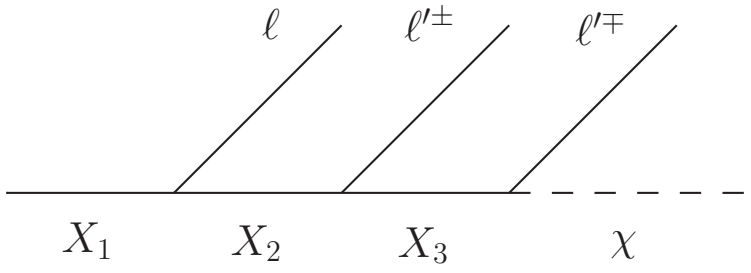


Figure 5.2: The specific realization of the event topology from Fig. 5.1(d) which will be studied in this paper. Here $\ell'^{\pm}$ and $\ell'^{\mp}$ is a pair of opposite-sign, same-flavor leptons, while ℓ is a third lepton of a different flavor.

the visible particles to be leptons. It is worth reiterating that our choice of final state is simply a choice of convenience in order to demonstrate the applicability of our methods, but the methods can be applied to photons, jets or even unstable SM particles with fully visible decays (such as visibly decaying Z-bosons) as well, at the potential cost of worse detector energy resolution and combinatorics. Our analysis *w*ill take into account the effect of finite energy resolution for leptons, as well as the combinatoric ambiguity about which lepton is emitted at the various decay stages. In particular, there will not in general be a way to distinguish which of the same-flavor, opposite-charge leptons is emitted higher upstream in the cascade. On the other hand, the lepton emitted in the first stage of the cascade *c*an be distinguished by demanding it to carry a flavor different from the same-flavor, opposite-charge lepton pair.

Since we aim to focus on improving signal selection in the case of compressed spectra, we adopt the following benchmark spectrum: $m_{X_1} = 390$ GeV, $m_{X_2} = 360$ GeV, $m_{X_3} = 330$ GeV and $m_\chi = 300$ GeV. Note that the choice of spectrum is mainly intended to demonstrate how well the kinematic variables in question compare to one another. Our conclusions would not be affected by raising all masses in the spectrum (while preserving the mass gaps), if we wanted to assign additional significance to this mass benchmark and avoid existing exclusion constraints for various potential underlying models, such as supersymmetry.

The outline of this paper is as follows: In the next section we will review the theoretical aspects of multidimensional phase space and formally introduce the Δ_4 variable. In section 5.3, we will then perform a preliminary study with simplified

assumptions to outline the salient features of Δ_4 as a discovery variable. In section 5.4 we will address a subtlety about the use of a hypothesis spectrum in order to calculate Δ_4. Once this is done, we will then perform a realistic study of the performance of Δ_4 as a discovery variable in section 5.5. We conclude in section 5.6.

5.2 Mathematical description of four-body phase space

Let us start by introducing a *m*anifestly Lorentz-invariant parametrization of the phase space for the cascade decay of our benchmark decay topology. Using the formalism of ref. [6],[6] we introduce the matrix

$$\mathcal{Z} = \{z_{ij}\} \quad with \quad z_{ij} = p_i \cdot p_j \,, \tag{5.2}$$

where the $\{p_i\}$ are the four momenta of the final state particles ℓ, $\ell'^{\pm}$, and χ. The variables Δ_i can then be defined as

$$\det\left[\lambda I_{4\times 4} - \mathcal{Z}\right] \equiv \lambda^4 - \left(\sum_{i=1}^{4} \Delta_i \lambda^{4-i}\right). \tag{5.3}$$

Among these variables, Δ_4 will play a special role in the rest of this paper. As described in ref. [6], the kinematically allowed region is given by $\Delta_{1,2,3,4} > 0$, with the boundary located at[7]

$$\Delta_4 = 0, \qquad \Delta_{1,2,3} > 0\,. \tag{5.4}$$

With the requirement that all $m_{ij}^2 \geq 0$, outside of the kinematically allowed region the values of Δ_4 are negative and become arbitrarily large in magnitude as one moves towards infinity.

The general four-body phase space volume element is given by

$$d\Pi_4 = \left(\prod_{i<j} dm_{ij}^2\right) \frac{8}{(4\pi)^{10} M_{X_1}^2 \Delta_4^{1/2}} \, \delta\left(\sum_{i<j} m_{ij}^2 - \left(M_{X_1}^2 + 2\sum_{i=1}^{4} m_i^2\right)\right), \tag{5.5}$$

[6]For an alternative derivation, the curious reader is invited to follow Exercise 11 on page 574 in [174].

[7]Alternative equivalent parametrizations of this kinematic boundary were previously derived in [98, 174, 99]. However, those results were not used to study the *interior* of the kinematically allowed phase space, as we will be doing here.

where $m_{ij}^2 = (p_i + p_j)^2 = 2z_{ij} + m_i^2 + m_j^2$ [8]. Note the factor of $\Delta_4^{-1/2}$, which causes an enhancement near the boundary $\Delta_4 = 0$.

Of course, the physically observable quantities depend not only on $d\Pi_4$ but on $|\mathcal{M}|^2$, the quantum mechanical matrix element squared for the decay:

$$d\Gamma = |\mathcal{M}|^2 \, d\Pi_4 \, . \tag{5.6}$$

In particular, for the benchmark decay topology of Fig. 5.2, the volume element will be combined with the squares of the internal propagators in the cascade, which in the narrow width approximation are given as delta functions with arguments linear in the m_{ij}^2 and can therefore be used to perform some of the m_{ij}^2 integrals. As a result, the events fill out a three-dimensional phase space that can conveniently be fully parameterized in terms of the observables m_{12}^2, m_{13}^2 and m_{23}^2.

The enhancement in the phase space volume element near the boundary should make it clear why it is promising to consider Δ_4 as a discovery variable. The prominent features in the edge-and-endpoint variable distributions happen at the extremes of linear slicings of the three dimensional phase space, and therefore only a small fraction of signal events contribute to these features. In contrast, the prominent feature in the Δ_4 distribution at $\Delta_4 = 0$ captures the full boundary of phase space, where the density of signal events is enhanced, so it is reasonable to expect that selecting for events near $\Delta_4 = 0$, one could significantly enhance signal over background.

It is worth remarking that the phase space for any known SM background process does not develop a singular structure like the one described in eq. (5.5). Furthermore, there is no reason to expect the $|\mathcal{M}|^2$ factor for the background to have any sharp features over the kinematically accessible signal region (the location of which depends on the signal spectrum). In particular, for a compressed signal spectrum which results in a relatively small signal region, the variation of the background matrix element over this region will in all likelihood be mild.

Note that for a given event, Δ_4 cannot be calculated from the observable data alone. As can be seen from eq. (5.3), Δ_4 is equal to $-det[\mathcal{Z}]$, and the last column and row of $\mathcal{Z}$ contain the four momentum of the lightest supersymmetric particle (LSP) χ, which is unobservable. However, if one starts with a hypothesis for the spectrum $\{m_{X_1}, m_{X_2}, m_{X_3}, m_\chi\}$, the on-shell constraints allow one to solve for all

[8] This is the general formula. For our analysis, while $m_\chi > 0$, we will take the leptons to be massless.

entries of $\mathcal{Z}$, and thus a mass hypothesis dependent value of Δ_4 can be calculated. The obvious question to ask then is whether this requirement for a spectrum hypothesis significantly weakens the usefulness of the Δ_4 variable. We will take up this question in section 5.4, drawing the conclusion that Δ_4 is a powerful variable despite this caveat.

5.3 Preliminary study with uniform background

In order to illustrate the usefulness of Δ_4, we wish to compare its performance as a discovery variable to the conventional edge-and-endpoint variables using the benchmark cascade decay and spectrum specified in the introduction. The performance of all variables will depend on the differential distribution of signal and background events, which as mentioned in the previous section will in turn depend on both the geometry of phase space as well as the matrix elements for signal and background. Again as emphasized in the previous section, the usefulness of Δ_4 originates from the phase space geometry for signal, in particular, the enhancement of the signal event density near the boundary of the kinematically allowed region where there is no strong reason to expect a feature in the density of background events. Therefore, we devote this section to a toy study where we minimize the effects of the matrix elements and of the background event distribution, by taking all particles in the signal decay chain to be scalars, and we make the highly simplifying approximation that the background varies not only slowly over the signal region but is in fact uniformly distributed over phase space (parameterized in terms of the coordinates m_{ij}^2). We will also use the true signal spectrum in calculating Δ_4 and return to the issue of having to scan over spectrum hypotheses in the next section, before we do a full analysis with SM backgrounds and a signal model with spins of new particles assigned SUSY-like in section 5.5.

Since we use a uniformly distributed background, we need to define a finite box in the three-dimensional space formed by the three m_{ij}^2 variables in order to deal with only a finite number of background events. We choose the box size as twice the maximal possible signal value in each of the m_{ij}^2 variables. This choice ensures that finite energy resolution in the detector does not push signal events outside the box, and that no artificial features are introduced in background distributions at small

but negative values of Δ_4, close to but outside the signal region. We generate high statistics samples with one million signal and background events each, where in the signal the flavors of the leptons ℓ and ℓ' are randomly assigned as electrons or muons. We only consider events where those two flavors are distinct.

Even in this preliminary study, we will need to face two complications. One is finite energy resolution, as mentioned, while the other complication arises from combinatoric ambiguities. Note that in our benchmark topology of Fig. 5.2, it cannot be experimentally determined in which order the particles ℓ'^+ and ℓ'^- are emitted in the cascade, leading to a combinatoric ambiguity. As argued in ref. [24], in such a case it is advantageous to work with *o*rdered variables instead, so we define and work with the variables

$$m^2_{1(hi)} \equiv \mathrm{m}ax(m^2_{12}, m^2_{13}), \qquad m^2_{1(lo)} \equiv \mathrm{m}in(m^2_{12}, m^2_{13}). \tag{5.7}$$

Note that there is no combinatorial ambiguity in defining m^2_{23} as we require ℓ' and ℓ to have distinct flavors. Due to the combinatorial ambiguity, there are two possible values of Δ_4 for every event, and both of them will be used when populating Δ_4 histograms. In setting up our study, we will choose to start by using perfect energy resolution and by ignoring the combinatoric ambiguity, before introducing them below. We do this because there are a few important lessons we can learn even before the analysis is made more complicated by these effects.

As mentioned in the introduction, an ideal discovery variable that eliminates the need for precise background modeling would exhibit a strong feature in the distribution of the signal while the background distribution is smooth at the same position, such that a sideband analysis can pick out the signal as in a bump-hunting analysis. At first sight, Δ_4 seems to be a promising variable along these lines, since the signal event density is enhanced near $\Delta_4 = 0$ while the background event density has no reason to be enhanced at the same surface, the location of which after all is dependent on the signal spectrum. Unfortunately, this line of thinking misses a potential problem, namely that even though the *d*ensity of background events may be smooth near the surface $\Delta_4 = 0$, the phase space in which signal and background events are distributed is three-dimensional, and in making a one-dimensional histogram of Δ_4, one has to integrate the phase space volume between surfaces of constant Δ_4. This can still introduce a feature into the background Δ_4 histogram if the volume between contours itself exhibits a feature near $\Delta_4 = 0$. This does in fact happen to be the case, since

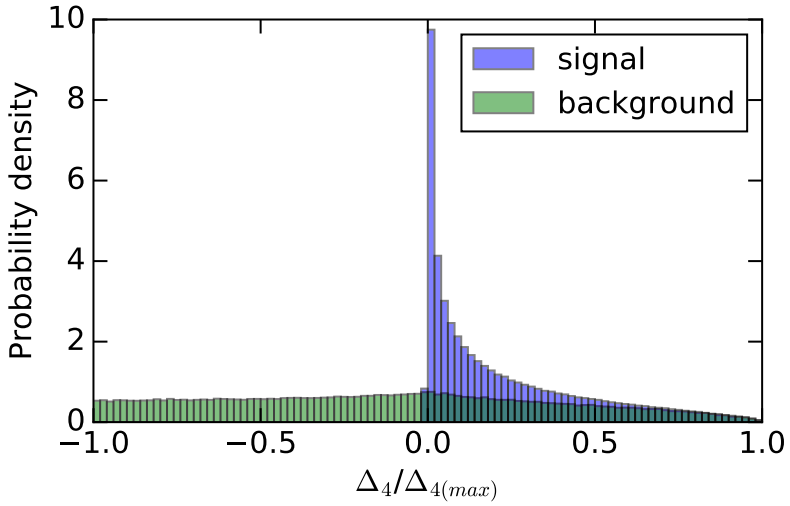


Figure 5.3: The Δ_4 histograms for signal (blue) and (uniformly distributed) background (green). The distributions are normalized by the maximum Δ_4 value for the chosen mass spectrum, $(m_{X_1}, m_{X_2}, m_{X_3}, m_\chi) = (390, 360, 330, 300)$ GeV. The feature in the background distribution near $\Delta_4 = 0$ is caused by the volume between constant Δ_4 surfaces becoming maximal.

the gradient of Δ_4 is small on a significant portion of the boundary surface, increasing the volume between Δ_4 contours there. The resulting Δ_4 histogram for signal and background (uniform density) is shown in Fig. 5.3, where the normalization of the signal and background histograms has been chosen such that they both contain the same total number of events. Here Δ_4 values are normalized by the maximum Δ_4 for the chosen mass spectrum, $(m_{X_1}, m_{X_2}, m_{X_3}, m_\chi) = (390, 360, 330, 300)$ GeV. When the number of background events are significantly higher than the number of signal events, as is often the case for searches for new physics, and when the distributions become smeared due to finite energy resolution, the presence of the background feature at $\Delta_4 = 0$ will make a simple bump hunt based on a sideband analysis difficult, since the signal can be misinterpreted as a background systematic [174].

We therefore switch to a different approach for a search strategy. In order to compare the effectiveness of the different variables in selecting signal events, we construct a performance curve of each variable as follows[9]. For a given variable, a histogram is

[9]The spirit of these curves is similar to a receiving operator characteristic (ROC) curve, even

made of the signal and background. For the m^2 variables, the interval of interest in the histogram is between the maximum and minimum possible values predicted by the spectrum, and for Δ_4 it is the interval between $\pm\Delta_{4(max)}$, also as predicted by the spectrum. The interval of interest is divided into 100 bins[10]. The first entry in the performance curve is the ratio of signal to background events (S/B) in the bin with the highest number of signal events. To obtain the second entry in the performance curve, this bin is combined with the bin to its left or to its right, whichever of the two has the larger number of signal events, and S/B is calculated for the combined two-bin region. For the third entry in the performance curve, these two bins are combined with the neighboring bin with the higher number of signal events, and so on. The procedure stops when all bins containing signal events are exhausted, and therefore the last entry in the performance curve corresponds to S/B over the full signal region for the variable in question. Note that the ordering of the bins in terms of signal events (as opposed to S/B) reduces the reliance on background modeling.

We point out that the performance curves of any two variables may be meaningfully compared independently of the overall signal and background normalizations, since any change in the signal and background normalizations will multiply the performance curve of all variables by the same common factor. Using the same procedure, for completeness we also produce performance curves for the $S/\sqrt{B}$ metric[11]. These performance curves are shown in Fig. 5.4. Note that by construction, the background has a flat distribution in all m_{ij}^2 variables, and in the absence of spin correlations, the signal has an exactly flat distribution in m_{12}^2 and m_{23}^2, and a nearly flat distribution in m_{13}^2 as well. This explains the near-flatness of the S/B performance curves of the m_{ij}^2 variables, as well as the $\sqrt{N_{bins}}$ scaling for the $S/\sqrt{B}$ performance curves. As can be seen from the figures, Δ_4 performs significantly better than these with respect to both metrics.

Encouraged by this result, we proceed to check whether it is robust in the presence of finite detector energy resolution and combinatorial ambiguities. We use the EM

though they are not technically ROC curves.

[10]We verify that the procedure outlined here is not sensitive to the choice of binning.

[11]S/B and $S/\sqrt{B}$ are the relevant quantities measuring signal significance in searches that are systematics and statistics dominated, respectively, and we wish to remain agnostic as to which case may apply in the experimental search of interest.

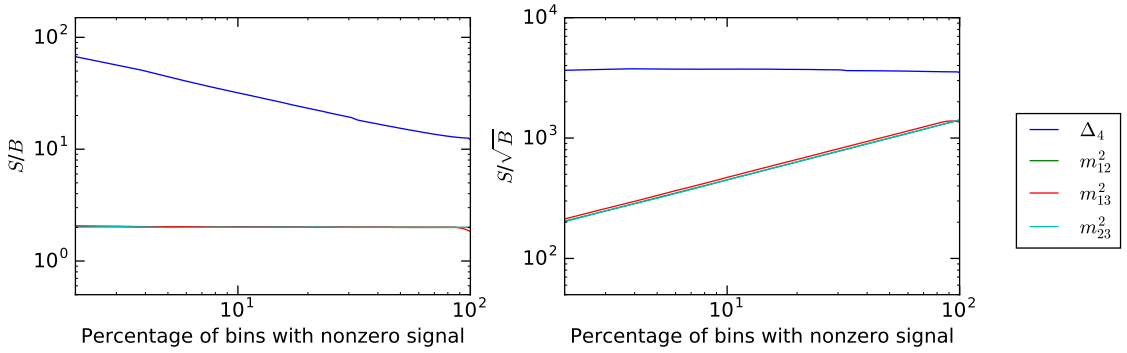


Figure 5.4: Performance curves for Δ_4 and the invariant mass variables using the S/B (left panel) and S/$\sqrt{\text{B}}$ (right panel) metrics, with perfect energy resolution. See the main text for the way in which we construct these curves.

calorimeter resolution based on the CMS-TDR [176]

$$\frac{\sigma_E}{E} = (0.0026) \oplus \frac{0.0363 \text{ GeV}^{1/2}}{\sqrt{E}} \oplus \frac{0.124 \text{ GeV}}{E}, \tag{5.8}$$

where the energy E is defined in GeV. For the muon resolution we utilized values (in terms of muon momentum and pseudorapidity) summarized in Figure 1.5 of the CMS-TDR [176]. Since the background that we consider in this preliminary study is not physical and has no four-vectors associated with it, we leave it unmodified. To incorporate combinatorial ambiguities into the analysis, we use the ordered m^2 variables as defined in eq. (5.7), and we populate Δ_4 histograms by both possible values for each event as mentioned above. The effect of smearing and combinatorics on the Δ_4 distribution of figure 5.3 is shown in Fig. 5.5.

As a result of both smearing and combinatorics, the performance curves for Δ_4 in Fig. 5.4 are mildly degraded, which can be seen in Fig. 5.6. In Fig. 5.7, the performance curves of Δ_4 and the edge-and-endpoint variables are compared with energy resolution and combinatorics included. Δ_4 is seen to still outperform the edge-and-endpoint variables, but by a smaller margin.

After this preliminary comparison among single kinematic variables as discovery tools, it is also interesting to look at how well *p*airs of variables compare to one another. In particular we will be interested in whether pairing Δ_4 with the m^2 variables will be more effective than pairing one of the m^2 variables with another one. The procedure we use to perform this comparison closely mirrors the procedure

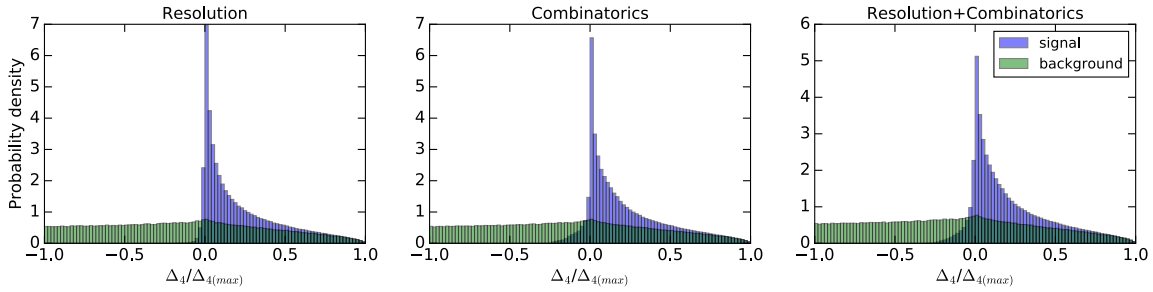


Figure 5.5: The Δ_4 histograms for signal (blue) and (uniformly distributed) background (green), with energy resolution and combinatoric ambiguities included. To be compared to Fig. 5.3

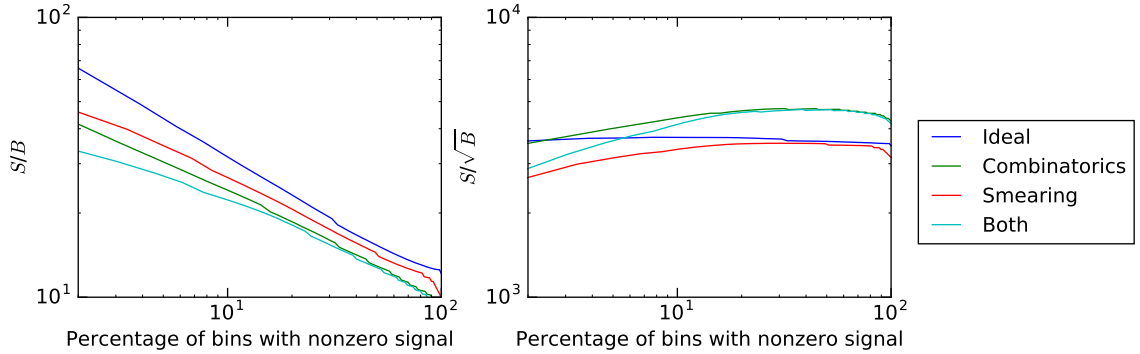


Figure 5.6: The effect of energy resolution and combinatorics on the significance performance curve of Δ_4 is shown using the S/B (left panel) and S/$\sqrt{\text{B}}$ (right panel) metrics.

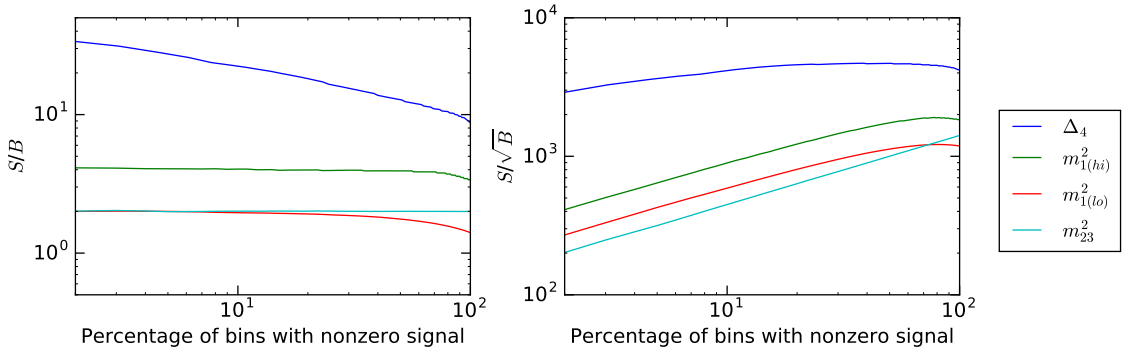


Figure 5.7: The same as Fig. 5.4, but taking the finite energy resolution and combinatoric effects into account.

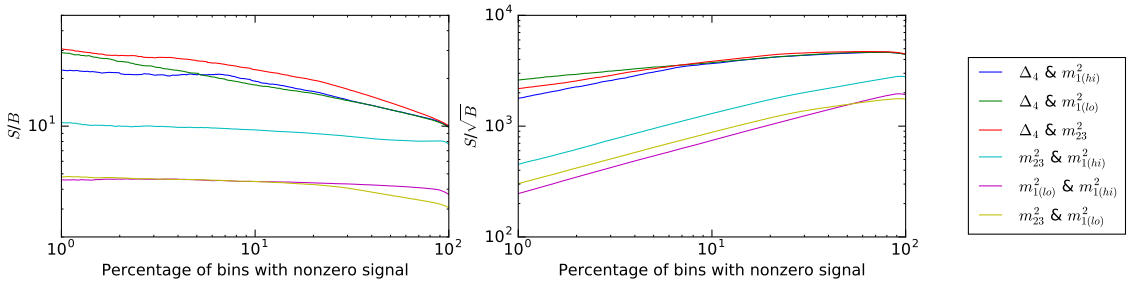


Figure 5.8: Performance curves for *pairs of* variables among Δ_4 and the invariant mass variables, using the S/B (left panel) and S/$\sqrt{\text{B}}$ (right panel) metrics, taking finite energy resolution and combinatoric effects into account.

outlined above for the case of a single variable. In particular, for any pair of variables, signal and background events populate a double histogram in the two variables in question (the same binning parameters are used in each variable as described earlier in this section). The (double) bins are then ordered in order of their signal contribution, but *w*ithout demanding that the bins that are combined neighbor one another, and performance curves of S/B and of S/$\sqrt{\text{B}}$ are made. The effects of both smearing and of combinatorics are included. We exhibit the results in Fig. 5.8 from which it is easy to see that variable pairs including Δ_4 perform better than variable pairs not including Δ_4 with respect to both metrics.

5.4 Scanning over mass spectra

Encouraged by the promising results of our preliminary study described in the previous section, we will devote this section to address the issue of the spectrum dependence in calculating Δ_4. In particular, since the true signal spectrum is not known a priori, analyses involving Δ_4 will need to scan over all possible (correctly ordered) spectra. Below, we will show that the significance is maximized at least locally for the true spectrum, a result which is consistent with the conclusions of ref. [175]. Therefore, if one were to scan over all spectra and use the spectrum that yields the highest significance, then the performance curve based on the true spectrum offers a guaranteed, and in fact potentially conservative (should other spectra exist far from the true spectrum that lead to even higher significance), benchmark for comparison

against the performance curves of the m^2 variables. The significances we report will be *l*ocal. The calculation of a global significance requires the use of a trials factor which is tricky to define for this analysis and is beyond the scope of this paper.

The question of the potential existence of other local (or even global) maxima of significance requires extensive calculational resources, since a fine scan over four masses is required[12]. However, since we will show below that the true spectrum yields at least a local maximum, with a high significance value, then if other local maxima with even higher significance should exist, this would only strengthen the discovery potential, not reduce it, but at the cost of having to give up the claim that the spectrum can be simultaneously measured in the same analysis. We will therefore not make this claim in this study.

To demonstrate that the true spectrum yields a local maximum of significance, we will compare the performance curves of Δ_4 for a range of hypothesized spectra obtained by *l*ocal deformations around the true spectrum. A background uniform in the m_{ij}^2 variables is used as in the previous section, and finite energy resolution as well as combinatorial ambiguities are included in the analysis.

For the local scan near the true spectrum, we allow each of the four masses to change up or down by 10 GeV, resulting in 8 variations. The performance curves obtained as a result of the scan are shown on the left-hand side of Fig. 5.9. It is easy to see that for any low or moderate number of bins in the performance curve, the true spectrum yields the highest significance. The strong reduction in the performance as one goes away from the true spectrum (along any direction other than the flat direction, see the next paragraph) can be traced to the fact that the sharp peak at $\Delta_4 = 0$ is only present when Δ_4 is calculated for the true spectrum, and is severely distorted otherwise, thereby erasing the most distinctive feature in the signal distribution compared to the background distribution.

We also perform a finer one-dimensional scan along a special direction. In particular, while the m_{ij}^2 variables are sensitive to changes in the mass gaps in the spectrum, there is a direction where the endpoints of all three m_{ij}^2 distributions remain fixed. We parameterize this direction in terms of the change in the mass of the LSP from its benchmark value. As shown in ref. [175], Δ_4 is sensitive to changes along the flat direction, while the effect on the shape of the m_{ij}^2 distributions is minimal. These

[12]We expect such resources to be available to the LHC collaborations, however most the analysis in this paper is performed entirely on standalone computers.

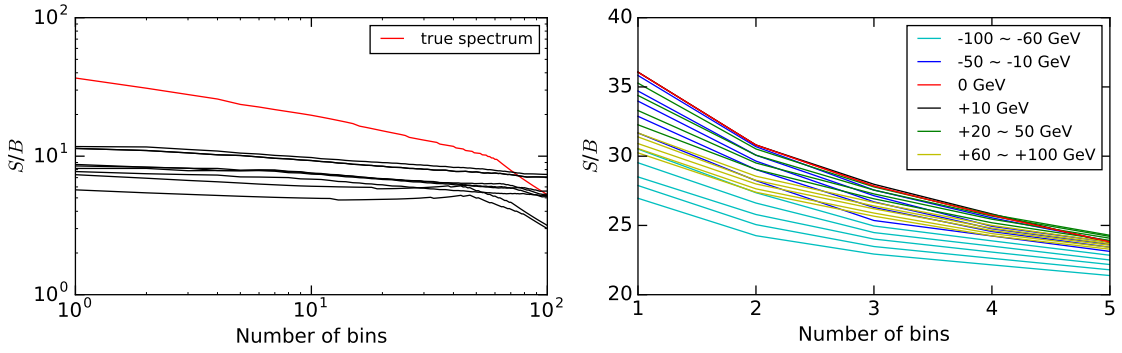


Figure 5.9: Performance curves for Δ_4 calculated by using a range of hypothesis spectra and the S/B metric. Left: Each one of the plotted curves corresponds to deforming the spectrum by changing each of the four masses up or down by $\pm$ 10 GeV. For comparison, the red curve highlights the true spectrum. Right: Each one of the curves corresponds to deforming the spectrum along the flat direction described in the main text over a wide range. The color scheme corresponds to the change in the mass of the LSP.

results are shown in the right-hand side of Fig. 5.9, with the conclusion that small deformations along the flat direction leave the performance curve unchanged (within statistical errors) while more substantial deformations reduce the significance. The results of the scans presented above thus confirm our claim that the Δ_4 performance has a local maximum for the true spectrum.

5.5 Study with SM background

Having obtained encouraging results in our toy study with uniform background, and having dealt with the subtlety of scanning over spectrum hypotheses in calculating Δ_4, we are now in the position to conduct a much more realistic study, with SM backgrounds, matrix element effects in the signal, finite detector resolution, and combinatorics taken into consideration. For the signal, we consider a benchmark model where X_1 is a scalar muon partner, X_2 is a heavy fermion, X_3 is a scalar electron partner, and χ is the fermionic LSP. It should be emphasized again that we are not arguing for this as a signal model to be taken literally; as argued in the introduction, this model is chosen to make an apples-to-apples comparison between Δ_4 and the

m^2 variables possible, without introducing distracting complications. Nevertheless, we believe that our proposed analysis is straightforwardly applicable to the SUSY signal searches in the channel we study here. This signal model guarantees the flavor arrangement of the three leptons in our benchmark cascade. The dominant SM background for this final state is $WZ^{(*)}$ production followed by their leptonic decays. Since our benchmark spectrum ensures that the opposite sign, same flavor lepton pair invariant mass remains well below m_Z, we impose a Z-veto in simulating the background, so that the region with off-shell Z's can be scanned efficiently.

We perform our parton-level simulation for signal and background using `MG5@aMC` [140], and apply energy resolution for final state leptons according to the CMS-TDR [176] [see also eq. (5.8)]. We use the following selection cuts on the events:

$$p_{T,\ell} > 10 \text{ GeV}, \quad |\eta_\ell| < 2.5, \quad \Delta R_{\ell\ell} \geq 0.4, \quad 15\text{GeV} < m_{\ell^+\ell^-} < 65 \text{ GeV} \quad (\ell = e,\, \mu). \tag{5.9}$$

Here the invariant mass cut in the second line is relevant only to same-flavor opposite-sign lepton pairs.

For the generated signal and background event samples, we plot the Δ_4 distributions, as well as the effect of smearing and combinatorics on these distributions, in Fig. 5.10. The resulting performance curves for Δ_4 are obtained following the same steps as in section 5.3, and shown in Fig. 5.11. We then compare the performance of Δ_4 to the edge-and-endpoint variables in Fig. 5.12. We observe that the Δ_4 variable becomes less powerful than it was in our preliminary study with uniform background. The main reason for this degradation is because the matrix elements and the parton distribution functions that govern the phase space distribution of SM background events lead more events to lie close to the regions in which Δ_4 is smaller than that for the uniform background distribution [174]; for example, the event population in the same-flavor lepton pair invariant mass is enhanced at small values due to the mixing between γ and Z, resulting in more background population at small values of Δ_4. Nonetheless, Δ_4 shows a comparable performance to the strongest m^2 variable with respect to both metrics.

Furthermore, as we pointed out in our preliminary exercise, some m^2 variable, when combined with Δ_4, may outperform traditional approaches with m^2 variables only. Indeed, the same expectation goes through for the signal under consideration, which is supported by the results presented in Fig. 5.13. As one would expect based

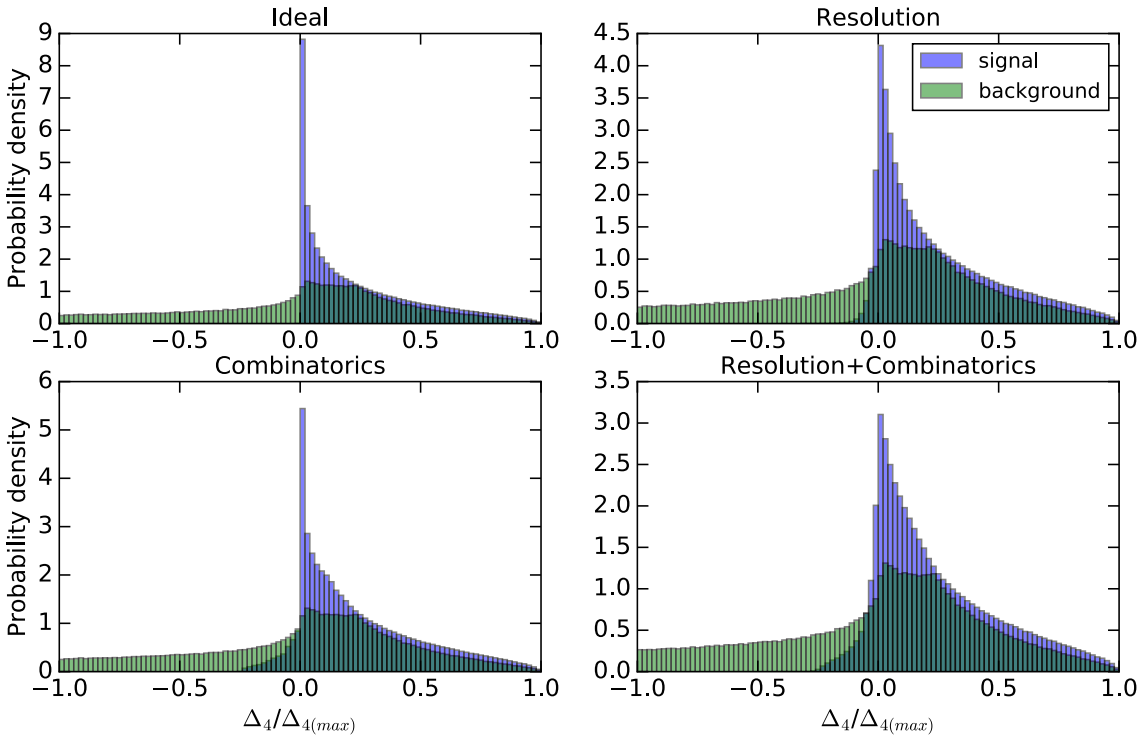


Figure 5.10: The Δ_4 histograms for signal (blue) and the SM background (green), with energy resolution and combinatoric ambiguities included.

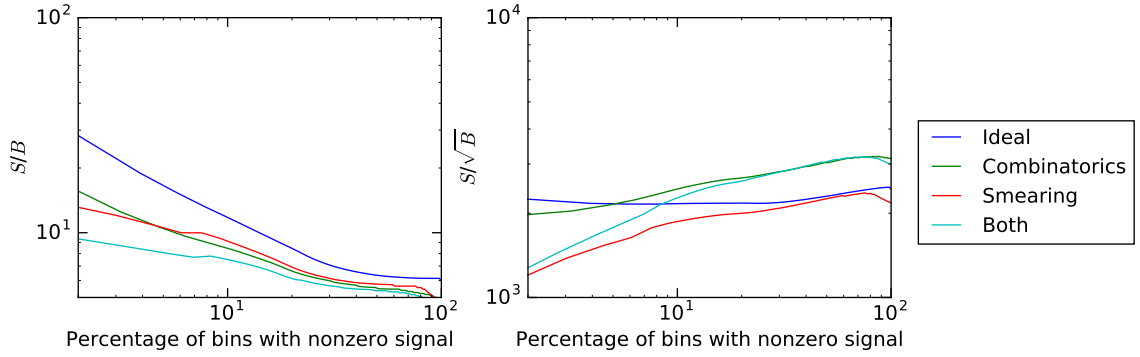


Figure 5.11: The effect of energy resolution and combinatorics on the significance performance curve of Δ_4 is shown using the S/B (left panel) and $S/\sqrt{B}$ (right panel) metrics.

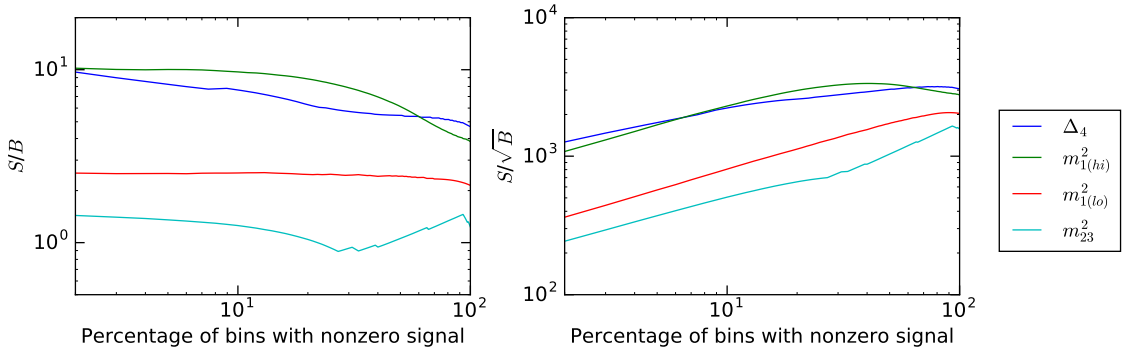


Figure 5.12: Performance curves for Δ_4 and the m^2 variables using the S/B and S/$\sqrt{\text{B}}$ metrics.

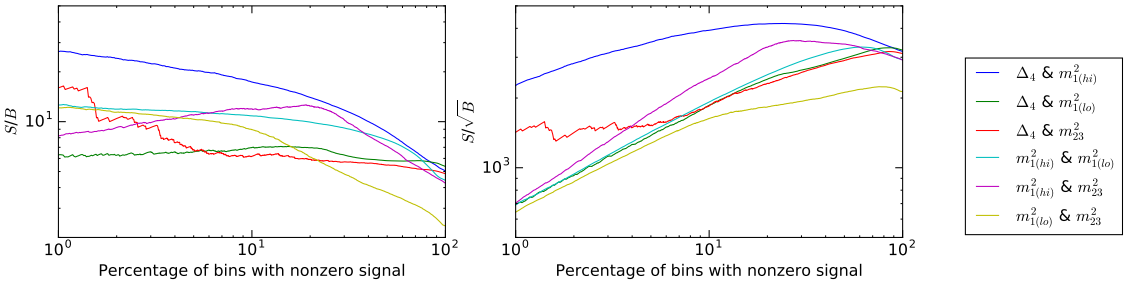


Figure 5.13: Performance curves for pairs of variables among Δ_4 and the m^2 variables, using the S/B and S/$\sqrt{\text{B}}$ metrics.

on the single variable results of Fig. 5.12, the best performance is achieved by the combination between $m^2_{1(hi)}$ and Δ_4 (blue lines) in both the S/B (left panel) and the S/$\sqrt{\text{B}}$ (right panel) metrics. Therefore, we find that Δ_4 can play, at least, a *c*omplementary role in separating signal from background, hence expediting a discovery of new physics.

5.6 Conclusions

As we approach the end of Run II in the LHC experiment, the absence of a discovery of new physics makes it increasingly more imperative to focus on scenarios where a new physics signal may exist in the data, but not be distinctive enough to register in searches looking for high momentum particles. This happens for example when

the new particles that are produced decay in a cascade with a compressed spectrum. We argued that using the variable Δ_4, which arises naturally in describing four-body phase space, allows one to design a search strategy in such a scenario that is quite inclusive and does not rely strongly on background modeling.[13] We do this by focusing our attention on only the part of the event containing the cascade decay, using Lorentz-invariant variables, and by not using detailed properties of the background in designing our search strategy. We have argued that even though the calculation of Δ_4 requires a hypothesis for the mass spectrum in the cascade decay, the significance has a local maximum for the true signal spectrum which can be used as a benchmark of comparison against the performance of other variables. We have compared the performance of the variable Δ_4, both singly and paired with conventional edge-and-endpoint variables, in a study using SM backgrounds, spin correlations, finite energy resolution and combinatoric effects, concluding that Δ_4 can significantly enhance the signal both for systematics-dominated (S/B metric) and statistics-dominated ($S/\sqrt{B}$ metric) searches.

[13] In place of Δ_4, one could in principle also use the geometrical distance to the kinematical boundary (5.4), a possibility which was entertained in [174]. However, that choice has disadvantages: the geometrical distance is suboptimal in terms of performance and cannot be easily computed by analytical means.

Appendix A: Curves and Sensitivity Measures

ROC (receiver operating characteristic) curves are a useful tool for measuring the sensitivity of a variable[1]. As shown in Fig. A.1, the ROC curve is found by (i) defining an event selection procedure ("cut") based on the variable in question and (ii) determining the fraction, ε_S, of signal events, and the fraction, ε_B, of background events that pass the given cut. The coordinates of each point along the curve are then provided by (ϵ_B, ϵ_S). It is easy to see that the ROC curve must include the point $(0, 0)$, "A" in Figs. A.1 and A.2, where all events, signal and background, have been disallowed by the cut. The point $(1, 1)$, where all events pass the cut ("C" in Figs. A.1 and A.2), is also part of every ROC curve. ROC curves have a number of important and useful properties which we shall explore in the remainder of this section.

A.1 Comparing ROC curves

If we consider two ROC curves, $R_1(\varepsilon_B)$ and $R_2(\varepsilon_B)$, obtained for the same signal and background processes using different choices of variable and/or the cut procedure, then if

$$R_1(\varepsilon_B) \geq R_2(\varepsilon_B) \text{ for all } \varepsilon_B, \tag{A.1}$$

the variable/ cut combination used to produce R_1 is clearly more sensitive than the variable/cut combination used to produce R_2. This statement is uncontroversial, but the comparison is not always applicable, as a pair of ROC curves, R_1 and R_2 may have points $\varepsilon_{B,1} \neq \varepsilon_{B,2}$ such that

$$R_1(\varepsilon_1) > R_2(\varepsilon_1), \tag{A.2}$$

but

$$R_1(\varepsilon_2) < R_2(\varepsilon_2). \tag{A.3}$$

[1]Much of the information in this section can be found elsewhere [116], however we present a unified and self-contained exposition of the main facts about ROC curves here for the convenience of the reader.

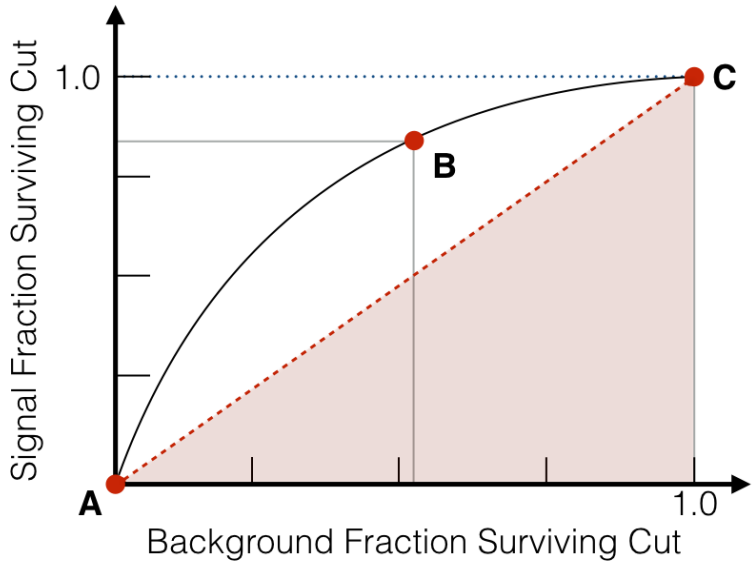


Figure A.1: Schematic illustrating the construction of the ROC curve. Each point along the curve indicates the fractions of signal and background events that pass a parameterizable cut. If the cut parameter is chosen to disallow all signal and all background events, we are at the point labelled "A", while if the cut parameter is chosen to allow signal and background events, then the appropriate point is "C". An intermediate point, where a certain fraction of signal and background events are allowed is labelled by "B".

We will therefore need to develop other procedures to compare ROC curves; we will present several approaches in section A.4. First, however, we must investigate the connection between ROC curves and likelihood and explore some of the more important consequences of this relationship.

A.2 ROC curves and likelihood

There are many ways to perform an analysis using a given variable, and hence many ROC curves that may be constructed with no other information than the value of a given variable for signal and background events. We note that the ratio of signal and background likelihoods is an optimal test statistic, i.e., choosing to accept an event, e, when

$$\frac{L_S(e)}{L_B(e)} \geq l(\varepsilon_B), \tag{A.4}$$

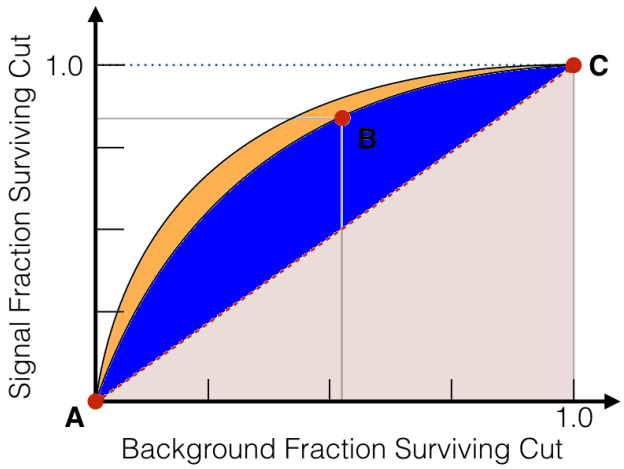


Figure A.2: Schematic illustrating the connection between the Gini coefficient (A.25) and sensitivity. Points "A", "B", and "C" from Fig. A.1 together with the ROC curve from Fig. A.1 are reproduced here. We have also indicated the ROC curve of a perfectly insensitive cut, which is shown by the line from point "A" to point "C". Finally we have shown a ROC curve for a cut which gives a higher signal fraction for every choice of background fraction than the ROC curve from Fig. A.1. We see that the area under the ROC curve of the perfectly insensitive variable (shown in pink) is 1/2, that there is additional area under the ROC curve shown in Fig. A.1 (shown in blue), and that there is even an even greater area under the ROC curve for the more sensitive variable.

where L_S is the signal likelihood and L_B is the background likelihood, is the procedure which accepts the maximum fraction of signal events (ε_S) for a given choice of ε_B (which in turn determines the numerical value of $l(\varepsilon_B)$ in eq. (A.4)). This fact is known as the Neyman-Pearson lemma [117] and is equivalent to the statement that the likelihood ratio produces the optimal ROC curve.

We now provide a heuristic proof of this assertion in terms of ROC curves, which hopefully provides useful insights into both ROC curves and the Neyman-Pearson lemma. We consider the situation where signal and background likelihoods are calculated numerically from samples of signal and background events, S and B, which can be assumed arbitrarily large in order to approximate analytic expressions with any desired accuracy. First, we create a set of "variable bins" in our variable of interest,

i.e.,

$$V_i = [v_{min,i}, v_{max,i}), \tag{A.5}$$

such that for every event, e, in our signal and background samples,

$$v(e) \in [v_{min,i}, v_{max,i}) \tag{A.6}$$

for exactly one choice of i, where $v(e)$ refers to the value of the variable under consideration obtained from the event, e. We then divide our signal and background events into bins using the values of the variable, v, under consideration. I.e., we obtain "signal bins"

$$S_i = \{e \mid v(e) \in V_i\} \text{ for } e \in S \tag{A.7}$$

and "background bins"

$$B_i = \{e \mid v(e) \in V_i\} \text{ for } e \in B. \tag{A.8}$$

Since every signal and every background event must be in some bin, by construction, we have

$$\sum_i |S_i| = |S| \tag{A.9}$$

and

$$\sum_i |B_i| = |B|. \tag{A.10}$$

Hence the estimators of the signal and background likelihood,

$$L_{S,i} = L_S(v_i) \approx |S_i|/|S| \text{ for } v_i \in V_i \tag{A.11}$$

and

$$L_{B,i} = L_B(v_i) \approx |B_i|/|B| \text{ for } v_i \in V_i, \tag{A.12}$$

are automatically normalized. Further, we note that the contribution to the ROC curve from bin i is a line segment from initial point $(\varepsilon_{B,0},\ \varepsilon_{S,0})$ to $(\varepsilon_{B,0}+|B_i|/|B|,\ \varepsilon_{S,0}+|S_i|/|S|)$, i.e. from $(\varepsilon_{B,0},\ \varepsilon_{S,0})$ to $(\varepsilon_{B,0} + L_{B,i},\ \varepsilon_{S,0} + L_{S,i})$. We therefore see from eqs. (A.11) and (A.12) that the slope of the line segment corresponding to a particular bin is given by the ratio of signal and background likelihoods calculated for that bin. An interesting corollary, since $L_{S,i}$ and $L_{B,i}$ are always non-negative, is that the ROC curve must be monotonically increasing, i.e.,

$$\frac{d\varepsilon_S}{d\varepsilon_B} \geq 0. \tag{A.13}$$

We now assert that any physically reasonable cut procedure[2] accepts a set of variable bins

$$V_A = v \in V_i \text{ for } i \in \{i_1, i_2, ...\}, \tag{A.14}$$

signal events

$$S_A = e \in S_i \text{ for } i \in \{i_1, i_2, ...\}, \tag{A.15}$$

and background events

$$B_A = e \in B_i \text{ for } i \in \{i_1, i_2, ...\}, \tag{A.16}$$

where $\{i_1, i_2, ...\}$ are the indices of the variable bins which pass the cut. Further if we have a cut that can be parameterized and that ranges from a cut in which all events are accepted to a cut in which no events are accepted, then we can express the cut function as a permutation, p of $1, 2, ..., n$, where n is the number of bins, such that the variable bin labelled by $p(1)$ is the last to be eliminated, the variable bin eliminated by $p(2)$ is the next-to-last to be eliminated, etc.

The optimal cut function identified by the Neyman-Pearson lemma, i.e., eq. (A.4) corresponds to the case where $p(1)$ labels a bin with the maximum value of signal to background likelihood ratio and

$$\frac{L_S(v_{p(i)})}{L_B(v_{p(i)})} \geq \frac{L_S(v_{p(j)})}{L_B(v_{p(j)})} \tag{A.17}$$

if $i > j$. We claim that this gives the ROC curve for which $\varepsilon_S(\varepsilon_B)$ is maximized for any choice of ε_B. This is geometrically obvious given that the likelihood ratio gives the slope of the line segment in the ROC curve; the optimal ROC curve is one in which the steepest segments occur first. A more rigourous proof notes that the ROC curve based on the likelihood ratio has

$$\varepsilon_S = \sum_i^m \frac{L_{S,p(i)}}{L_{B,p(i)}} L_{B,i} + f_{m+1} \frac{L_{S,p(m+1)}}{L_{B,p(m+1)}} L_{B,(m+1)}, \tag{A.18}$$

where

$$\varepsilon_B = \sum_i^m L_{B,i} + f_{m+1} L_{B,(m+1)}. \tag{A.19}$$

[2]We may have to choose different, and, in particular, smaller bins, e.g., we may not be able to model the cut function $v > v_0$ if v_0 is not on a bin boundary. In principle this limitaiton can be circumvented by choosing the variable bins to be smaller than the detector resolution, which must be finite. Since physically reasonable cut functions can only use the output of such a detector, we can eliminate pathological cut functions, like only accepting events with rational v.

and $f_{m+1} \in [0,1]$ is the fraction of bin $m+1$ that must be included to give the desired value of ε_B.

We now consider the value of ε_S at the same value of ε_B in a different ROC curve. Since this ROC curve differs from the optimal ROC curve described by eqs. (A.18-A.19), it must include some different bins or different fractions of bins. These "new bins" give a contribution to ε_B of $\Delta\varepsilon_B$, which is also the contribution to ε_B from the "replcaed original bins" they replace. The contribution to ε_S from the replaced original bins was $\Delta\varepsilon_S$, and can be found by multiplying $\Delta\varepsilon_B$ by the (appropriately weighted) average slope of the replaced original bins. The contribution to ε_S from the new bins is also the weighed average slope of the bins. However every bin in the new bins has a slope less than or equal to the slope of any bin in the replaced original bins. Thus the value of ε_S in the new ROC curve is less than or equal to its value in the ROC curve based on the likelihood ratio.

A.3 Subdividing variable bins

We now consider the effect on our ROC curve of subdividing a bin, by which we mean that we take a bin, V_i, and break it into two bins, which we label V_i and V_{n+1} as follows:

$$V_i = [v_{i,min}, v_{i,max}) \to [v_{i,min}, v_{i',max}) \text{ and } [v_{i',max}, v_{i,max}) = V_{i'} \text{ and } V_{i''}. \tag{A.20}$$

We then obtain the signal bins S_i and S_{n+1} following eq. (A.7) and the background bins B_i and B_{n+1} following eq. (A.8). Using these new bins will always allow for the construction of a ROC curve that is as good as or better than the curve before subdivision (in the sense of eq. (A.1)). This is clear as if the permutation of bins that led to the original ROC curve is p, then the permutation

$$p'(j) = p(j) \text{ for } j < i \tag{A.21}$$

$$p'(i) = i \text{ (which now refers to a fraction of the original bin, } i) \tag{A.22}$$

$$p'(i+1) = i' = n+1 \tag{A.23}$$

$$p'(j) = p(j-1) \text{ for } j > i+1 \tag{A.24}$$

yields a ROC curve which is the same everywhere except in the region corresponding to the subdivided bin. If the two "daughter" bins have the same slope (likelihood

ratio), then this curve will be exactly the same as the original ROC curve. However, if this is not the case, then if we choose (without loss of generality) the daughter bin with the greater slope to come first (i.e., to be labelled as i), we have a ROC curve which has strictly greater signal values than the original ROC curve in the regime corresponding to the subdivided bin. Finally, it is possible that daughter bin i may have a greater slope than some of the other bins, that daughter bin $n+1$ may have a lesser slope than some of the other bins, or both of these possibilities may be realized. In all of these cases the ROC curve may be further imporoved by sorting the bins by slope.

As a consequence of this result considering smaller bins will increase sensitivity; the extreme limit is to use an unbinned likelihood rather than binned likelihood. While we have not proved it, this fact also motivates the (true) statement that considering additional relevant, but still independent, variables will improve sensitivity (as subdividing a bin using the new variable may yield daughter bins with different slopes).

However, it should be noted, when using actual samples of signal and background events to construct likelihoods, that one can reach false conclusions about sensitivity by subdividing bins too far. An extreme example of this is provided by the case where all the signal and background events we are considering have their own bin. In this case, the likelihood ratio is infinite for bins with one signal event and zero for bins with one background event and the ROC curve is vertical from $(0,0)$ to $(0,1)$, and then horizontal from $(0,1)$ to $(1,1)$. Obviously this is too good to be true. What has happened is an example of posterior statistics and is clearly an invalid inference. Clearly such a situation should be avoided, e.g., by using bins large enough to be robust with respect to statistical fluctuations.

A.4 Measures of ROC curve sensitivity

In section A.1, we noted that sometimes a ROC curve will indicate that one variable/ cut procedure is absolutely better than another variable/ cut procedure. However, this may not always be the case. Thus it is important to have procedures for comparing ROC curves (and hence sensitivity) that always allow the comparison to be made. All of these procedures can be viewed as functionals, i.e., they describe a ROC

curve by a number that corresponds directly to the quality of a variable.

Simple functionals that can be used are (i) the value of ε_S at a fixed value of ε_B and (ii) the value of ε_B at a fixed value of ε_S. We can view (i) as the fraction of signal events chosen with a fixed false positive rate, while in (ii) we demand that the cut accept a fixed fraction of signal events. A better variable will then reject a larger fraction of background events.

In particle physics, we are generally interested in the significance with which we can claim the discovery (or the exclusion) of some process. Therefore we can describe ROC curves by (i) the maximum value of $S/\sqrt{B}$ attained on the ROC curve, (ii) the maximum value of S/B attained on the ROC curve, or (iii) the maximum discovery significance on the ROC curve. Choice (i) is appropriate as a measure of significance in the (Gaussian) limit of a large number of events in situations where systematic uncertainties do not have a large effect. Choice (ii) is appropriate in the situation where systematic errors dominate. There are different appropaches to implementing choice (iii) to model the statistical and systematic components of the significance. The difference between the Gaussian and Poisson distributions for small numbers of events may also be modelled. While choice (iii) represents the actual quantity to be maximized for a specifc analysis, it has the drawbacks of (i) being complicated and (ii) being dependent on specific experimental conditions, such as the luminosity gathered, systematic uncertainties, etc.

To make comparisons between ROC curves in a general way, we note that the worst possible likelihood-based ROC curve is the ROC curve that reflects throwing away signal and background events indiscriminately, i.e. a straight line from ("A" to "C" in Figs. A.1 and A.2). On the other hand, a perfect variable would take us from "A" to $(0, 1)$ to "C". This ROC curve contains the entire range of ε_B and ε_S underneath it, suggesting the use of the integral of the ROC curve as a sensitivity measure. Further, as indicated in Fig. A.2, variables which lead to "better" ROC curves, in the sense of eq. (A.1) in section A.1 have more area under their ROC curves. In order for the worst possible likelihood ROC curve to have a value of 0 and the best possible ROC curve to have a value of 1, we multiply the integral under the ROC curve by 2 and subtract 1. The resulting expression,

$$G_1 \equiv 2\,\mathrm{A}UROC - 1 = 2\int_0^1 d\varepsilon_B \times \varepsilon_S(\varepsilon_B) - 1, \tag{A.25}$$

where AUROC referes to the "area under the ROC curve", e.g., under curve ABC in

Figs. A.1 and A.2 gives the measure of sensitivity known as the Gini coefficient; it is our main quantifier of ROC curves (and hence of variable/ cut procedure sensitivity) in the bulk of this work.

Appendix B: Inverse formulas

In this appendix we derive the inverse relations which define m_B, m_C and m_D in terms of the three measured endpoints

$$a \equiv (m_{ll}^{max})^2\,, \qquad b \equiv \left(m_{j\ell\ell}^{max}\right)^2\,, \qquad c \equiv \left(m_{jl(low)}^{max}\right)^2\,, \tag{B.1}$$

and the remaining mass parameter m_A. For simplicity of notation, in this appendix we shall omit the tildes on the trial mass parameters m_A, m_B, m_C and m_D.

B.1 The case of region $(3,1)$

Region $(3,1)$ is defined by the following conditions

$$R_{AB} \;\leq\; R_{BC}R_{CD} = R_{BD}, \tag{B.2}$$

$$R_{BC} \;\geq\; \frac{1}{2-R_{AB}}. \tag{B.3}$$

The kinematic endpoints are given by the following formulas:

$$a \;=\; m_D^2\, R_{CD}\,(1-R_{AB})(1-R_{BC}), \tag{B.4}$$

$$b \;=\; m_D^2\,(1-R_{BD})(1-R_{AB}), \tag{B.5}$$

$$c \;=\; m_D^2\,(1-R_{BC})(1-R_{CD}), \tag{B.6}$$

$$d \;=\; m_D^2\,(1-R_{CD})(1-R_{AB}). \tag{B.7}$$

The masses of B, C and D are given by

$$m_B^2 \;=\; \frac{1}{2c}\left\{2m_A^2 c + a(b-a-c) + \left[4m_A^2 ac(b-a) + a^2(b-a-c)^2\right]^{1/2}\right\}, \tag{B.8}$$

$$m_C^2 \;=\; m_B^2\left(1+\frac{a}{m_B^2-m_A^2}\right), \tag{B.9}$$

$$m_D^2 \;=\; m_B^2\left(1+\frac{b}{m_B^2-m_A^2}\right), \tag{B.10}$$

where in the right hand sides of the last two equations m_A is an input, while m_B is calculated from (B.8). Using (B.2) it is easy to show that in this region, we always have

$$b - a - c = m_D^2 \, (1 - R_{CD})(R_{BC} - R_{AB}) \geq 0,$$

so that (B.8) always gives a non-negative result for m_B^2. Substituting (B.8-B.10) into (B.7), one can explicitly check that the m_A dependence drops out and we recover the "bad" relation (4.29) in the form

$$d = b - a. \tag{B.11}$$

B.2 The case of region $(3, 2)$

Region $(3, 2)$ is defined by the following two conditions:

$$R_{AB} \leq R_{BC} R_{CD} = R_{BD}, \tag{B.12}$$

$$R_{BC} \leq \frac{1}{2 - R_{AB}}. \tag{B.13}$$

The kinematic endpoints are given by the following formulas:

$$a = m_C^2 (1 - R_{AB})(1 - R_{BC}), \tag{B.14}$$

$$b = m_D^2 (1 - R_{AB})(1 - R_{BD}), \tag{B.15}$$

$$c = m_D^2 (1 - R_{CD})(1 - R_{AB})(2 - R_{AB})^{-1}, \tag{B.16}$$

$$d = m_D^2 (1 - R_{CD})(1 - R_{AB}). \tag{B.17}$$

The masses of B, C and D are given by [139]

$$m_B^2 = \frac{c m_A^2}{2c - b + a}, \tag{B.18}$$

$$m_C^2 = m_B^2 \left(1 + \frac{a}{m_B^2 - m_A^2}\right), \tag{B.19}$$

$$m_D^2 = m_B^2 \left(1 + \frac{b}{m_B^2 - m_A^2}\right). \tag{B.20}$$

In this region, we always have

$$2c - b + a = m_D^2 \, (1 - R_{CD}) \left(\frac{2}{2 - R_{AB}} - 1\right) \geq 0,$$

so that (B.18) always gives a non-negative result for m_B^2. The "bad" relation (4.29) is again satisfied in this region, so that along the flat direction the endpoint d is again constant and given by (B.11), providing a useful cross-check on the obtained solution (B.18-B.20).

B.3 The case of region $(2,3)$.

Region $(2,3)$ is defined by the following condition:

$$R_{BC} \leq R_{AB}R_{CD}, \tag{B.21}$$

The kinematic endpoints are given by the following formulas:

$$a = m_C^2(1-R_{AB})(1-R_{BC}), \tag{B.22}$$

$$b = m_D^2(1-R_{BC})(1-R_{AB}R_{CD}), \tag{B.23}$$

$$c = m_D^2(1-R_{CD})(1-R_{AB})(2-R_{AB})^{-1}, \tag{B.24}$$

$$d = m_D^2(1-R_{CD})(1-R_{BC}). \tag{B.25}$$

The masses of B, C and D are given by

$$m_B^2 = \frac{2m_A^2(b-a)+a(2c-b+a)+\left[4m_A^2ac(b-a)+a^2(2c-b+a)^2\right]^{1/2}}{2(b-a)} \tag{B.26}$$

$$m_C^2 = m_B^2\left(1+\frac{a}{m_B^2-m_A^2}\right), \tag{B.27}$$

$$m_D^2 = \left(1+\frac{a}{m_B^2-m_A^2}\right)\left[\frac{b}{a}\left(m_B^2-m_A^2\right)+m_A^2\right]. \tag{B.28}$$

Using (B.22), (B.23) and (B.25) it is easy to show that in this region, we always have

$$b-a=d\geq 0,$$

and while the term $(2c-b+a)$ can have either sign, the discriminant (i.e., the term inside the square root in (B.26)) is always larger than $a^2(2c-b+a)^2$, which guarantees a non-negative result for m_B^2. In this region, the relation (4.29) is again satisfied, so that d is again given by (B.11), which can be used to cross-check the result (B.26-B.28).

B.4 The case of region $(4,1)$.

Region $(4,1)$ is defined by the following conditions:

$$R_{AB} \geq R_{BC}R_{CD} = R_{BD}, \tag{B.29}$$

$$R_{BC} \geq \frac{1}{2 - R_{AB}}, \tag{B.30}$$

$$R_{CD} \geq R_{AB}R_{BC} = R_{AC}. \tag{B.31}$$

For the case (4,1) the endpoints are given by the following formulas:

$$a = m_C^2(1 - R_{AB})(1 - R_{BC}), \tag{B.32}$$

$$b = m_D^2(1 - \sqrt{R_{AD}})^2, \tag{B.33}$$

$$c = m_D^2(1 - R_{CD})(1 - R_{BC}), \tag{B.34}$$

$$d = m_D^2(1 - R_{CD})(1 - R_{AB}). \tag{B.35}$$

The masses of B, C and D in terms of a, b, c and m_A are given by

$$m_B^2 = \frac{am_D^2 + cm_A^2 + (m_A^2 - a)(a + c) + a\left[(a + c - m_A^2 - m_D^2)^2 - 4m_A^2m_D^2\right]^{1/2}}{2(a + c)}, \tag{B.36}$$

$$m_C^2 = m_B^2\left(1 + \frac{a}{m_B^2 - m_A^2}\right), \tag{B.37}$$

$$m_D = m_A + \sqrt{b}, \tag{B.38}$$

where in (B.36) m_D should be taken from (B.38), and the obtained result for m_B should be used in (B.37). The d endpoint is given by

$$d = b - a - m_D^2\left(\sqrt{R_{AB}} - \sqrt{R_{BD}}\right)^2 \tag{B.39a}$$

$$= \frac{ac\left\{(b - a - c)\left[b(1 + \sqrt{r}) + 2m_A\sqrt{b}\sqrt{r}\right] + 2(2b - a - c)(m_A\sqrt{b} + m_A^2)\right\}}{(a + c)\left[a(b - a - c)(1 + \sqrt{r}) + 2am_A\sqrt{b} + 2(a + c)m_A^2\right]}, \tag{B.39b}$$

where

$$r \equiv 1 + \frac{4m_A(m_A + \sqrt{b})}{b - a - c}. \tag{B.40}$$

B.5 The case of region $(4,2)$.

Region $(4,2)$ is defined by the following conditions:

$$R_{AB} \geq R_{BC}R_{CD} = R_{BD}, \tag{B.41}$$

$$R_{BC} \geq R_{AB}, \tag{B.42}$$

$$R_{BC} \leq \frac{1}{2 - R_{AB}}, \tag{B.43}$$

$$R_{CD} \geq R_{AB}R_{BC} = R_{AC}. \tag{B.44}$$

The kinematic endpoints are given by the following formulas:

$$a = m_C^2(1 - R_{AB})(1 - R_{BC}), \tag{B.45}$$

$$b = m_D^2(1 - \sqrt{R_{AD}})^2, \tag{B.46}$$

$$c = m_D^2(1 - R_{CD})(1 - R_{AB})(2 - R_{AB})^{-1}, \tag{B.47}$$

$$d = m_D^2(1 - R_{CD})(1 - R_{AB}). \tag{B.48}$$

The masses of B, C, D in terms of a, b, c, m_A are the following

$$m_B^2 = \frac{1}{2}\left[m_D^2 + m_A^2 - a - 2c \pm \sqrt{(m_D^2 + m_A^2 - a - 2c)^2 - 4m_A^2(m_D^2 - c)}\right] \tag{B.49}$$

$$m_C^2 = m_B^2\left(1 + \frac{a}{m_B^2 - m_A^2}\right), \tag{B.50}$$

$$m_D = m_A + \sqrt{b}. \tag{B.51}$$

Here, just as in (B.36-B.38), one should first find m_D from (B.51), then use the result in (B.49) to obtain m_B, which will be needed in (B.50). The d endpoint is then given by

$$d = b - a - m_D^2\left(\sqrt{R_{AB}} - \sqrt{R_{BD}}\right)^2 \tag{B.52a}$$

$$= \frac{c\left[a + 3b - 2c + 6m_A\sqrt{b} + 2m_A^2 - \sqrt{(b - a - 2c + 2m_A\sqrt{b})^2 - 4(a + c)m_A^2}\right]}{2(b - c + 2m_A\sqrt{b} + m_A^2)} \tag{B.52b}$$

B.6 The case of region $(4,3)$.

Region $(4,3)$ is defined by the following conditions:

$$R_{AB} \geq R_{BC}, \tag{B.53}$$

$$R_{BC} \geq R_{AB}R_{CD}, \tag{B.54}$$

$$R_{CD} \geq R_{AB}R_{BC} = R_{AC}. \tag{B.55}$$

For the case (4,3) the endpoints are given by the following formulas:

$$a = m_C^2(1-R_{AB})(1-R_{BC}), \tag{B.56}$$

$$b = m_D^2(1-\sqrt{R_{AD}})^2, \tag{B.57}$$

$$c = m_D^2(1-R_{CD})(1-R_{AB})(2-R_{AB})^{-1}, \tag{B.58}$$

$$d = m_D^2(1-R_{CD})(1-R_{BC}). \tag{B.59}$$

The masses of B, C, D in terms of a, b, c, m_A are the following

$$m_B^2 = \frac{1}{2}\left[m_D^2+m_A^2-a-2c-\sqrt{(m_D^2+m_A^2-a-2c)^2-4m_A^2(m_D^2-c)}\right] \tag{B.60}$$

$$m_C^2 = m_B^2\left(1+\frac{a}{m_B^2-m_A^2}\right), \tag{B.61}$$

$$m_D = m_A+\sqrt{b}, \tag{B.62}$$

where again the masses are calculated in the order m_D, m_B and then m_C. The d endpoint is given by

$$d = b-a-m_D^2\left(\sqrt{R_{BC}}-\sqrt{R_{AB}R_{CD}}\right)^2 \tag{B.63a}$$

$$= \frac{4ac\left[b-a-2c+2m_A\sqrt{b}+m_A^2-\sqrt{(b-a-2c+2m_A\sqrt{b})^2-4(a+c)m_A^2}\right]}{\left(b-2c+2m_A\sqrt{b}-\sqrt{(b-a-2c+2m_A\sqrt{b})^2-4(a+c)m_A^2}\right)^2-a^2} \tag{B.63b}$$

Bibliography

[1] P. Agrawal, C. Kilic, C. White and J. H. Yu, "Improved mass measurement using the boundary of many-body phase space," Phys. Rev. D **8**9, no. 1, 015021 (2014) doi:10.1103/PhysRevD.89.015021 [arXiv:1308.6560 [hep-ph]]. 3, 22, 23, 24, 36, 38, 47, 58, 67, 83, 114

[2] E. Byckling and K. Kahantie, "Particle Kinematics" John Wikey & Sons 5, 6

[3] D. Griffiths, "Introduction to Elementary Particles," Wiley-VCH, 2008.

[4] M. Tanabashi et al. (Particle Data Group), "The Review of Particle Physics (2018)," Phys. Rev. D 98, 030001 (2018). 5

[5] Tim Gershon, "Introduction to Dalitz Plot Analysis," https://warwick.ac.uk/fac/sci/physics/staff/academic/gershon/talks/gershon-BadHonnef.pdf. ix, 7

[6] N. Byers and C. N. Yang, "Physical Regions in Invariant Variables for n Particles and the Phase-Space Volume element", Rev. Mod. Phys. **3**6, 595 (1964). ix, 8

[7] For a basic introduction to Voronoi tessellations, see, e.g., S. Okabe, B. Boots and K. Sugihara, "Spatial Tessellations: Concepts and Applications of Voronoi Diagrams," John Wiley & Sons, 1992. 9, 36, 37, 116

20, 67

[8] L. Schlicht, M. Valcu and B. Kempenaers, "Thiessen polygons as a model for animal territory estimation," Ibis **1**56, no. 1, 215-219 (2014). 20

[9] M. Ramella, W. Boschin, D. Fadda and M. Nonino, "Finding galaxy clusters using voronoi tessellations," Astron. Astrophys. **3**68, 776 (2001) doi:10.1051/0004-6361:20010071 [astro-ph/0101411]. 20

[10] M. Cappellari and Y. Copin, "Adaptive spatial binning of integral-field spectroscopic data using voronoi tessellations," Mon. Not. Roy. Astron. Soc. **3**42, 345 (2003) doi:10.1046/j.1365-8711.2003.06541.x [astro-ph/0302262]. 20

[11] L. Giomi and M. Bowick, "Crystalline Order On Riemannian Manifolds With Variable Gaussian Curvature And Boundary," Phys. Rev. B **7**6, 054106 (2007) [arXiv:cond-mat/0702471]. 20

[12] M. Cacciari, G. P. Salam and G. Soyez, "FastJet User Manual," Eur. Phys. J. C **7**2, 1896 (2012) doi:10.1140/epjc/s10052-012-1896-2 [arXiv:1111.6097 [hep-ph]]. 20

[13] B. Abbott *e*t al. [D0 Collaboration], "Search for new physics in $e\mu X$ data at DØ using Sherlock: A quasi model independent search strategy for new physics," Phys. Rev. D **6**2, 092004 (2000) doi:10.1103/PhysRevD.62.092004 [hep-ex/0006011]. 20

[14] B. Abbott *e*t al. [D0 Collaboration], "A quasi-model-independent search for new high p_T physics at DØ," Phys. Rev. Lett. **8**6, 3712 (2001) doi:10.1103/PhysRevLett.86.3712 [hep-ex/0011071]. 20

[15] T. Aaltonen *e*t al. [CDF Collaboration], "Model-Independent and Quasi-Model-Independent Search for New Physics at CDF," Phys. Rev. D **7**8, 012002 (2008) doi:10.1103/PhysRevD.78.012002 [arXiv:0712.1311 [hep-ex]]. 20

[16] T. Aaltonen *e*t al. [CDF Collaboration], "Global Search for New Physics with 2.0 fb^{-1} at CDF," Phys. Rev. D **7**9, 011101 (2009) doi:10.1103/PhysRevD.79.011101 [arXiv:0809.3781 [hep-ex]]. 20

[17] D. Debnath, J. S. Gainer, D. Kim and K. T. Matchev, "Edge Detecting New Physics the Voronoi Way," Europhys. Lett. **1**14, no. 4, 41001 (2016) doi:10.1209/0295-5075/114/41001 [arXiv:1506.04141 [hep-ph]]. 20, 23, 24, 26, 27, 29, 30, 31, 36, 45, 58, 67, 68, 69, 89, 108, 114

[18] S. P. Martin, "A Supersymmetry primer," Adv. Ser. Direct. High Energy Phys. **2**1, 1 (2010) [Adv. Ser. Direct. High Energy Phys. **1**8, 1 (1998)] [hep-ph/9709356]. 20, 60

[19] T. Appelquist, H. C. Cheng and B. A. Dobrescu, "Bounds on universal extra dimensions," Phys. Rev. D **6**4, 035002 (2001) doi:10.1103/PhysRevD.64.035002 [hep-ph/0012100]. 21, 60

[20] H. C. Cheng and I. Low, "TeV symmetry and the little hierarchy problem," JHEP **0**309, 051 (2003) doi:10.1088/1126-6708/2003/09/051 [hep-ph/0308199]. 21, 60

[21] A. J. Barr and C. G. Lester, "A Review of the Mass Measurement Techniques proposed for the Large Hadron Collider," J. Phys. G **37**, 123001 (2010) doi:10.1088/0954-3899/37/12/123001 [arXiv:1004.2732 [hep-ph]]. 21, 60

[22] I. Hinchliffe, F. E. Paige, M. D. Shapiro, J. Soderqvist and W. Yao, "Precision SUSY measurements at CERN LHC," Phys. Rev. D **55**, 5520 (1997) doi:10.1103/PhysRevD.55.5520 [hep-ph/9610544]. 21, 64, 114

[23] H. Bachacou, I. Hinchliffe and F. E. Paige, "Measurements of masses in SUGRA models at CERN LHC," Phys. Rev. D **62**, 015009 (2000) doi:10.1103/PhysRevD.62.015009 [hep-ph/9907518]. 21, 64, 114

[24] B. C. Allanach, C. G. Lester, M. A. Parker and B. R. Webber, "Measuring sparticle masses in nonuniversal string inspired models at the LHC," JHEP **0**009, 004 (2000) doi:10.1088/1126-6708/2000/09/004 [hep-ph/0007009]. 21, 38, 54, 64, 83, 114, 119

[25] C. G. Lester, "Model independent sparticle mass measurements at ATLAS," CERN-THESIS-2004-003. 21, 64

[26] B. K. Gjelsten, D. J. Miller and P. Osland, "Measurement of SUSY masses via cascade decays for SPS 1a," JHEP **0**412, 003 (2004) doi:10.1088/1126-6708/2004/12/003 [hep-ph/0410303]. 21, 54, 64, 71, 73, 83, 114

[27] B. K. Gjelsten, D. J. Miller and P. Osland, "Measurement of the gluino mass via cascade decays for SPS 1a," JHEP **0**506, 015 (2005) doi:10.1088/1126-6708/2005/06/015 [hep-ph/0501033]. 21, 64, 114

[28] C. G. Lester, M. A. Parker and M. J. White, "Determining SUSY model parameters and masses at the LHC using cross-sections, kinematic edges and other observables," JHEP **0**601, 080 (2006) doi:10.1088/1126-6708/2006/01/080 [hep-ph/0508143]. 21, 64, 114

[29] D. J. Miller, P. Osland and A. R. Raklev, "Invariant mass distributions in cascade decays," JHEP **0**603, 034 (2006) doi:10.1088/1126-6708/2006/03/034 [hep-ph/0510356]. 21, 64, 114

[30] C. G. Lester, "Constrained invariant mass distributions in cascade decays: The Shape of the '$m(qll)$-threshold' and similar distributions," Phys. Lett. B **6**55, 39 (2007) doi:10.1016/j.physletb.2007.08.061 [hep-ph/0603171]. 21, 66

[31] G. G. Ross and M. Serna, "Mass determination of new states at hadron colliders," Phys. Lett. B **6**65, 212 (2008) doi:10.1016/j.physletb.2008.06.003 [arXiv:0712.0943 [hep-ph]]. 21, 63

[32] A. J. Barr, G. G. Ross and M. Serna, "The Precision Determination of Invisible-Particle Masses at the LHC," Phys. Rev. D **7**8, 056006 (2008) doi:10.1103/PhysRevD.78.056006 [arXiv:0806.3224 [hep-ph]]. 21, 63

[33] A. J. Barr, A. Pinder and M. Serna, "Precision Determination of Invisible-Particle Masses at the CERN LHC. II.," Phys. Rev. D **7**9, 074005 (2009) doi:10.1103/PhysRevD.79.074005 [arXiv:0811.2138 [hep-ph]]. 21, 63

[34] W. S. Cho, D. Kim, K. T. Matchev and M. Park, "Probing Resonance Decays to Two Visible and Multiple Invisible Particles," Phys. Rev. Lett. **1**12, no. 21, 211801 (2014) doi:10.1103/PhysRevLett.112.211801 [arXiv:1206.1546 [hep-ph]]. 21, 61

[35] B. K. Gjelsten, D. J. Miller and P. Osland, "Resolving ambiguities in mass determinations at future colliders," eConf C **0**50318, 0211 (2005) [hep-ph/0507232]. 21, 66

[36] M. Burns, K. T. Matchev and M. Park, "Using kinematic boundary lines for particle mass measurements and disambiguation in SUSY-like events with missing energy," JHEP **0**905, 094 (2009) doi:10.1088/1126-6708/2009/05/094 [arXiv:0903.4371 [hep-ph]]. 21, 23, 55, 66, 71, 75, 84, 94, 114

[37] K. T. Matchev, F. Moortgat, L. Pape and M. Park, "Precise reconstruction of sparticle masses without ambiguities," JHEP **0**908, 104 (2009) doi:10.1088/1126-6708/2009/08/104 [arXiv:0906.2417 [hep-ph]]. 21, 23, 54, 64, 114

[38] I. Hinchliffe and F. E. Paige, "Measurements in gauge mediated SUSY breaking models at CERN LHC," Phys. Rev. D **6**0, 095002 (1999) doi:10.1103/PhysRevD.60.095002 [hep-ph/9812233]. 21, 63

[39] M. M. Nojiri, G. Polesello and D. R. Tovey, "Proposal for a new reconstruction technique for SUSY processes at the LHC," hep-ph/0312317. 21, 63

[40] K. Kawagoe, M. M. Nojiri and G. Polesello, "A New SUSY mass reconstruction method at the CERN LHC," Phys. Rev. D **7**1, 035008 (2005) doi:10.1103/PhysRevD.71.035008 [hep-ph/0410160]. 21, 63

[41] H. C. Cheng, J. F. Gunion, Z. Han, G. Marandella and B. McElrath, "Mass determination in SUSY-like events with missing energy," JHEP **0**712, 076 (2007) doi:10.1088/1126-6708/2007/12/076 [arXiv:0707.0030 [hep-ph]]. 21, 63

[42] M. M. Nojiri and M. Takeuchi, "Study of the top reconstruction in top-partner events at the LHC," JHEP **0**810, 025 (2008) doi:10.1088/1126-6708/2008/10/025 [arXiv:0802.4142 [hep-ph]]. 21, 63

[43] H. C. Cheng, D. Engelhardt, J. F. Gunion, Z. Han and B. McElrath, "Accurate Mass Determinations in Decay Chains with Missing Energy," Phys. Rev. Lett. **1**00, 252001 (2008) doi:10.1103/PhysRevLett.100.252001 [arXiv:0802.4290 [hep-ph]]. 21, 63

[44] H. C. Cheng, J. F. Gunion, Z. Han and B. McElrath, "Accurate Mass Determinations in Decay Chains with Missing Energy. II," Phys. Rev. D **8**0, 035020 (2009) doi:10.1103/PhysRevD.80.035020 [arXiv:0905.1344 [hep-ph]]. 21, 63

[45] B. Webber, "Mass determination in sequential particle decay chains," JHEP **0**909, 124 (2009) doi:10.1088/1126-6708/2009/09/124 [arXiv:0907.5307 [hep-ph]]. 21, 63

[46] C. Autermann, B. Mura, C. Sander, H. Schettler and P. Schleper, "Determination of supersymmetric masses using kinematic fits at the LHC," arXiv:0911.2607 [hep-ph]. 21, 63

[47] Z. Kang, N. Kersting, S. Kraml, A. R. Raklev and M. J. White, "Neutralino Reconstruction at the LHC from Decay-frame Kinematics," Eur. Phys. J. C **7**0, 271 (2010) doi:10.1140/epjc/s10052-010-1477-1 [arXiv:0908.1550 [hep-ph]]. 21, 63

[48] M. M. Nojiri, K. Sakurai and B. R. Webber, "Reconstructing particle masses from pairs of decay chains," JHEP **1**006, 069 (2010) doi:10.1007/JHEP06(2010)069 [arXiv:1005.2532 [hep-ph]]. 21, 63

[49] Z. Kang, N. Kersting and M. White, "Mass Estimation without using MET in early LHC data," arXiv:1007.0382 [hep-ph]. 21, 63

[50] J. Hubisz and J. Shao, "Mass Measurement in Boosted Decay Chains with Missing Energy," Phys. Rev. D **8**4, 035031 (2011) doi:10.1103/PhysRevD.84.035031 [arXiv:1009.1148 [hep-ph]]. 21, 63

[51] H. C. Cheng, Z. Han, I. W. Kim and L. T. Wang, "Missing Momentum Reconstruction and Spin Measurements at Hadron Colliders," JHEP **1**011, 122 (2010) doi:10.1007/JHEP11(2010)122 [arXiv:1008.0405 [hep-ph]]. 21, 63

[52] B. Gripaios, K. Sakurai and B. Webber, "Polynomials, Riemann surfaces, and reconstructing missing-energy events," JHEP **1**109, 140 (2011) doi:10.1007/JHEP09(2011)140 [arXiv:1103.3438 [hep-ph]]. 21, 63

[53] A. Barr, C. Lester and P. Stephens, "m(T2): The Truth behind the glamour," J. Phys. G **2**9, 2343 (2003) doi:10.1088/0954-3899/29/10/304 [hep-ph/0304226]. 22, 63

[54] M. Baumgart, T. Hartman, C. Kilic and L. T. Wang, "Discovery and measurement of sleptons, binos, and winos with a Z-prime," JHEP **0**711, 084 (2007) doi:10.1088/1126-6708/2007/11/084 [hep-ph/0608172]. 22, 63

[55] C. Lester and A. Barr, "MTGEN: Mass scale measurements in pair-production at colliders," JHEP **0**712, 102 (2007) doi:10.1088/1126-6708/2007/12/102 [arXiv:0708.1028 [hep-ph]]. 22, 63

[56] D. R. Tovey, "On measuring the masses of pair-produced semi-invisibly decaying particles at hadron colliders," JHEP **0**804, 034 (2008) doi:10.1088/1126-6708/2008/04/034 [arXiv:0802.2879 [hep-ph]]. 22, 63

[57] M. Serna, "A Short comparison between $m(T2)$ and $m(CT)$," JHEP **0**806, 004 (2008) doi:10.1088/1126-6708/2008/06/004 [arXiv:0804.3344 [hep-ph]]. 22, 63

[58] M. M. Nojiri, K. Sakurai, Y. Shimizu and M. Takeuchi, "Handling jets + missing E(T) channel using inclusive m(T2)," JHEP **0**810, 100 (2008) doi:10.1088/1126-6708/2008/10/100 [arXiv:0808.1094 [hep-ph]]. 22, 63

[59] W. S. Cho, K. Choi, Y. G. Kim and C. B. Park, "M(T2)-assisted on-shell reconstruction of missing momenta and its application to spin measurement at the LHC," Phys. Rev. D **7**9, 031701 (2009) doi:10.1103/PhysRevD.79.031701 [arXiv:0810.4853 [hep-ph]]. 22, 63

[60] H. C. Cheng and Z. Han, "Minimal Kinematic Constraints and m(T2)," JHEP **0**812, 063 (2008) doi:10.1088/1126-6708/2008/12/063 [arXiv:0810.5178 [hep-ph]]. 22, 63

[61] M. Burns, K. Kong, K. T. Matchev and M. Park, "Using Subsystem MT2 for Complete Mass Determinations in Decay Chains with Missing Energy at Hadron Colliders," JHEP **0**903, 143 (2009) doi:10.1088/1126-6708/2009/03/143 [arXiv:0810.5576 [hep-ph]]. 22, 63

[62] K. Choi, S. Choi, J. S. Lee and C. B. Park, "Reconstructing the Higgs boson in dileptonic W decays at hadron collider," Phys. Rev. D **8**0, 073010 (2009) doi:10.1103/PhysRevD.80.073010 [arXiv:0908.0079 [hep-ph]]. 22, 63

[63] K. T. Matchev and M. Park, "A General method for determining the masses of semi-invisibly decaying particles at hadron colliders," Phys. Rev. Lett. **1**07, 061801 (2011) doi:10.1103/PhysRevLett.107.061801 [arXiv:0910.1584 [hep-ph]]. 22, 23, 63

[64] G. Polesello and D. R. Tovey, "Supersymmetric particle mass measurement with the boost-corrected contransverse mass," JHEP **1**003, 030 (2010) doi:10.1007/JHEP03(2010)030 [arXiv:0910.0174 [hep-ph]]. 22, 63

[65] P. Konar, K. Kong, K. T. Matchev and M. Park, "Superpartner Mass Measurement Technique using 1D Orthogonal Decompositions of the Cambridge Transverse Mass Variable M_{T2}," Phys. Rev. Lett. **1**05, 051802 (2010) doi:10.1103/PhysRevLett.105.051802 [arXiv:0910.3679 [hep-ph]]. 22, 63

[66] W. S. Cho, J. E. Kim and J. H. Kim, "Amplification of endpoint structure for new particle mass measurement at the LHC," Phys. Rev. D **8**1, 095010 (2010) doi:10.1103/PhysRevD.81.095010 [arXiv:0912.2354 [hep-ph]]. 22, 63

[67] M. M. Nojiri and K. Sakurai, "Controlling ISR in sparticle mass reconstruction," Phys. Rev. D **8**2, 115026 (2010) doi:10.1103/PhysRevD.82.115026 [arXiv:1008.1813 [hep-ph]]. 22, 63

[68] K. Agashe, R. Franceschini and D. Kim, "Simple ÒinvarianceÓ of two-body decay kinematics," Phys. Rev. D **8**8, no. 5, 057701 (2013) doi:10.1103/PhysRevD.88.057701 [arXiv:1209.0772 [hep-ph]]. 22, 63

[69] K. Agashe, R. Franceschini and D. Kim, "Using Energy Peaks to Measure New Particle Masses," JHEP **1**411, 059 (2014) doi:10.1007/JHEP11(2014)059 [arXiv:1309.4776 [hep-ph]]. 22, 63

[70] K. Agashe, R. Franceschini, D. Kim and K. Wardlow, "Mass Measurement Using Energy Spectra in Three-body Decays," JHEP **1**605, 138 (2016) doi:10.1007/JHEP05(2016)138 [arXiv:1503.03836 [hep-ph]]. 22, 63

[71] K. Agashe, R. Franceschini, S. Hong and D. Kim, "Energy spectra of massive two-body decay products and mass measurement," JHEP **1**604, 151 (2016) doi:10.1007/JHEP04(2016)151 [arXiv:1512.02265 [hep-ph]]. 22, 63

[72] K. Kondo, "Dynamical Likelihood Method for Reconstruction of Events With Missing Momentum. 1: Method and Toy Models," J. Phys. Soc. Jap. **5**7, 4126 (1988). doi:10.1143/JPSJ.57.4126 22

[73] R. H. Dalitz and G. R. Goldstein, "The Decay and polarization properties of the top quark," Phys. Rev. D **4**5, 1531 (1992). doi:10.1103/PhysRevD.45.1531 22

[74] V. M. Abazov *et* al. [D0 Collaboration], "A precision measurement of the mass of the top quark," Nature **4**29, 638 (2004) doi:10.1038/nature02589 [hep-ex/0406031]. 22

[75] J. Alwall, A. Freitas and O. Mattelaer, "Measuring Sparticles with the Matrix Element," AIP Conf. Proc. **1**200, 442 (2010) doi:10.1063/1.3327611 [arXiv:0910.2522 [hep-ph]]. 22

[76] P. Artoisenet, V. Lemaitre, F. Maltoni and O. Mattelaer, "Automation of the matrix element reweighting method," JHEP **1**012, 068 (2010) doi:10.1007/JHEP12(2010)068 [arXiv:1007.3300 [hep-ph]]. 22

[77] J. Alwall, A. Freitas and O. Mattelaer, "The Matrix Element Method and QCD Radiation," Phys. Rev. D **8**3, 074010 (2011) doi:10.1103/PhysRevD.83.074010 [arXiv:1010.2263 [hep-ph]]. 22

[78] F. Fiedler, A. Grohsjean, P. Haefner and P. Schieferdecker, "The Matrix Element Method and its Application in Measurements of the Top Quark Mass," Nucl. Instrum. Meth. A **6**24, 203 (2010) doi:10.1016/j.nima.2010.09.024 [arXiv:1003.1316 [hep-ex]]. 22

[79] J. S. Gainer, J. Lykken, K. T. Matchev, S. Mrenna and M. Park, "The Matrix Element Method: Past, Present, and Future," arXiv:1307.3546 [hep-ph]. 22

[80] W. S. Cho, K. Choi, Y. G. Kim and C. B. Park, "Gluino Stransverse Mass," Phys. Rev. Lett. **1**00, 171801 (2008) doi:10.1103/PhysRevLett.100.171801 [arXiv:0709.0288 [hep-ph]]. 22, 63

[81] B. Gripaios, "Transverse observables and mass determination at hadron colliders," JHEP **0**802, 053 (2008) doi:10.1088/1126-6708/2008/02/053 [arXiv:0709.2740 [hep-ph]]. 22, 63

[82] A. J. Barr, B. Gripaios and C. G. Lester, "Weighing Wimps with Kinks at Colliders: Invisible Particle Mass Measurements from Endpoints," JHEP **0**802, 014 (2008) doi:10.1088/1126-6708/2008/02/014 [arXiv:0711.4008 [hep-ph]]. 22, 63

[83] W. S. Cho, K. Choi, Y. G. Kim and C. B. Park, "Measuring superparticle masses at hadron collider using the transverse mass kink," JHEP **0**802, 035 (2008) doi:10.1088/1126-6708/2008/02/035 [arXiv:0711.4526 [hep-ph]]. 22, 63

[84] M. M. Nojiri, Y. Shimizu, S. Okada and K. Kawagoe, "Inclusive transverse mass analysis for squark and gluino mass determination," JHEP **0**806, 035 (2008) doi:10.1088/1126-6708/2008/06/035 [arXiv:0802.2412 [hep-ph]]. 22, 63

[85] A. J. Barr, B. Gripaios and C. G. Lester, "Transverse masses and kinematic constraints: from the boundary to the crease," JHEP **0**911, 096 (2009) doi:10.1088/1126-6708/2009/11/096 [arXiv:0908.3779 [hep-ph]]. 22, 63

[86] K. T. Matchev, F. Moortgat, L. Pape and M. Park, "Precision sparticle spectroscopy in the inclusive same-sign dilepton channel at LHC," Phys. Rev. D **8**2, 077701 (2010) doi:10.1103/PhysRevD.82.077701 [arXiv:0909.4300 [hep-ph]]. 22, 63

[87] P. Konar, K. Kong, K. T. Matchev and M. Park, "Dark Matter Particle Spectroscopy at the LHC: Generalizing M_{T2} to Asymmetric Event Topologies," JHEP **1**004, 086 (2010) doi:10.1007/JHEP04(2010)086 [arXiv:0911.4126 [hep-ph]]. 22, 63

[88] P. Konar, K. Kong and K. T. Matchev, "$\sqrt{\hat{s}}_{min}$: A Global inclusive variable for determining the mass scale of new physics in events with missing energy at hadron colliders," JHEP **0**903, 085 (2009) doi:10.1088/1126-6708/2009/03/085 [arXiv:0812.1042 [hep-ph]]. 22, 63

[89] P. Konar, K. Kong, K. T. Matchev and M. Park, "RECO level $\sqrt{s}_{min}$ and subsystem $\sqrt{s}_{min}$: Improved global inclusive variables for measuring the new physics mass scale inE_T events at hadron colliders," JHEP **1**106, 041 (2011) doi:10.1007/JHEP06(2011)041 [arXiv:1006.0653 [hep-ph]]. 22, 63

[90] A. J. Barr, T. J. Khoo, P. Konar, K. Kong, C. G. Lester, K. T. Matchev and M. Park, "Guide to transverse projections and mass-constraining variables," Phys. Rev. D **8**4, 095031 (2011) doi:10.1103/PhysRevD.84.095031 [arXiv:1105.2977 [hep-ph]]. 22, 63, 112

[91] T. Robens, "$\sqrt{\hat{s}}_{min}$ resurrected," JHEP **1**202, 051 (2012) doi:10.1007/JHEP02(2012)051 [arXiv:1109.1018 [hep-ph]]. 22, 63

[92] R. Mahbubani, K. T. Matchev and M. Park, "Re-interpreting the Oxbridge stransverse mass variable MT2 in general cases," JHEP **1**303, 134 (2013) doi:10.1007/JHEP03(2013)134 [arXiv:1212.1720 [hep-ph]]. 22, 63, 112

[93] Y. Bai, H. C. Cheng, J. Gallicchio and J. Gu, "A Toolkit of the Stop Search via the Chargino Decay," JHEP **1**308, 085 (2013) doi:10.1007/JHEP08(2013)085 [arXiv:1304.3148 [hep-ph]]. 22, 63

[94] W. S. Cho, J. S. Gainer, D. Kim, K. T. Matchev, F. Moortgat, L. Pape and M. Park, "On-shell constrained M_2 variables with applications to

mass measurements and topology disambiguation," JHEP **1**408, 070 (2014) doi:10.1007/JHEP08(2014)070 [arXiv:1401.1449 [hep-ph]]. 22, 61, 112

[95] W. S. Cho, J. S. Gainer, D. Kim, K. T. Matchev, F. Moortgat, L. Pape and M. Park, "Improving the sensitivity of stop searches with on-shell constrained invariant mass variables," JHEP **1**505, 040 (2015) doi:10.1007/JHEP05(2015)040 [arXiv:1411.0664 [hep-ph]]. 22, 63, 111

[96] W. S. Cho, J. S. Gainer, D. Kim, S. H. Lim, K. T. Matchev, F. Moortgat, L. Pape and M. Park, "OPTIMASS: A Package for the Minimization of Kinematic Mass Functions with Constraints," JHEP **1**601, 026 (2016) doi:10.1007/JHEP01(2016)026 [arXiv:1508.00589 [hep-ph]]. 22, 63

[97] P. Konar and A. K. Swain, "Mass reconstruction with M_2 under constraint in semi-invisible production at a hadron collider," Phys. Rev. D **9**3, no. 1, 015021 (2016) doi:10.1103/PhysRevD.93.015021 [arXiv:1509.00298 [hep-ph]]. 22, 63

[98] D. Costanzo and D. R. Tovey, "Supersymmetric particle mass measurement with invariant mass correlations," JHEP **0**904, 084 (2009) doi:10.1088/1126-6708/2009/04/084 [arXiv:0902.2331 [hep-ph]]. 23, 48, 50, 66, 114, 116

[99] D. Kim, K. T. Matchev and M. Park, "Using sorted invariant mass variables to evade combinatorial ambiguities in cascade decays," JHEP **1**602, 129 (2016) doi:10.1007/JHEP02(2016)129 [arXiv:1512.02222 [hep-ph]]. 23, 39, 48, 50, 62, 66, 106, 114, 116

[100] I. W. Kim, "Algebraic Singularity Method for Mass Measurement with Missing Energy," Phys. Rev. Lett. **1**04, 081601 (2010) doi:10.1103/PhysRevLett.104.081601 [arXiv:0910.1149 [hep-ph]]. 23

[101] G. Voronoi, "Nouvelles applications des paramètres continus à la théorie des formes quadratiques," Journal für die Reine und Angewandte Mathematik, **1**33, 97 (1908). 25

[102] G. L. Dirichlet, "Über die Reduktion der positiven quadratischen Formen mit drei unbestimmten ganzen Zahlen," Journal für die Reine und Angewandte Mathematik, **4**0, 209 (1850). 25

[103] R. Descartes, Principia philosophiae (1644). 25

[104] See, e.g., R. E. Miles and R. J. Maillardet, "The basic structures of Voronoi and generalized Voronoi polygons," J. of Appl. Prob. **1**9A, 97-112 (1982). 29

[105] J. M. Drouffe and C. Itzykson, "Random Geometry and the Statistics of Two-dimensional Cells," Nucl. Phys. B **2**35, 45 (1984). doi:10.1016/0550-3213(84)90147-0 29

[106] A. L. Hinde and R. E. Miles, "Monte Carlo estimates of the distributions of the random polygons of the Voronoi tessellation with respect to a Poisson process," J. Statist. Comput. Simul. **1**0, 205-223 (1980). 30

[107] D. Debnath, J. S. Gainer, D. Kim and K. T. Matchev, "Discovering New Physics with Voronoi Tessellations," arXiv:1511.02724 [hep-ph]. 30, 68

[108] U. Lorz, "Distribution of cell characteristics of the spatial Poisson-Voronoi tessellation and plane sections," Geometrical Problems of Image Processing, Proc. 1991 Georgenthal workshop, ed. U. Eckhardt, A. Hybler, W. Nagel and G. Werner, Akademie-Verlag, 1991, pp. 171-178. 33

[109] K. A. Olive *et* al. [Particle Data Group Collaboration], "Review of Particle Physics," Chin. Phys. C **3**8, 090001 (2014). doi:10.1088/1674-1137/38/9/090001 37

[110] D. Debnath, J. .S. Gainer, C. Kilic, D. Kim, K. T. Matchev, and Y.-P. Yang, in progress. 41, 43, 58, 109

[111] J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer and T. Stelzer, "MadGraph 5 : Going Beyond," JHEP **1**106, 128 (2011) doi:10.1007/JHEP06(2011)128 [arXiv:1106.0522 [hep-ph]]. 47

[112] G. Aad *et* al. [ATLAS Collaboration], "Search for supersymmetry in events containing a same-flavour opposite-sign dilepton pair, jets, and large missing transverse momentum in $\sqrt{s} = 8$ TeV pp collisions with the ATLAS detector," Eur. Phys. J. C **7**5, no. 7, 318 (2015) Erratum: [Eur. Phys. J. C **7**5, no. 10, 463 (2015)] doi:10.1140/epjc/s10052-015-3661-9, 10.1140/epjc/s10052-015-3518-2 [arXiv:1503.03290 [hep-ex]]. 47

[113] T. Han, I. W. Kim and J. Song, "Kinematic Cusps: Determining the Missing Particle Mass at Colliders," Phys. Lett. B **6**93, 575 (2010) doi:10.1016/j.physletb.2010.09.010 [arXiv:0906.5009 [hep-ph]]. 59

[114] T. Han, I. W. Kim and J. Song, "Kinematic Cusps With Two Missing Particles I: Antler Decay Topology," Phys. Rev. D **8**7, no. 3, 035003 (2013) doi:10.1103/PhysRevD.87.035003 [arXiv:1206.5633 [hep-ph]]. 59

[115] T. Han, I. W. Kim and J. Song, "Kinematic Cusps with Two Missing Particles II: Cascade Decay Topology," Phys. Rev. D **8**7, no. 3, 035004 (2013) doi:10.1103/PhysRevD.87.035004 [arXiv:1206.5641 [hep-ph]]. 59

[116] "Receiver operating characteristic", Accessed June 8, 2016. 131

[117] J. Neyman and E. S. Pearson, "On the problem of the most efficient tests of statistical hypotheses", Philosophical Transactions of the Royal Society of London. Series A, Vol. 231 (1933), pp. 289-337. 133

[118] P. Cushman *et* al., "Working Group Report: WIMP Dark Matter Direct Detection," arXiv:1310.8327 [hep-ex]. 60

[119] J. Buckley *et* al., "Working Group Report: WIMP Dark Matter Indirect Detection," arXiv:1310.7040 [astro-ph.HE]. 60

[120] S. Arrenberg *et* al., "Working Group Report: Dark Matter Complementarity," arXiv:1310.8621 [hep-ph]. 60

[121] D. Abercrombie *et* al., "Dark Matter Benchmark Models for Early LHC Run-2 Searches: Report of the ATLAS/CMS Dark Matter Forum," arXiv:1507.00966 [hep-ex]. 60

[122] F. Moortgat and L. Pape, CMS Physics TDR, Vol. II, Report No. CERN-LHCC-2006, Chap. 13.4, p. 410. 62

[123] S. Matsumoto, M. M. Nojiri and D. Nomura, "Hunting for the Top Partner in the Littlest Higgs Model with T-parity at the CERN LHC," Phys. Rev. D **7**5, 055006 (2007) [hep-ph/0612249]. 62

[124] A. Rajaraman and F. Yu, "A New Method for Resolving Combinatorial Ambiguities at Hadron Colliders," Phys. Lett. B **7**00, 126 (2011) [arXiv:1009.2751 [hep-ph]]. 62

[125] Y. Bai and H. C. Cheng, "Identifying Dark Matter Event Topologies at the LHC," JHEP **1**106, 021 (2011) [arXiv:1012.1863 [hep-ph]]. 62

[126] P. Baringer, K. Kong, M. McCaskey and D. Noonan, "Revisiting Combinatorial Ambiguities at Hadron Colliders with M_{T2}," JHEP **1**110, 101 (2011) [arXiv:1109.1563 [hep-ph]]. 62

[127] K. Choi, D. Guadagnoli and C. B. Park, "Reducing combinatorial uncertainties: A new technique based on MT2 variables," JHEP **1**111, 117 (2011) [arXiv:1109.2201 [hep-ph]]. 62

[128] M. D. Klimek, "Ordered Kinematic Endpoints for 5-body Cascade Decays," arXiv:1610.08603 [hep-ph]. 62, 106

[129] H. C. Cheng, K. T. Matchev and M. Schmaltz, "Bosonic supersymmetry? Getting fooled at the CERN LHC," Phys. Rev. D **6**6, 056006 (2002) doi:10.1103/PhysRevD.66.056006 [hep-ph/0205314]. 62

[130] A. Freitas and D. Wyler, "Phenomenology of mirror fermions in the littlest Higgs model with T-parity," JHEP **0**611, 061 (2006) doi:10.1088/1126-6708/2006/11/061 [hep-ph/0609103]. 62

[131] A. Airapetian *et* al. [ATLAS Collaboration], "ATLAS: Detector and physics performance technical design report. Volume 2," CERN-LHCC-99-15, ATLAS-TDR-15. 65

[132] M. M. Nojiri, D. Toya and T. Kobayashi, "Lepton energy asymmetry and precision SUSY study at hadron colliders," Phys. Rev. D **6**2, 075009 (2000) doi:10.1103/PhysRevD.62.075009 [hep-ph/0001267]. 65

[133] C. Lester, "Mass and Spin Measurement Techniques (for the LHC)", in "The Dark Secrets of the Terascale : Proceedings, TASI 2011, Boulder, Colorado, USA, Jun 6 - Jul 11, 2011," Eds. T. Tait and K. Matchev. 66

[134] D. Debnath, J. S. Gainer, C. Kilic, D. Kim, K. T. Matchev and Y. P. Yang, "Identifying Phase Space Boundaries with Voronoi Tessellations," arXiv:1606.02721 [hep-ph]. 67, 68, 69, 83, 87, 89, 108, 114

[135] B. Altunkaynak, C. Kilic and M. D. Klimek, "Multidimensional phase space methods for mass measurements and decay topology determination," Eur. Phys. J. C **77**, no. 2, 61 (2017) doi:10.1140/epjc/s10052-017-4631-1 [arXiv:1611.09764 [hep-ph]]. 67, 108, 114

[136] C. B. Barber, D. P Dobkin and H. T. Huhdanpaa, "The Quickhull algorithm for convex hulls," ACM Trans. on Mathematical Software, 22(4):469-483, Dec 1996, http://www.qhull.org. 67

[137] TGenPhaseSpace Class Reference, https://root.cern.ch/doc/v608/classTGenPhaseSpace.html 79

[138] S. Chatrchyan *et* al. [CMS Collaboration], "Measurement of masses in the $t\bar{t}$ system by kinematic endpoints in pp collisions at $\sqrt{s} = 7$ TeV," Eur. Phys. J. C **7**3, 2494 (2013) doi:10.1140/epjc/s10052-013-2494-7 [arXiv:1304.5783 [hep-ex]]. 81

[139] B. K. Gjelsten, D. J. Miller, P. Osland and A. R. Raklev, "Mass Determination in Cascade Decays Using Shape Formulas," AIP Conf. Proc. **9**03, 257 (2007) doi:10.1063/1.2735174 [hep-ph/0611259]. 83, 141

[140] J. Alwall *et* al., "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations," JHEP **1**407, 079 (2014) doi:10.1007/JHEP07(2014)079 [arXiv:1405.0301 [hep-ph]]. 91, 127

[141] J. Reuter and D. Wiesler, "Distorted mass edges at LHC from supersymmetric leptoquarks," Phys. Rev. D **8**4, 015012 (2011) doi:10.1103/PhysRevD.84.015012 [arXiv:1010.4215 [hep-ph]]. 91

[142] B. C. Allanach, S. Grab and H. E. Haber, "Supersymmetric Monojets at the Large Hadron Collider," JHEP **1**101, 138 (2011) Erratum: [JHEP **1**107, 087 (2011)] Erratum: [JHEP **1**109, 027 (2011)] doi:10.1007/JHEP07(2011)087,

10.1007/JHEP09(2011)027, 10.1007/JHEP01(2011)138 [arXiv:1010.4261 [hep-ph]]. 91

[143] C. Macesanu, C. D. McMullen and S. Nandi, "Collider implications of universal extra dimensions," Phys. Rev. D **6**6, 015009 (2002) doi:10.1103/PhysRevD.66.015009 [hep-ph/0201300]. 91

[144] M. ElKacimi, D. Goujdami, H. Przysiezniak and P. Z. Skands, "One Universal Extra Dimension in Pythia," Comput. Phys. Commun. **1**81, 122 (2010) doi:10.1016/j.cpc.2009.08.008 [arXiv:0901.4087 [hep-ph]]. 91

[145] J. Beuria, A. Datta, D. Debnath and K. T. Matchev, "LHC Collider Phenomenology of Minimal Universal Extra Dimensions," arXiv:1702.00413 [hep-ph]. 91

[146] G. L. Bayatian *et* al. [CMS Collaboration], "CMS physics: Technical design report," CERN-LHCC-2006-001, CMS-TDR-008-1. 105

[147] E. Gildener, "Gauge Symmetry Hierarchies," Phys. Rev. D **1**4, 1667 (1976). doi:10.1103/PhysRevD.14.1667 110

[148] G. Jungman, M. Kamionkowski and K. Griest, "Supersymmetric dark matter," Phys. Rept. **2**67, 195 (1996) doi:10.1016/0370-1573(95)00058-5 [hep-ph/9506380]. 110

[149] See the recent results reported by ATLAS and CMS Collaborations at ICHEP 2018, https://indico.cern.ch/event/686555/contributions/3028060/attachments/1683138/2705249/ICHEP_TancrediCarli.pdf and https://indico.cern.ch/event/686555/contributions/3028061/attachn rahatlou-ichep-20180709-7.pdf 110

[150] S. P. Martin, "Compressed supersymmetry and natural neutralino dark matter from top squark-mediated annihilation to top quarks," Phys. Rev. D **7**5, 115005 (2007) doi:10.1103/PhysRevD.75.115005 [hep-ph/0703097 [HEP-PH]]. 111

[151] H. Baer, A. Box, E. K. Park and X. Tata, "Implications of compressed supersymmetry for collider and dark matter searches," JHEP **0**708, 060 (2007) doi:10.1088/1126-6708/2007/08/060 [arXiv:0707.0618 [hep-ph]]. 111

[152] P. Schwaller and J. Zurita, "Compressed electroweakino spectra at the LHC," JHEP **1**403, 060 (2014) doi:10.1007/JHEP03(2014)060 [arXiv:1312.7350 [hep-ph]]. 111

[153] Z. Han, G. D. Kribs, A. Martin and A. Menon, "Hunting quasidegenerate Higgsinos," Phys. Rev. D **8**9, no. 7, 075007 (2014) doi:10.1103/PhysRevD.89.075007 [arXiv:1401.1235 [hep-ph]]. 111

[154] Z. Han and Y. Liu, "MT2 to the Rescue – Searching for Sleptons in Compressed Spectra at the LHC," Phys. Rev. D **9**2, no. 1, 015010 (2015) doi:10.1103/PhysRevD.92.015010 [arXiv:1412.0618 [hep-ph]]. 111

[155] H. An, J. Gu and L. T. Wang, "Exploring the nearly degenerate stop region with sbottom decays," doi:10.3204/PUBDB-2016-05828 arXiv:1611.09868 [hep-ph]. 111

[156] P. Konar, T. Mondal and A. K. Swain, "Demystifying compressed top squark region with kinematic variables," arXiv:1612.03269 [hep-ph]. 111

[157] M. Zarucki [CMS Collaboration], "Search for Supersymmetry with a Highly Compressed Mass Spectrum in the Single Soft Lepton Channel with the CMS Experiment at the LHC," PoS EPS **-**HEP2017, 728 (2017). doi:10.22323/1.314.0728 111

[158] M. Aaboud *e*t al. [ATLAS Collaboration], "Search for electroweak production of supersymmetric states in scenarios with compressed mass spectra at $\sqrt{s} = 13$ TeV with the ATLAS detector," Phys. Rev. D **9**7, no. 5, 052010 (2018) doi:10.1103/PhysRevD.97.052010 [arXiv:1712.08119 [hep-ex]]. 111

[159] R. Schöfbeck [ATLAS and CMS Collaborations], "Search for supersymmetry with extremely compressed spectra with the ATLAS and CMS detectors," Nucl. Part. Phys. Proc. **2**73-275, 631 (2016). doi:10.1016/j.nuclphysbps.2015.09.095 111

[160] J. Fan, M. Reece and J. T. Ruderman, "Stealth Supersymmetry," JHEP **1**111, 012 (2011) doi:10.1007/JHEP11(2011)012 [arXiv:1105.5135 [hep-ph]]. 111

[161] D. S. M. Alves, M. R. Buckley, P. J. Fox, J. D. Lykken and C. T. Yu, "Stops and $\not{E}_T$: The shape of things to come," Phys. Rev. D **8**7, no. 3, 035016 (2013) doi:10.1103/PhysRevD.87.035016 [arXiv:1205.5805 [hep-ph]]. 111

[162] S. Macaluso, M. Park, D. Shih and B. Tweedie, "Revealing Compressed Stops Using High-Momentum Recoils," JHEP **1**603, 151 (2016) doi:10.1007/JHEP03(2016)151 [arXiv:1506.07885 [hep-ph]]. 111

[163] H. C. Cheng, L. Li and Q. Qin, "Second Stop and Sbottom Searches with a Stealth Stop," JHEP **1**611, 181 (2016) doi:10.1007/JHEP11(2016)181 [arXiv:1607.06547 [hep-ph]]. 111

[164] V. Khachatryan *et* al. [CMS Collaboration], "Search for stealth supersymmetry in events with jets, either photons or leptons, and low missing transverse momentum in pp collisions at 8 TeV," Phys. Lett. B **7**43, 503 (2015) doi:10.1016/j.physletb.2015.03.017 [arXiv:1411.7255 [hep-ex]]. 111

[165] G. Zevi Della Porta [ATLAS and CMS Collaborations], "Compressed and Stealth SUSY searches with ATLAS and CMS," PoS LHCP **2**016, 155 (2016). doi:10.22323/1.276.0155 111

[166] G. Aad *et* al. [ATLAS Collaboration], "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC," Phys. Lett. B **7**16, 1 (2012) doi:10.1016/j.physletb.2012.08.020 [arXiv:1207.7214 [hep-ex]]. 112

[167] S. Chatrchyan *et* al. [CMS Collaboration], "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC," Phys. Lett. B **7**16, 30 (2012) doi:10.1016/j.physletb.2012.08.021 [arXiv:1207.7235 [hep-ex]]. 112

[168] J. Smith, W. L. van Neerven and J. A. M. Vermaseren, "The Transverse Mass and Width of the W Boson," Phys. Rev. Lett. **5**0 (1983) 1738. doi:10.1103/PhysRevLett.50.1738 112

[169] V. D. Barger, A. D. Martin and R. J. N. Phillips, Z. Phys. C **2**1, 99 (1983). doi:10.1007/BF01648783 112

[170] A. Betancur, D. Debnath, J. S. Gainer, K. T. Matchev and P. Shyamsundar, "Measuring the mass, width, and couplings of semi-invisible resonances with the Matrix Element Method," arXiv:1708.07641 [hep-ph]. 112

[171] C. G. Lester and D. J. Summers, "Measuring masses of semiinvisibly decaying particles pair produced at hadron colliders," Phys. Lett. B **4**63, 99 (1999) doi:10.1016/S0370-2693(99)00945-4 [hep-ph/9906349]. 22, 63, 112

[172] W. S. Cho, D. Kim, K. Kong, S. H. Lim, K. T. Matchev, J. C. Park and M. Park, "750 GeV Diphoton Excess May Not Imply a 750 GeV Resonance," Phys. Rev. Lett. **1**16, no. 15, 151805 (2016) doi:10.1103/PhysRevLett.116.151805 [arXiv:1512.06824 [hep-ph]]. 112

[173] I. Hinchliffe and F. E. Paige, "Measurements in SUGRA models with large tan beta at CERN LHC," Phys. Rev. D **6**1, 095011 (2000) doi:10.1103/PhysRevD.61.095011 [hep-ph/9907519]. 114

[174] C. Lester, "Mass and Spin Measurement Techniques (for the LHC)", in "The Dark Secrets of the Terascale : Proceedings, TASI 2011, Boulder, Colorado, USA, Jun 6 - Jul 11, 2011," Eds. T. Tait and K. Matchev. 114, 116, 120, 127, 130

[175] D. Debnath, J. S. Gainer, C. Kilic, D. Kim, K. T. Matchev and Y. P. Yang, "Detecting kinematic boundary surfaces in phase space: particle mass measurements in SUSY-like events," JHEP **1**706, 092 (2017) doi:10.1007/JHEP06(2017)092 [arXiv:1611.04487 [hep-ph]]. 114, 124, 125

[176] G. L. Bayatian *et* al. [CMS Collaboration], "CMS Physics : Technical Design Report Volume 1: Detector Performance and Software," CERN-LHCC-2006-001, CMS-TDR-8-1. 122, 127