PHYSICAL REVIEW D **80,** 035020 (2009)

Accurate mass determinations in decay chains with missing energy: II

Hsin-Chia Cheng,[1] John F. Gunion,[1] Zhenyu Han,[1] and Bob McElrath[2]
[1]*Department of Physics, University of California, Davis, California 95616, USA*
[2]*CERN, Geneva 23, Switzerland*
(Received 20 May 2009; published 25 August 2009)

We discuss kinematic methods for determining the masses of the particles in events at a hadron collider in which a pair of identical particles is produced with each decaying via a series of on shell intermediate beyond-the-standard model (BSM) particles to visible standard model (SM) particles and an invisible particle (schematically, $pp \to ZZ + \text{jets}$ with $Z \to Aa \to Bba \to Ccba \to \ldots \to cba \ldots + N$ where $a, b, c, \ldots$ are visible SM particles or groups of SM particles, $A, B, C, \ldots$ are on shell BSM particles, and N is invisible). This topology arises in many models including supersymmetry (SUSY) processes such as squark and gluino pair production and decay. We present the detailed procedure for the case of $Z \to 3$ visible particles $+ N$ and demonstrate that the masses obtained from the kinematic procedure are independent of the model by comparing SUSY to universal extra dimensions.

DOI: 10.1103/PhysRevD.80.035020 PACS numbers: 14.80.Ly, 12.60.Cn

I. INTRODUCTION

Many solutions to the hierarchy problem require new particles whose loop corrections to the Higgs mass squared cancel the quadratically divergent standard model (SM) loop corrections. The masses of the new particles (especially the particle or particles which cancel the SM top loop) should be sufficiently small that the Higgs mass can be naturally below the TeV scale. On the other hand, if the new particles have masses below $\mathcal{O}$(TeV) then LEP observables will be strongly affected if they are exchanged at tree level or can be singly produced. Both these latter possibilities are automatically removed if there is a symmetry under which the new particles are odd and the SM particles are even. In particular, the new particles can then only contribute to the electroweak observables at the loop level, and new particles with masses of order of a few hundreds of GeV can be compatible with the data. In such scenarios the lightest of the new particles is automatically stable and it should be neutral for consistency with bounds on new charged stable matter. Typically, it is also weakly interacting. Such a weakly interaction massive stable particle will be invisible, leading to "missing energy" in particle detectors. This scenario is also highly desirable since such a weakly interaction massive stable particle can readily provide the dark matter known to be present in the Universe.

Almost all the models with dark matter candidates also contain additional particles that are charged not only under the new symmetry but also carry SM "charges," most often including color. At a collider, these new particles must (and will) be pair produced, and since they are heavier than the dark matter particle, they will cascade decay down to it. In many cases, this cascade radiates SM particles in a series of $A \to Bc$, $1 - \text{body} \to 2 - \text{body}$ decays, in which A and B are new physics particles while c is a SM particle. (In some cases, phase space restrictions force one of the new particles off shell and $A \to B^* c \to Cdc$, $1 - \text{body} \to 3 - \text{body}$ decays are relevant.) Since the final step in the chain will yield a dark matter particle, the typical collider signals for such a scenario will be jets and/or leptons plus missing energy.

Supersymmetry (SUSY) is the most popular model of this type. In SUSY, the new symmetry is termed matter parity (sometimes called R parity). Its conservation implies that the lightest supersymmetric particle is stable. In most supersymmetric models, the lightest supersymmetric particle is the lightest neutralino, which is a good dark matter candidate. It appears at the end of every supersymmetric particle decay chain and escapes the detector. All supersymmetric particles are produced in pairs, resulting in at least two missing particles in each event.

Other theories of TeV scale physics with dark matter candidates have been recently proposed. They have experimental signatures very similar to SUSY, i.e., multileptons and/or jets plus missing energy. For instance, universal extra dimensions (UEDs) [1,2], little Higgs theories with T parity [3], and warped extra dimensions with a Z_3 parity [4] belong to this category of models.

Clearly, being able to reconstruct events with missing energy is an important first step to distinguish various scenarios and establish the underlying theory. In addition, studies [5] suggest that the mass of the dark matter particle, and the masses of any other particles with which it can coannihilate, need to be determined to within a few GeV in order to be able to compute the dark matter density in the context of a given model. A very important question is then whether or not the LHC can achieve such accuracy or will it be necessary to wait for threshold scan data from the ILC. The goal of this paper will be to provide details regarding the kinematic techniques developed in Refs. [6,7] that provide the needed accuracy using just LHC data. For the case of 3 visible particles per decay chain, the focus of this paper, we also show that the kinematic technique gives

 © 2009 The American Physical Society

masses for the beyond-the-standard model (BSM) particles that are completely insensitive to the particular model by comparing a SUSY case to a UED case where the decaying BSM and final invisible BSM particles in the two cases have the same masses. This implies that it is unnecessary to determine the overall mass scale of the BSM particles using model-dependent information, such as total cross sections. Indeed, to fully test a potential model, it is *necessary* to first determine the masses of the produced particles just based on kinematic information. Once the masses are known, there are many chain decay configurations for which it will be possible to determine the four-momenta of all the particles on an event-by-event basis. The four-momenta can then be employed in computing the matrix element squared for different possible spin assignments. In this way, a spin determination may be possible which, in combination with cross section information, can be used to distinguish different models.

In recent years there have been numerous studies in the context of SUSY-like theories of how to measure the superpartner masses just based on kinematic information [8–28]. In some cases the procedures employ a single long decay chain of superparticles, usually requiring 3 or more visible particles in the decay chain in order to have enough invariant mass combinations of the visible particles to achieve sensitivity to the absolute mass scale, as opposed to simply mass differences. Squark decay, Fig. 1, is an example of one such chain.

Our approach has been to pursue alternative procedures that employ events in which a pair of identical (particle and antiparticle, e.g., squark plus antisquark) BSM particles is produced and both decay in the same manner. In such an event, information from both decay chains in the event can be included at once. In our first paper [6], we tackled the difficult case where we assumed that only two particles appeared in each chain decay, e.g., making use only of the leptons appearing in Fig. 1. In this case, kinematic constraints alone cannot give a discrete solution for the unknown masses. Nonetheless, the space of the allowed solutions does contain enough information about the new particle masses and they can be extracted using a statistical procedure [6]. The mass determination can be further improved by combining with other kinematic variables such as M_{T2} [9,24]. However, these kind of analyzes usually require large statistics in order to achieve a reasonable precision. In this paper, we provide details on the case where 3 visible particles are present in each chain decay. In this case, a single pair of events provides enough information to yield a discrete set of possible masses. Therefore, very few events are needed for a rough determination of the masses, in contrast to the statistical methods which rely on the availability of a large number of events.

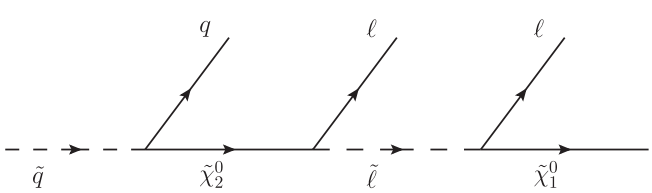

FIG. 1. A decay chain in SUSY.

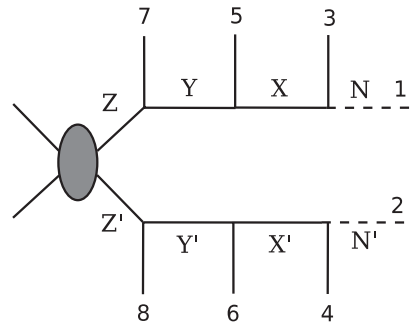

FIG. 2. The event topology we consider.

The general topology on which we focus is then that of Fig. 2. After including combinatorics and resolution, we will achieve root-mean-square accuracies on the three underlying masses in the decay chain of the order of a few GeV (depending upon the number of available events) with a systematic shift that can be easily corrected. This result is fairly stable when backgrounds are included so long as $S/B \gtrsim 2$.

The organization of the paper is as follows. In Sec. II, we give the general counting of constraints and unknowns for single chain and multiple chain events. In Sec. III, we give a more detailed exposition regarding solving the topology of Fig. 2. In Sec. IV, we first demonstrate how the masses of the Z, Y, X, and N particles in Fig. 2 can be very precisely determined using just a few events if there are no effects associated with combinatorics, particle momentum measurement resolutions, or backgrounds. We then develop the very crucial strategies for dealing with the realistic situation where combinatorics, resolution effects, and backgrounds are present. We still find good accuracies for all the masses using only the kinematic information contained in the available events. We study the accuracy of the mass determinations as a function of the available number of events and as a function of the signal to background ratio. In Sec. V, we compare results for the SUSY and UED cases and show that the masses determined are independent (to within 1 to 2 GeV) of which the model is employed. We summarize and present additional discussion in Sec. VI. Some of the material in Secs. III and IV has appeared in Ref. [7], but is included in the present article for completeness and to simplify some of the discussions.

II. CONSTRAINTS COUNTING

To begin, it is useful to perform a general counting of observables and constraints for various different configurations. We consider first the counting when only one decay chain in the event is considered at a time. We then

show the increase in constraints possible if both decay chains in each event are considered at once.

A. Single decay chain case

We begin with the chain decay $X \to aY \to abN$, where the four-momenta of the SM particles a and b are directly measured. For each event, there are four unknowns due to the unobserved four-momentum of the N. In addition, we have the three unknown masses (the same for every event) m_X, m_Y, and m_N. These are subject to $3n$ constraints coming from requiring that the X, Y, and N be on their mass shell. Thus, after n events we have $3 + 4n - 3n = 3 + n$ free parameters. No matter how many events we examine, we will not be able to obtain a discrete solution (or set of solutions) for m_X, m_Y, and m_N.

Next, consider a decay chain with three observable SM particles: $Z \to aY \to abX \to abcN$. In this case, the number of unknowns after n events is $4 + 4n$ and the number of constraints is $4n$ (the Z, Y, X, and N masses, which are the same for every event). After n events we then have $4 + 4n - 4n = 4$ free parameters, which basically correspond to the 4 unknown masses of the decaying particles. In this case, each event will determine a region in the (m_Z, m_Y, m_X, m_N) mass space and, as more events are accumulated, in an ideal world the region of mass space consistent with all events would become more and more restricted, but it will never reach a discrete point (or a set of discrete points) with any number of events. To pin down the actual mass point, one needs to use additional information by examining the end points of certain kinematic distributions.

If one considers (as in [13]) a chain with four fully measured SM particles, $A \to aZ \to abY \to abcX \to abcdN$, then we have 5 on shell particles, and after n events, we end up with $5 + 4n$ unknowns and $5n$ constraints. The number of free parameters is then $5 - n$, implying that (up to discrete ambiguities associated with a high order polynomial and ignoring combinatorics and resolution) $n = 5$ events would be sufficient to solve for the resonance masses. However, combinatorics and resolution effects will considerably complicate the situation, as already apparent from the study of [13], where they assume that m_Y, m_X, and m_N are known, leaving, in principle, $2 + 4n$ unknowns and $5n$ constraints after n events, implying that only $n = 2$ events would be needed to solve for the remaining 2 masses, m_A and m_Z. After including combinatorics and resolution, Ref. [13] needed many more than 2 events in order to get a mass determination.

The general counting procedure is apparent. If the decay chain has $N_A - 1$ on shell decaying particles and a final invisible particle, then after n events there will be $4n$ unknowns associated with the four-momenta of the invisible particle. There will also be N_A unknowns corresponding to the unknown masses of all the BSM particles. These unknowns will be subject to $N_A n$ constraints from the requirement that all the BSM particles have the same on shell masses in each event. The number of unknowns after n signal events will then be

$$N_U = N_A + 4n - N_A n = N_A + (4 - N_A)n. \quad (1)$$

For $N_A \leq 4$, a discrete solution or set of solutions is not possible regardless of how many events are available. The actual masses may still be obtained by combining additional information from the kinematic distributions for $N_A = 4$ (e.g., the squark decay case of Fig. 1). For $N_A = 5$ (e.g., the SUSY $\tilde{g} \to q\tilde{q} \to qq\tilde{\chi}_2^0 \to qql\tilde{\ell} \to qqll\tilde{\chi}_1^0$ decay chain) $n = 5$ events will give a discrete set of solutions for the masses. Still longer decay chains would require fewer events to obtain a set of discrete possibilities for the BSM particle masses. An example of a longer decay chain with $N_A = 6$ (implying that 3 events would give a discrete set of solutions for the masses) would be

$$\tilde{g} \to q\tilde{q} \to qq\tilde{\chi}_3^0 \to qqZ\tilde{\chi}_2^0 \to qqZl\tilde{\ell} \to qqZll\tilde{\chi}_1^0, \quad (2)$$

where the Z would be observed in one of its visible decay modes. By considering the full event at once through inclusion of information from both decay chains, one need not resort to such long decay chains in order to get to the point of having a discrete set of solutions for the masses.

B. Using the whole event, i.e., both decay chains

When considering the whole event at once, the constraint counting proceeds differently. Assuming that there are two invisible particles present in the final state, the number of unknowns associated with their four-momenta in all n events is $8n$. Requiring that the sum of the invisible transverse momenta equal minus the sum of the visible transverse momenta imposes $2n$ constraints on these unknowns. In addition, let us suppose that the topology is such that N_B masses are unknown (N_B includes the unknown masses of the invisible particles, which we do not require to be the same at this point in our counting). Each event will also be subject to a number N_A of on shell mass constraints, including the requirement that the two invisible particles have masses equal to their on shell (unknown) values. Then, after n events there are $N_A n$ constraints. Thus, the number of unknowns before imposing constraints is $N_B + 6n$ and the number of constraints is $N_A n$, leaving

$$N_U = N_B + (6 - N_A)n \quad (3)$$

unknowns. If we consider an event with $N_A - 2 > 4$ on shell decays and require that there be no unknowns, i.e., $N_U \leq 0$, after n_S events, we find

$$n_S \geq \frac{N_B}{N_A - 6}. \quad (4)$$

Of course, in general $N_B \leq N_A$. For symmetric chains, $N_B = N_A/2$. For chains in which only the final missing

particles are assumed to have the same mass, $N_B = N_A - 1$. The particular case we focus on in this paper, Fig. 2, corresponds $N_A = 8$ and $N_B = 4$. In this case, $n_S = 2$ events will lead to $N_U = 0$, implying a discrete set of solutions for the unknown masses of Z, Y, X, N. Were we to consider the case of two identical chains with only 2 visible particles in each and unknown masses for Y, X, N, one would have $N_A = 6$ and $N_B = 3$, leading to $N_U = 3$ (corresponding to the unknown Y, X, N masses). This is the case considered in [6]; a discrete set of solutions will never emerge with any finite number of signal events. Additional information from kinematic distributions is needed to pin down the masses. Conversely, if one goes to symmetric longer decay chains [such as that of Eq. (2)] with $N_A = 12$ and $N_B = 6$, then just one event will be sufficient to give a discrete set of mass solutions. The general problem with longer decay chains is that they are less likely to occur and harder to identify even if they occur, as the visible particles may be lost or hard to isolate. The shortest decay chains that can give rise to discrete solutions, i.e., the topology of Fig. 2, with $N_A = 8$ and $N_B = 4$ yielding $n_S = 2$, is likely to be the optimal configuration for the mass determination.

Extension to more missing particles, or quadratic constraints is straightforward, but increases the order of the equations to be solved, requiring more advanced polynomial solvers. It also requires that more masses be specified. For example, let us assume some definite topology for which each event contains 3 "missing" particles. If we do not presume any mass equalities among the total number, N_A, of decaying resonances and missing particles, then the number of unknowns after n events is $N_A + (4 \times 3 - 2)n$, where the -2 is the transverse momentum constraint setting the visible transverse momentum equal to minus the invisible transverse momentum in each event and the $4 \times 3 = 12$ just corresponds to the 4 unknown components of each of the 3 invisible particles' four-momenta. The number of mass-shell constraints is (by definition) $N_A n$. If the topology is such that $N_A > 10$ on shell masses can be reconstructed from the visible momenta and the invisible momenta and if only N_B of the N_A masses are independent, then the number of unknowns after n events is $N_B + (4 \times 3 - 2)n$ and the number of constraints is $N_A n$. Again neglecting possible relations among these masses and requiring the final number of unknowns, $N_U = N_B + 10n - N_A n$, to be 0 or negative in order to have fewer unknowns than constraints after $n = n_S$ events, one finds that

$$n_S \geq \frac{N_B}{(N_A - 10)} \tag{5}$$

events would lead to a certain set of discrete mass solutions. If all the invisible particles are the same then $N_B \leq N_A - 2$.

We have classified many possible decay chains which fall into a category such that a small handful of events could give an essentially unique mass spectrum in the absence of combinatorial and experimental resolution effects. However, generically speaking one wishes to keep the chains as short as possible while consistent with a small number of events being sufficient to yield a discrete spectrum of mass solutions. This is because (a) shorter chains are easier to isolate on an event-by-event basis and (b) combinatorial and resolution smearing of the solutions may be lessened.

III. BASIC EQUATIONS FOR THE TOPOLOGY OF FIG. 2

The topology on which we focus in this paper is that given in Fig. 2. As sketched in the previous section, this is an ideal topology for precise mass reconstruction. Assuming $m_N = m_{N'}$, $m_X = m_{X'}$, $m_Y = m_{Y'}$, $m_Z = m_{Z'}$, and denoting the four-momenta for particles $i (i = 1 \ldots 8)$ with p_i, we have

$$p_1^2 = p_2^2 (= m_N^2), \tag{6}$$

$$(p_1 + p_3)^2 = (p_2 + p_4)^2 (= m_X^2), \tag{7}$$

$$(p_1 + p_3 + p_5)^2 = (p_2 + p_4 + p_6)^2 (= m_Y^2), \tag{8}$$

$$(p_1 + p_3 + p_5 + p_7)^2 = (p_2 + p_4 + p_6 + p_8)^2 (= m_Z^2). \tag{9}$$

We assume further that the only invisible particles are particles 1 and 2, and thus have two more constraints,

$$p_1^x + p_2^x = p_{\rm miss}^x, \qquad p_1^y + p_2^y = p_{\rm miss}^y. \tag{10}$$

There are 8 unknowns in Eqs. (6)–(10), namely, the four-momenta p_1 and p_2 of the missing particles. Therefore the system is underconstrained and we cannot solve the equations. This situation changes if we add a second event with the same decay chains. Denoting the four-momenta in the second events as q_i $(i = 1 \ldots 8)$, we have 8 more unknowns, q_1 and q_2, but 10 more equations,

$$q_1^2 = q_2^2 = p_1^2, \tag{11}$$

$$(q_1 + q_3)^2 = (q_2 + q_4)^2 = (p_2 + p_4)^2, \tag{12}$$

$$(q_1 + q_3 + q_5)^2 = (q_2 + q_4 + q_6)^2 = (p_2 + p_4 + p_6)^2, \tag{13}$$

$$(q_1 + q_3 + q_5 + q_7)^2 = (q_2 + q_4 + q_6 + q_8)^2 = (p_2 + p_4 + p_6 + p_8)^2, \tag{14}$$

$$q_1^x + q_2^x = q_{\rm miss}^x, \qquad q_1^y + q_2^y = q_{\rm miss}^y. \tag{15}$$

Altogether, we have 16 unknowns and 16 equations. The system can be solved numerically and we obtain discrete solutions for p_1, p_2, q_1, and q_2 and thus the masses m_N, m_X, m_Y, m_Z. Note that the equations always have 8 com-

plex solutions, but we will keep only the real and positive-energy ones, which we simply call "solutions" in the rest of the paper. Thus, up to a certain number of discrete ambiguities we can determine the Z, Y, X, N masses by pairing any two signal events. Even a few pairs of events are typically sufficient to eliminate the discrete ambiguities due to higher order equations. However, effects such as wrong combinations and solutions, initial and final state radiation, experimental resolutions, and background events will add complications, which we address in Sec. IV.

Eqs. (6)–(15) can be easily reduced to 3 quadratic equations plus 13 linear equations,

$$p_1^2 = p_2^2 = q_1^2 = q_2^2, \tag{16}$$

$$\begin{aligned} 2p_1 \cdot p_3 + p_3^2 = 2p_2 \cdot p_4 + p_4^2 = 2q_1 \cdot q_3 + q_3^2 \\ = 2q_2 \cdot q_4 + q_4^2, \end{aligned} \tag{17}$$

$$\begin{aligned} 2(p_1 + p_3) \cdot p_5 + p_5^2 &= 2(p_2 + p_4) \cdot p_6 + p_6^2 \\ &= 2(q_1 + q_3) \cdot q_5 + q_5^2 \\ &= 2(q_2 + q_4) \cdot q_6 + q_6^2, \end{aligned} \tag{18}$$

$$\begin{aligned} 2(p_1 + p_3 + p_5) \cdot p_7 + p_7^2 &= 2(p_2 + p_4 + p_6) \cdot p_8 + p_8^2 \\ &= 2(q_1 + q_3 + q_5) \cdot q_7 + q_7^2 \\ &= 2(q_2 + q_4 + q_6) \cdot q_8 + q_8^2, \end{aligned} \tag{19}$$

$$p_1^x + p_2^x = p_{\rm miss}^x, \qquad p_1^y + p_2^y = p_{\rm miss}^y, \tag{20}$$

$$q_1^x + q_2^x = q_{\rm miss}^x, \qquad q_1^y + q_2^y = q_{\rm miss}^y, \tag{21}$$

where all but the first line are linear equations because $p_{3,4,5,6,7,8}$ and $q_{3,4,5,6,7,8}$ are all visible measured momenta. In general, the above equation system has 8 complex solutions, each of which could be real. This can be shown by calculating the Gröbner basis [29], in which the system is transformed to an 8th order univariate equation plus 15 linear equations. Since the other 15 equations are linear, it is straightforward to solve for the other 15 variables once the 8th order equation is solved. Commercial software such as MATHEMATICA uses this method. However, it consumes an intolerably long time for a single or small number of PCs. We take a simpler and faster approach which is described in detail in the appendixes. In our method, instead of ending up with an 8th order equation, we obtain a 9th order univariate polynomial equation and therefore introduce a fake solution in addition to the true solutions. The 9th order univariate polynomial equation is numerically solved using the algorithm TOMS/493 [30]. The fake solution can be easily eliminated by substituting back all solutions in the original equations.

IV. APPLICATIONS

A. SUSY point SPS1a

For illustration and easy comparison to the literature, we apply our method for the SUSY point, SPS1a [31], although many of the discussions below apply for generic cases. For SPS1a, the particles corresponding to N, X, Y, Z are $\tilde{\chi}_1^0$, $\tilde{\ell}_R(\ell = e/\mu)$, $\tilde{\chi}_2^0$, $\tilde{q}_L(q = d, u, s, c)$, respectively. The masses are

$$\begin{aligned} m_N &= 97.4 \text{ GeV}, \qquad m_X = 142.5 \text{ GeV}, \\ m_Y &= 180.3 \text{ GeV}, \qquad m_Z = 564.8/570.8 \text{ GeV}, \end{aligned} \tag{22}$$

with the final two numbers corresponding to up/down type squarks, respectively. Since $m_{\tilde{\tau}} \neq m_{\tilde{e},\tilde{\mu}}$, the $\ell = \tau$ case is an important background. We generate events with PYTHIA 6.4 [32].

We first consider the ideal case: no background events, all visible momenta measured exactly, all intermediate particles on shell and each visible particle associated with the correct decay chain and position in the decay chain. We also restrict the squarks to be up-type only. In this case, we can solve for the masses exactly by pairing any two events. The only complication comes from there being 8 complex solutions for the system of equations, of which more than one can be real and positive. Of course, the wrong solutions are different from pair to pair, but the correct solution is common. The mass distributions for the ideal case with 100 events (no kinematic cuts applied) are shown in Fig. 3. Note the logarithmic scale. As expected, we observe δ-function-like mass peaks on top of small backgrounds coming from wrong solutions. On average, there are about 2 solutions per pair of events.

The δ functions in the mass distributions arise only when exactly correct momenta are input into the equations

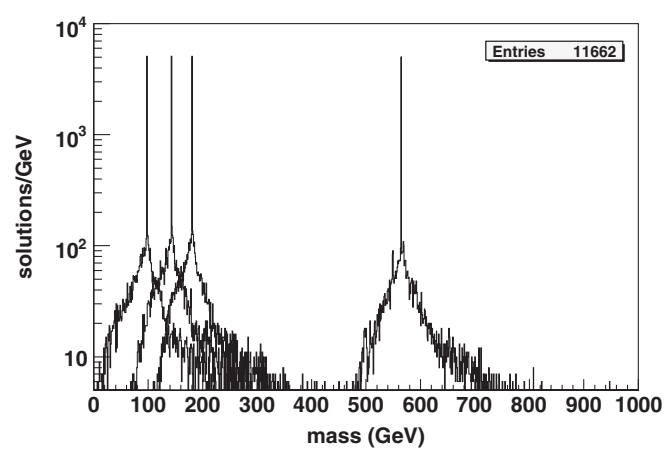

FIG. 3. We plot the number of mass solutions (in 1 GeV bins—the same binning is used for the other plots) vs mass in the ideal case. All possible pairs for 100 events are included. Signal events only.

we solve. To be experimentally realistic, we now include the following.

(1) *Wrong combinations.* For a given event a "combination" is a particular assignment of the jets and leptons to the external legs of Fig. 2. For each event, there is only one correct combination (excluding $1357 \leftrightarrow 2468$ symmetry). Assuming that we can identify the two jets that correspond to the two quarks, we have 8 (16) possible combinations for the $2\mu 2e$ (4μ or $4e$) channel. The total number of combinations for a pair of events is the product of the two, i.e., 64, 128, or 256. Adding the wrong combination pairings for the ideal case yields the mass distributions of Fig. 4. Compared to Fig. 3, there are 16 times more (wrong) solutions, but the δ-function-like mass peaks remain evident.

(2) *Finite widths.* For SPS1a, the widths of the intermediate particles are roughly 5 GeV, 20 MeV, and 200 MeV for $\tilde{q}_L$, $\tilde{\chi}_2^0$, and $\tilde{\ell}_R$. Thus, the widths are quite small in comparison to the corresponding masses.

(3) *Mass splitting between flavors.* The masses for up and down type squarks have a small difference of 6 GeV. Since it is impossible to determine flavors for the light jets, the mass determined should be viewed as the average value of the two squarks (weighted by the parton distribution functions).

(4) *Initial/final state radiation.* These two types of radiation not only smear the visible particles' momenta, but also provide a source for extra jets in the events. We will apply a p_T cut to get rid of soft jets.

(5) *Extra hard particles in the signal events.* In SPS1a, many of the squarks come from gluino decay ($\tilde{g} \to q\tilde{q}_L$), which yields another hard q in the event. Fortunately, for SPS1a $m_{\tilde{g}} - m_{\tilde{q}_L} = 40$ GeV is much smaller than $m_{\tilde{q}_L} - m_{\tilde{\chi}_2^0} = 380$ GeV. Therefore, the q from squark decay is usually much more energetic than the q from $\tilde{g}$ decay. We select the two jets with highest p_T in each event after cuts. Experimentally, one would want to justify this choice by examining the jet multiplicity to ensure that this analysis is dominated by 2-jet events, and not 3 or 4 jet events.

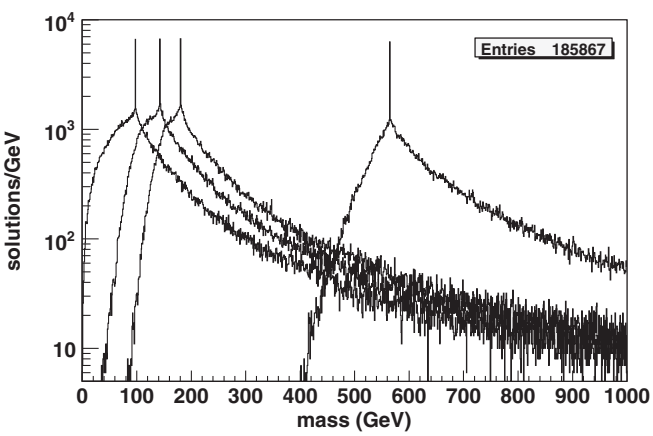

FIG. 4. Number of mass solutions vs mass after including all combination pairings for 100 events. Signal events only, with only combinatoric ambiguities included.

(6) *Background events.* The SM backgrounds are negligible for this signal in SPS1a. There are a few significant backgrounds from other SUSY processes:

(a) $\tilde{q}_L \to q\tilde{\chi}_2^0 \to q\tau\tilde{\tau} \to q\tau\tau\tilde{\chi}_1^0$ for one or both decay chains, with all τ's decaying leptonically. Indeed, $\tilde{\chi}_2^0 \to \tau\tilde{\tau}$ has the largest partial width, being 14 times that of $\tilde{\chi}_2^0 \to \mu\tilde{\mu}$. However, to be included in our selection the two τ's in one decay chain must both decay to leptons with the same flavor, which reduces the ratio. A cut on lepton p_T also helps to reduce this background, since leptons from τ decays are softer. Experimentally, one should perform a separate search for hadronically decaying tau's or nonidentical-flavor lepton decay chains to explicitly measure this background.

(b) Processes containing a pair of sbottoms, which have different masses from the first two generations. Since b jets are distinguishable, a separate analysis should be performed to determine the b squark masses. However, this presents a background to the light squark search since b-tagging efficiency is only about 50% at high p_T.

(c) Processes that contain a pair of $\tilde{\chi}_2^0$'s, not both coming from squark decays. For these events to fake signal events, extra jets need to come from initial and/or final state radiation or other particle decays. For example, direct $\tilde{\chi}_2^0$ pair production or $\tilde{\chi}_2^0 + \tilde{g}$ production. These are electroweak processes, but, since $\tilde{\chi}_2^0$ has a much smaller mass than squarks, the cross section is not negligible. In our SPS1a analysis, the large jet p_T cut reduces this kind of background due to the small $m_{\tilde{g}} - m_{\tilde{q}_L}$.

(7) *Experimental resolutions.* In order to estimate this experimental effect at the LHC, events in both signal and the aforementioned SUSY backgrounds are further processed with pretty good simulation (PGS) [33]. Note that in [7], we used ATLFAST [42] for the detector simulation. Compared with ATLFAST, PGS has more stringent lepton isolation cuts, therefore we obtain fewer events. Nevertheless, as shown below, the results turn out to be similar. All objects including jets, isolated leptons, and missing p_T are taken directly from PGS.

The cuts used to isolate the signal are:

(I) Four isolated leptons with $p_T > 10$ GeV, $|\eta| < 2.5$, and matching flavors and charges consistent with our assumed $\tilde{\chi}_2^0 \to \tilde{\ell} \to \tilde{\chi}_1^0$ decay.

(II) No b jets and ≥ 2 jets with $p_T > 100$ GeV, $|\eta| < 2.5$. The 2 highest-p_T jets are taken to be particles 7 and 8.

(III) Missing $p_T > 50$ GeV.

For a data sample with 300 fb^{-1} integrated luminosity, there are about 620 events left after the above cuts, out of which about 420 are signal events. After taking all possible pairs for all possible combinations and solving for the masses, we obtain the mass distributions in Fig. 5.

From Fig. 5, we see that the mass peaks are smeared but still present around the input masses. The analytical formula for the distributions are unknown, so we estimate the masses by reading the peak positions. To minimize the effect from statistical fluctuations, we fit each distribution using a sum of a Gaussian plus a (single) quadratic polynomial and taking the maximum positions of the fitted peaks as the estimated masses. We will use this function as the "standard fit" throughout this article. The fitted range is restricted to be above the half height. The fitted curves are superimposed on the mass distributions in Fig. 5, which yields $\{78.4, 134.2, 181.5, 553.9\}$ GeV for the masses. Averaging over 20 different data samples, we find

$$m_N = 76.7 \pm 2.0 \text{ GeV}, \qquad m_X = 134.6 \pm 2.2 \text{ GeV},$$
$$m_Y = 178.9 \pm 3.8 \text{ GeV}, \qquad m_Z = 561.6 \pm 5.4 \text{ GeV}. \tag{23}$$

The statistical uncertainties are very small, but there exist biases, especially for the two light masses. In practice, we can always correct the biases by comparing real data with Monte Carlo data. Nevertheless, we would like to reduce the biases as much as possible using data only. In some cases, the biases can be very large and it is essential to reduce them before comparing with Monte Carlo data—we will see an example later.

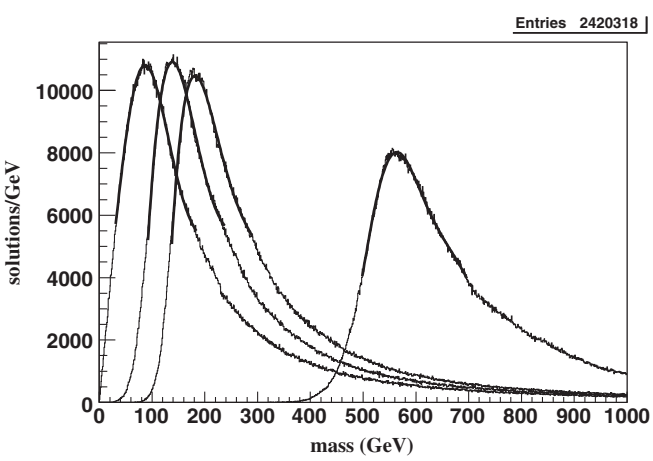

FIG. 5. Mass solutions with all effects 1–7 included and after cuts I–III for the SPS1a SUSY model and $L = 300$ fb^{-1}. All effects incorporated, including backgrounds.

The combinatorial background is an especially important source of bias since it yields peaked mass distributions that are not symmetrically distributed around the true masses, as can be seen from Fig. 4. This will introduce biases that survive even after smearing. Therefore, we concentrate on reducing wrong solutions.

First, we reduce the number of wrong combinations by the following procedure. For each combination choice, c, for a given event, i ($i = 1, N_{\text{evt}}$), we count the number, $N_{\text{pair}}(c, i)$, of events that can pair with it (for some combination choice for the 2nd events) and give us solutions. We repeat this for every combination choice for every event. Neglecting effects 2–7, $N_{\text{pair}}(c, i) = N_{\text{evt}} - 1$ if c is the correct combination for event i. After including backgrounds and smearing, $N_{\text{pair}}(c, i) < N_{\text{evt}} - 1$, but the correct combinations still have statistically larger $N_{\text{pair}}(c, i)$ than the wrong combinations. The frequency with which various values of $N_{\text{pair}}(c, i)$ occur is shown as a function of $N_{\text{pair}}(c, i)$ in Fig. 6.

To enhance the likelihood that a particular choice of c, i corresponds to a correct solution, we cut on $N_{\text{pair}}(c, i)$. For the SPS1a model point, if $N_{\text{pair}}(c, i) \leq 0.75 N_{\text{evt}}$ we discard the combination choice, c, for event i. If all possible c choices for event i fail this criterion, then we discard event i altogether (implying a smaller N_{evt} for the next analysis cycle). We then repeat the above procedure for the remaining events until no combinations can be removed. After this, for the example data sample, the number of events is reduced from 622 (424 signal + 198 background) to 430

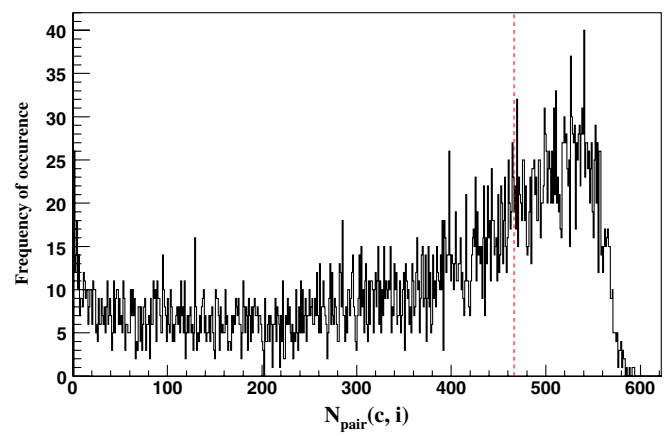

FIG. 6 (color online). For each event, i, and each combination, c, associated with that event, we count the number, $N_{\text{pair}}(c, i)$, of events that can pair with it and give at least one solution. The plot shows the frequency of occurrence of different values of $N_{\text{pair}}(c, i)$. All effects are incorporated, including backgrounds. The plot is for the SPS1a case (for which the total number of signal+background events is 620 for $L = 300$ fb^{-1}). In the bias-reduction procedure, any choices of c, i yielding $N_{\text{pair}}(c, i)$ to the left of the red (dashed) line (corresponding to 75% of the total number of events) are discarded.

(322 signal + 108 background), and the average number of combinations per event changes from 11 to 4.

Second, we increase the significance of the true solution by weighting each surviving pair of events by $1/n$ where n is the number of solutions for the given pair (using only the combination choices that have survived the previous cuts). This causes each pair (and therefore each event) to have equal weight in our histograms. Without this weighting, a pair with multiple solutions has more weight than a pair with a single solution, even though at most one solution would be correct for each pair.

Finally, we exploit the fact that wrong solutions and backgrounds are much less likely to yield M_N, M_X, M_Y, and M_Z values that are all simultaneously close to their true values. We plot the $1/n$-weighted number of solutions as a function of the three mass differences (Fig. 7). We define mass difference windows by $0.6 \times$ (peak height) and keep only those solutions for which *all three* mass differences fall within the mass difference windows. The surviving solutions are plotted (without the $1/n$ weighting) in Fig. 8. Compared with Fig. 5, the mass peaks are narrower, more symmetric and the fitted values are less biased. The fitted masses are $\{93.9, 140.3, 180.5, 559.2\}$ GeV. Repeating the procedure for 20 data sets, we find

$$m_N = 93.8 \pm 3.9 \text{ GeV}, \qquad m_X = 138.4 \pm 4.5 \text{ GeV},$$
$$m_Y = 178.7 \pm 4.6 \text{ GeV}, \qquad m_Z = 559.5 \pm 5.4 \text{ GeV}, \tag{24}$$

to be compared to the input masses of Eq. (22). Thus, the biases are reduced without significantly increasing the statistical errors.

Thus, we have shown that the masses can be measured with high precision for a few hundred events in the four-fermion decay channel. In the case of the SPS1a point, the number of events employed above corresponds to a high integrated luminosity, $L \sim 300$ fb^{-1}. The reason that such a high luminosity is required in the case of the SPS1a scenario is that the branching ratio for $\tilde{\chi}_2^0 \to \tilde{\tau}\tau$ is 14 times that for $\tilde{\chi}_2^0 \to \tilde{\mu}\mu$ or $\tilde{\chi}_2^0 \to \tilde{e}e$. More generally, the integrated luminosity needed to get a few hundred events is highly dependent on the branching ratios for the various SUSY particle decays in the model. For example, if one takes the SPS1a masses but requires that $\tilde{\chi}_2^0$ decays equally to the three lepton flavors instead, the same number of signal events as employed above can be obtained with just 10 fb^{-1} of data.

Although the errors in the mass determinations depend upon the number of events, our method is quite robust in that we get decent mass determinations even with a small number of events. In Fig. 9, the mass distributions for 50 events are shown, with evident mass peaks. By repeating our procedure for multiple data sets of a given size, we obtain the errors as functions of the number of events. Figure 10 shows the error for the $\tilde{\chi}_1^0$ mass determination

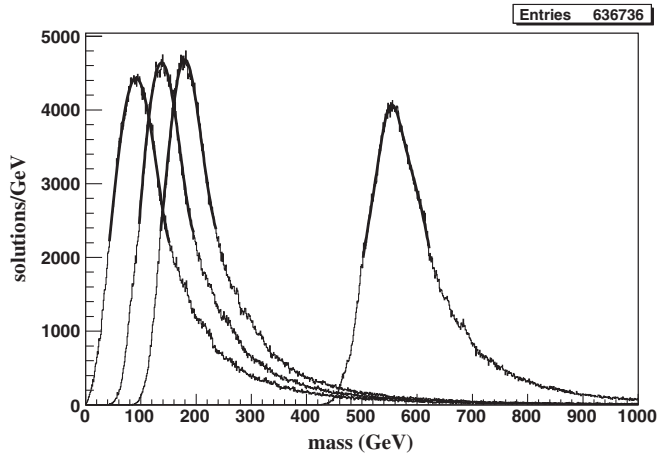

FIG. 8. Final mass distributions after the bias-reduction procedure for the SPS1a SUSY model and $L = 300$ fb^{-1}. All effects incorporated, including backgrounds.

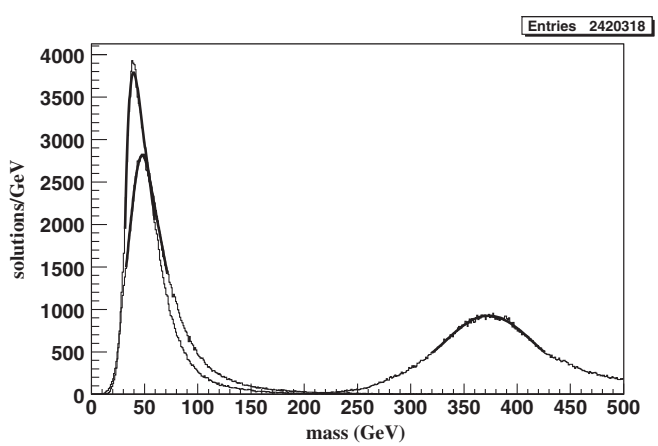

FIG. 7. SPS1a, $L = 300$ fb^{-1} mass difference distributions. All effects incorporated, including backgrounds.

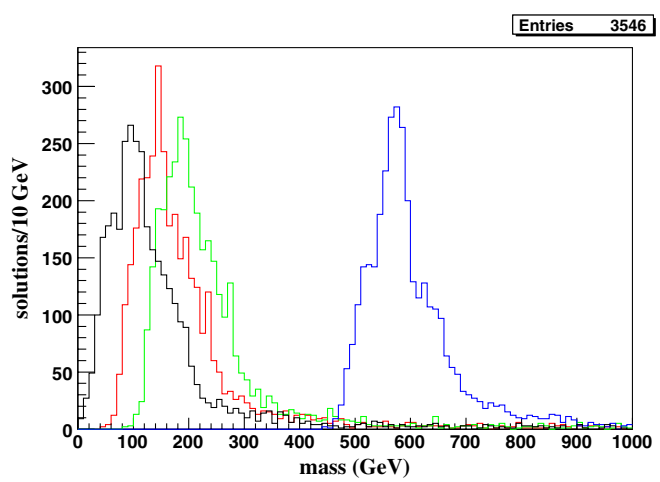

FIG. 9 (color online). Mass distributions for 50 events for SPS1a.

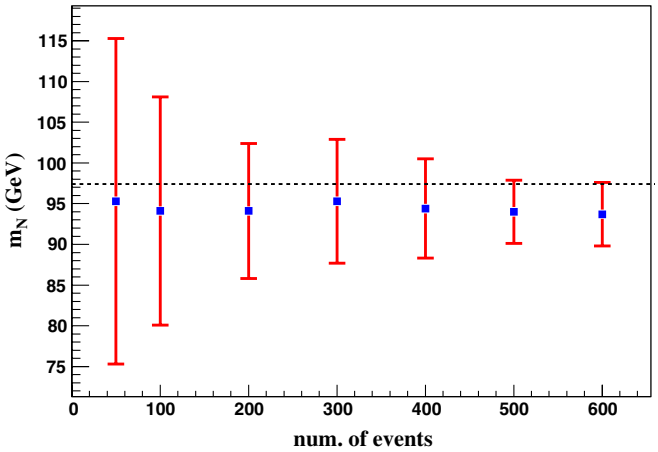

FIG. 10 (color online). Error bars for m_N as a function of the number of background + signal events for SPS1a. All effects and procedures included.

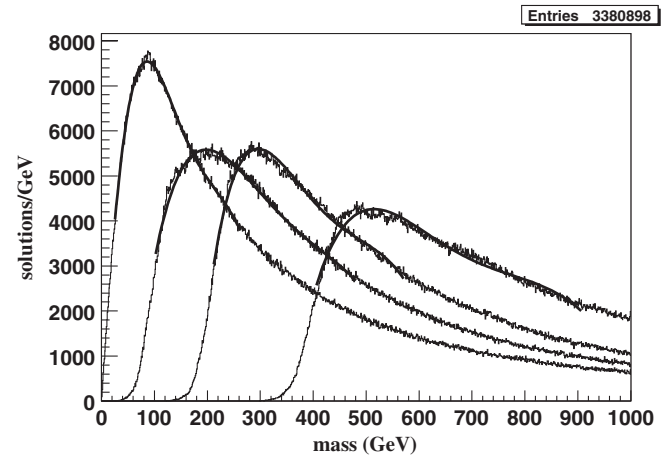

FIG. 11. Final mass distributions before the bias-reduction procedure for the point no. 1 SUSY model and $L = 100\ \mathrm{fb}^{-1}$. All effects incorporated, including backgrounds.

as a function of the number of signal+background events. Note that the central value for multiple data sets of the given size is quite insensitive to the data set size, but, of course, the possible deviation from this central value for any one data set increases as the data set size decreases.

B. SUSY Point No. 1

We have applied our method to other mass points to show its reliability. We quote here results for "point no. 1" defined in Ref. [6] with the following masses: $\{85.3, 128.4, 246.6, 431.1/438.6\}$ GeV. For $100\ \mathrm{fb}^{-1}$ data, we have about 800 events (770 signal events) after the same prebias-reduction cuts. The resulting mass plot before performing bias-reduction cuts is that given in Fig. 11. From Fig. 11, we see that the mass peaks are very broad and we get more than 50 GeV biases if we use the positions of the maxima as the true mass values. We then repeat the same bias-reduction procedure as for SPS1a except that we employ a looser cut on $N_{\mathrm{pair}}(c, i)$ than for the SPS1a case, despite the fact that there are more signal events for point no. 1. We require $N_{\mathrm{pair}}(c, i) > 0.6N_{\mathrm{evt}}$. The reason is that, unlike the SPS1a case, the gluino mass in point no. 1 (524 GeV) is significantly larger than the squark mass. Therefore the quark jet from gluino decay is often misidentified as the jet from squark decay, which reduces the chance to obtain solutions for a pair of events.[1] In practice, there is not a universal "best" cut on $N_{\mathrm{pair}}(c, i)$: a more stringent cut leads to smaller biases but larger statistical uncertainties. After the bias-reduction procedure using $N_{\mathrm{pair}}(c, i) > 0.6N_{\mathrm{evt}}$ we are left with 560 events (550 signal events). The mass distributions are shown in Fig. 12. They are much narrower and the biases are considerably reduced. After following the bias-reduction procedure and using 20 data samples to estimate the errors, we obtain $m_N = 82.8 \pm 3.2$ GeV, $m_X = 127.9 \pm 3.0$ GeV, $m_Y = 245.7 \pm 3.4$ GeV, $m_Z = 436.4 \pm 5.4$ GeV. The central values are in quite close agreement with the input masses except for m_N which comes out a bit low.

C. Comments and comparisons

We emphasize that the remaining biases in the above mass determinations can be removed by finding those input masses that yield the observed output masses after processing Monte Carlo generated data through our procedures. In this way, very accurate central mass values are obtained with the indicated statistical errors.

The above results for the N, Y, and X masses for the SPS1a point and point no. 1 can be compared to those obtained following the very different procedure of Ref. [6].

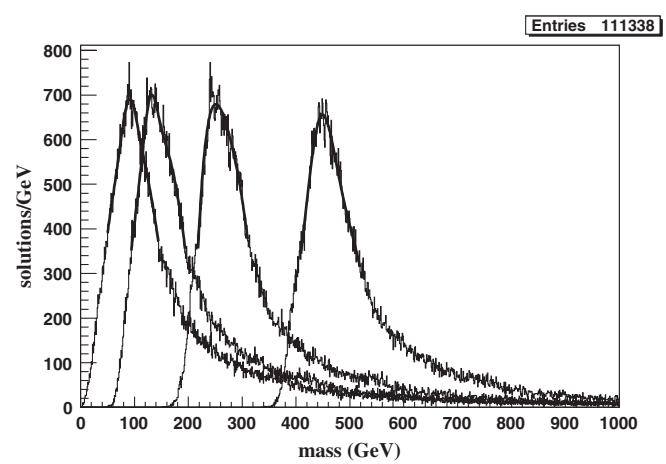

FIG. 12. Final mass distributions after the bias-reduction procedure for the point no. 1 SUSY model and $L = 100\ \mathrm{fb}^{-1}$. All effects incorporated, including backgrounds.

[1] It is possible to improve the results by considering all high p_T jets as candidates for the quarks from squark decays, instead of simply choosing the two highest p_T jets.

There, only the $X \to Y \to N$ parts of the two decay chains were employed and we used only 4μ events. For the SPS1a model point we obtained $m_N = 98 \pm 9$ GeV, $m_Y = 187 \pm 10$ GeV, and $m_X = 151 \pm 10$ GeV. And, for point no. 1 we found $m_N = 86.2 \pm 4.3$ GeV, $m_X = 130.4 \pm 4.3$ GeV, and $m_Y = 252.2 \pm 4.3$ GeV. Including the $4e$ and $2\mu 2e$ channels will reduce the indicated errors by a factor of ~ 2. The procedure of [6] can thus be used to verify the results for m_N, m_X, and m_Y from the present procedure and possibly the two can be combined to obtain smaller errors than from either one, with m_Z determined by the procedure of this paper.

We also compare the results for SPS1a with those given in Ref. [7] where exactly the same procedure and cuts are applied to the same model point. The difference is that we used ATLFAST for the detector simulation in Ref. [7] while we have switched to PGS in the current paper. The PGS simulation has more stringent lepton isolation cuts and therefore we obtain fewer events in the present analysis (620 *vs* 1050). In Ref. [7], we obtained

$$m_N = 76.7 \pm 1.4 \text{ GeV}, \qquad m_X = 135.4 \pm 1.5 \text{ GeV},$$
$$m_Y = 182.2 \pm 1.8 \text{ GeV}, \qquad m_Z = 564.4 \pm 2.5 \text{ GeV}$$

before the bias-reduction procedure and

$$m_N = 94.1 \pm 2.8 \text{ GeV}, \qquad m_X = 138.8 \pm 2.8 \text{ GeV},$$
$$m_Y = 179.0 \pm 3.0 \text{ GeV}, \qquad m_Z = 561.5 \pm 4.1 \text{ GeV}$$

after. Comparing the above numbers with those in Eqs. (23) and (24), we see that the masses obtained using PGS simulation have larger statistical errors, in accord with the smaller number of events. On the other hand, the central values agree well, indicating that the bias-reduction procedure affects the mass peaks in a nearly model-independent manner. We view this as evidence of the robustness of our method.

D. Removing biases using a dilepton edge cut

As we have discussed, the primary source of biases is the detector smearing of wrong solutions, especially those associated with wrong combinations. It will be possible to efficiently eliminate many of these wrong solutions if there is a significant structure associated with correct solutions in one or more distributions constructed from the visible particles' momenta. In the SUSY examples we consider here, such a structure is especially apparent in the distribution of $m_{\ell^+\ell^-}$, where $\ell = e, \mu$ (same flavor pairs only). The advantage of using only leptons is the much better resolution for the lepton momentum measurements. Ignoring resolution smearing, kinematics predicts that correct combinations should have

$$(m_{\ell^+\ell^-}^{\text{edge}})^2 = \frac{(m_{Y0}^2 - m_{X0}^2)(m_{X0}^2 - m_{N0}^2)}{m_{X0}^2}, \qquad (25)$$

where m_{Y0}, m_{X0}, and m_{N0} are the input masses. Note that there are many more dilepton events than four lepton events since dileptons require only a single decay chain of Fig. 1. The plot of $m_{\ell^+\ell^-}$ values for all solutions coming from 600 SPS1a events (after PGS smearing and general cuts, but before applying the bias-reduction procedure) is shown in Fig. 13. The edge at the predicted value of 80 GeV is apparent and its location can be determined quite accurately from the data.

Before employing the bias-reduction procedure, we apply a cut on the m_Y, m_X, m_N values obtained for a given solution of

$$\left|\sqrt{(m_Y^2 - m_X^2)(m_X^2 - m_N^2)/m_X^2} - m_{\ell^+\ell^-}^{\text{edge}}\right| < 20 \text{ GeV}, \quad (26)$$

where we have purposely employed a rather loose cut so as to not lose statistics and to take into account smearing of the input X, Y, N masses that will be present even for a correct combination, as well as the small error associated with determining the edge location experimentally. We *then* use the same bias-reduction procedure as discussed earlier using a sequence of choices for the cut f_{cut} defined by retaining only combinations with $N_{\text{pair}}(c, i) > f_{\text{cut}} N_{\text{evt}}$ in Table I.

We clearly observe that the dilepton edge cut has greatly reduced the bias in comparison to results obtained without the dilepton edge cut. Further, for the larger values of f_{cut}, the mass peak locations are not biased at all (within statistics) in comparison to the input masses of Eq. (22). This occurs because many of the wrong solutions have been eliminated. For example, after the dilepton edge cut and after employing the bias-reduction procedure using $f_{\text{cut}} = 0.75$, about 160 events are retained on average and the average number of solutions for the remaining pairs

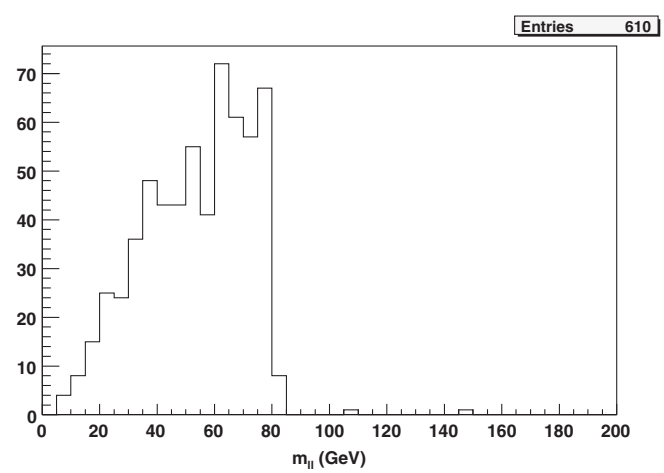

FIG. 13. We plot the number of events as a function of $m_{\ell^+\ell^-}$ for 600 SPS1a events in 5 GeV bins (after PGS smearing and general cuts, but before the bias-reduction procedure). Only events containing two muons and two electrons with opposite charges are used to avoid ambiguity, each of which contributes two entries to the histogram. The edge at 80 GeV is apparent.

TABLE I. Peak locations for various values of $f_{\rm cut}$ with and without the dilepton edge cut. Errors were determined using 20 distinct data sets.

	With dilepton edge cut				Without dilepton edge cut			
$f_{\rm cut}$	0.60	0.65	0.70	0.75	0.60	0.65	0.70	0.75
m_N (GeV)	93.0 ± 3.7	96.1 ± 3.9	97.5 ± 4.3	97.9 ± 4.9	85.6 ± 2.3	88.1 ± 3.5	90.7 ± 3.8	93.8 ± 3.9
m_X (GeV)	138.9 ± 3.9	141.4 ± 4.6	143.7 ± 4.6	144.3 ± 4.0	131.5 ± 2.7	133.9 ± 3.6	135.9 ± 4.3	138.4 ± 4.5
m_Y (GeV)	176.5 ± 3.8	178.8 ± 4.6	180.8 ± 5.1	181.5 ± 5.3	172.8 ± 2.8	174.8 ± 3.8	176.6 ± 4.4	178.7 ± 4.6
m_Z (GeV)	557.8 ± 4.4	559.9 ± 4.5	563.2 ± 5.0	565.6 ± 6.2	555.8 ± 5.2	557.2 ± 5.5	557.8 ± 5.1	559.5 ± 5.4

formed from these surviving events is only about 1.2. In other words, the dilepton edge cut is highly effective in removing wrong combinations. Errors for the peak mass values are, of course, slightly larger when a dilepton edge cut is imposed, implying that ultimately the best mass determinations may be those obtained using Monte Carlo determination of the bias corrections that should be applied to mass peak values obtained without the dilepton edge cut. Nonetheless, doing the analysis with a dilepton edge cut will provide a very important cross check of the bias determination.

As a final note, we observe that in the case of SPS1a there are also incorrect solutions coming from chains containing a pair of leptonically decaying τ's. Many, but not all of these wrong solutions are also be eliminated by the dilepton edge cut. The remaining background events contain mostly those events for which one chain has $\ell = e, \mu$ while the other has a pair of leptonically decaying τ's, since sometimes such events will give a solution with nearly correct mass values.

E. More on backgrounds

Because the SM background can be efficiently reduced by applying a large missing p_T cut, the most difficult backgrounds usually come from other SUSY processes that contain the same final state particles. In the above examples, we have already encountered such backgrounds. In the SPS1a case, the backgrounds are dominated by events that contain leptonically decaying τ's. For SUSY point no. 1, although most events are originally signal events, in many cases the jets from squark decays are not correctly identified, in which case these events should be viewed as background events. In both examples, the background events are closely related to the signal events and therefore also carry some information about the masses. It is also interesting to study the effects of background events of a completely different origin, and test the stability of our mass determination method.

In order to explore the issues that arise, we will perform an analysis in which we consider the SPS1a events as signal events, fixing the number of events to 600 (including the intrinsic background of SPS1a). For a possible SUSY background to the SPS1a events we employ events of the above SUSY point no. 1 as "background" events, varying the ratio of point no. 1 events with respect to the SPS1a events. Since this is only for illustration, we are not concerned with how this could happen in a specific SUSY model. The existence of two different types of events is immediately seen from the dilepton invariant mass distribution (Fig. 14), where two different edges are evident. The position of the edges are given by Eq. (25). In this case, the two signals give two different m_{Y0}, m_{X0}, and m_{N0} input mass sets. Again, there are many more dilepton events than four lepton events since a dilepton only requires a single decay chain of Fig. 1. From Fig. 14 we see that the position of the edge associated with the SPS1a signal can be determined quite precisely even when the background to signal ratio is of order one. Consequently, one can try to combine the edge location measurement with information from double chain events (see, also, Ref. [22]).

First, we repeat our fitting procedure on the mixed events without using the dilepton edge information. Since there are more background events, we cannot use the fixed $N_{\rm pair}(c, i) > 0.75 N_{\rm evt}$ cut as before. This is because $N_{\rm evt}$ now refers to the total number of events from both SPS1a and SUSY point no. 1 so that such a cut would amount to a much stronger $f_{\rm cut}$ value for the SPS1a signal on its own. Instead, we choose the $N_{\rm pair}(c, i)$ cut so that 60% of all the

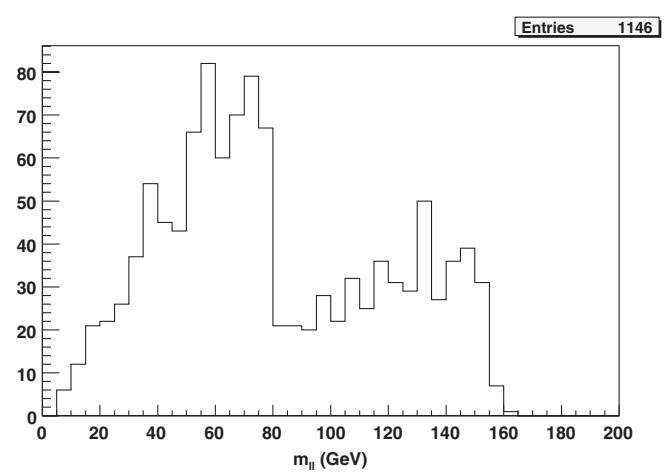

FIG. 14. Dilepton invariant mass distribution for 600 SPS1a events together with 600 point no. 1 events (after PGS smearing and general cuts, but before the bias-reduction procedure). Only events containing two muons and two electrons of opposite charge are used to avoid ambiguity. The two edges at 80 GeV and 157 GeV correspond to SPS1a and point no. 1, respectively.

events are left after the bias-reduction procedure. The corresponding value of $f_{\rm cut}$ varies according to the amount of SUSY point no. 1 background included. For 600 SPS1a events combined with 600 point no. 1 background events the 60% survival fraction corresponds to using $f_{\rm cut} \sim 0.58$ in the bias-reduction procedure. For the SPS1a signal alone, the corresponding $f_{\rm cut}$ value is somewhat larger. The measured m_N is shown in Fig. 15 as a function of the number of background events (after PGS smearing, but before the bias-reduction procedure). From Fig. 15, we see that the mass determination is not accurate when the background/signal ratio is high. However, as long as the number of background events after general cuts is less than about half the number of signal events, the bias-reduction procedure is effective in removing background events while retaining signal events, and the mass determination is quite good. Of course, in practice we will not know *a priori* what the number of background events is relative to the number of signal events and therefore we would need additional input in order to know if the fitted masses from mass peaks are reliable estimates for the true masses. In the present case, the dilepton mass plot of Fig. 14 would clearly have indicated the presence of two different classes of events and we would therefore know that it would help to use an additional cut to reduce the background class. We again employ the simple cut of Eq. (26), where in order to isolate the SPS1a component of the combined events we would employ $m_{\ell^+\ell^-}^{\rm edge} = 80$ GeV. Again note that this is a loose cut that does not require a precise knowledge of the edge position from the single chain events. Applying the cut in Eq. (26) on data sets with 600 signal + 600 background events (after general cuts, but before the bias-reduction procedure) and repeating the fit procedure, we obtain the masses

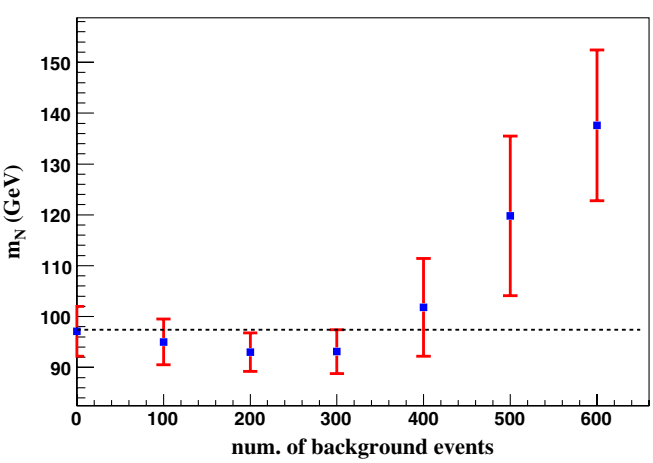

FIG. 15 (color online). The measured m_N as a function of the number of background events. The number of signal (SPS1a) events is fixed to 600. In both cases, these are the event numbers after general cuts but before the bias-reduction procedure.

$$m_N = 96.2 \pm 3.6 \text{ GeV}, \qquad m_X = 141.3 \pm 4.1 \text{ GeV},$$
$$m_Y = 178.4 \pm 4.1 \text{ GeV}, \qquad m_Z = 558.5 \pm 4.6 \text{ GeV}, \tag{27}$$

where errors were determined using 20 distinct data sets. The same $f_{\rm cut} = 0.58$ was used as in the case without dilepton edge cut. With the dilepton edge cut, we are left with averagely 363 SPS1a events and 8 SUSY point no. 1 events in the mass distributions used to get the mass peak locations of Eq. (27). Thus, we effectively isolated the signal of interest by employing the dilepton edge cut. We have also obtained central mass values that have almost no bias relative to the input SPS1a masses. This is because the $f_{\rm cut}$ for the SPS1a signal alone that yields the number of events (~ 360) after bias reduction is close to 0.65, which according to Table I should give a nearly unbiased mass determination.

V. COMPARISON OF SUSY AND UED

In this section, we address the question of whether or not the mass determinations (and the accuracy thereof) are sensitive to the model employed by comparing results for the SPS1a point to a UED model chosen so that the masses are exactly the same as for SPS1a (the corresponding decay chain in UED is KK quark $\to$ KK $Z \to$ KK lepton $\to$ KK photon). We have also adjusted the squarks/Kaluza-Klein (KK) quarks of different flavors to have the same mass (564.8 GeV, as for $\tilde{u}_L$ in SPS1a) and chosen squark/KK quark pair production as the only process (i.e., no gluinos/KK gluons). The finite widths of the involved particles are also turned off. Both SUSY and UED events are simulated with Herwig + + [34], with spin correlations included and confirmed by comparing with Ref. [35]. In Figs. 16 and 17, for SUSY and UED, respectively, we plot the mass distributions after employing identical smearing (PGS), general cuts and the bias-reduction procedure [using $N_{\rm pair}(c, i) > 0.75 N_{\rm evt}$]. The size of the event samples for SUSY and UED are set by requiring that both samples contain 400 events after PGS smearing and general cuts, but before the bias-reduction procedure. Visually, it is clear that the peaks are in very similar locations.

After employing our standard fitting techniques, we obtain masses of

$$m_N = 90.3 \pm 3.0 \text{ GeV}, \qquad m_X = 135.4 \pm 3.1 \text{ GeV},$$
$$m_Y = 176.0 \pm 2.9 \text{ GeV}, \qquad m_Z = 551.2 \pm 5.2 \text{ GeV}$$

for the actual SUSY SPS1a model and masses of

$$m_N = 89.5 \pm 3.5 \text{ GeV}, \qquad m_X = 134.8 \pm 3.5 \text{ GeV},$$
$$m_Y = 176.2 \pm 3.7 \text{ GeV}, \qquad m_Z = 548.7 \pm 5.8 \text{ GeV}$$

for the UED model with SPS1a masses. We note that the measured masses still have some biases of order of a few GeV (for the three smaller masses) up to more than

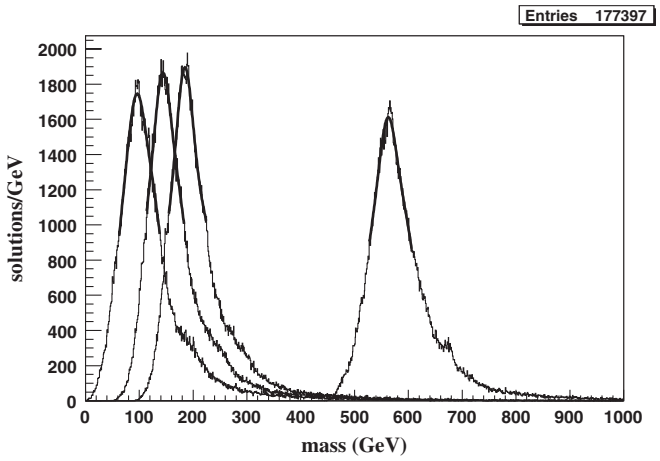

FIG. 16. Final mass distributions for signal events only after employing the bias-reduction procedure for the SPS1a mass choices in the context of the SUSY model and using 400 signal events after PGS smearing and general cuts but before bias reduction. All combinations and solutions are included, but backgrounds associated with the SUSY model are not included.

10 GeV. But the biases are very similar for the two models with different spins, indicating that the major part of these biases can be removed by comparing with Monte Carlo even before the underlying model is determined. Using the technique of repeating the procedure 20 times, we find that the SUSY model statistical errors in this case are somewhat smaller than quoted earlier in Eqs. (24) because we have not included any background from other SUSY processes. The biases are also slightly different from Eqs. (24) where the backgrounds are included. The UED model errors would also be increased by including backgrounds coming from other UED processes. In other words, the masses are very well determined (to within a few GeV) by our purely kinematic procedures, but the errors are mildly model dependent because of variation in the nature and magnitude of the new-physics-model backgrounds. The contribution from the backgrounds can often be inferred from real data and subtracted. For example, the SPS1a background is dominated by the chain decays that yield staus, whose existence and production rate can be determined from hadronically decayed taus.

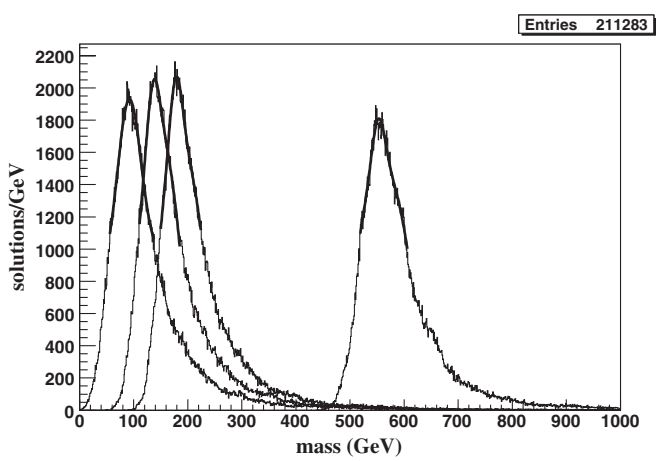

FIG. 17. Final mass distributions for signal events only after the bias-reduction procedure for the SPS1a mass choices, but in the case of the UED model. The UED event sample is scaled so that there are 400 signal events after PGS smearing and general cuts, but before bias reduction. All combinatorics and solutions are included, but backgrounds associated with the UED model are not included.

VI. DISCUSSION AND CONCLUSIONS

One important question is whether it is better to use one-chain or two-chain techniques. Our point of view is that one should use all available kinematic information, regardless of whether it is from one-chain or two-chain events. On the one hand, due to the fact that not all events contain two identical chains, one often obtains more one-chain events of a certain type than events with two identical decay chains of that type. However, if one considers only one chain at a time, information, in particular, that related to the measured missing transverse momentum, is always lost. The consequence is that either one cannot solve directly for all involved masses for a given length of decay chain, or one must employ longer decay chains, in which case the method becomes very complicated. For example, in the one-chain case one needs to employ decay chains with four visible particles (vs three visible particles in the two-chain case) and, in addition, one needs to combine five events to obtain discrete solutions for the unknown masses. Even assuming that such events do exist, there are more wrong combinations and wrong solutions than the two-chain case studied in this paper. Further, the existence of a certain type of decay chain implies that there are always events with two identical such decay chains. Events with two identical decay chains always provide more information for the masses of the particles in the decay chains. The challenge of two-chain techniques is that one needs to identify those events in which there are indeed two identical chain decays. Ideally, one would divide the observed events into different channels according to their event topologies (chain type 1 + chain type 1, chain type 1 + chain type 2, chain type 2 + chain type 2, ...), apply methods appropriate to each topology and, in order not to lose statistics, combine all these channels in the analysis. For this reason, it is very important to extend the studies in Ref. [6] and the current paper to other event topologies.

The importance of using one-chain decay information is illustrated in the SPS1a case. Since there are many more dilepton events than four-lepton events, the dilepton edge given in Eq. (25) can be measured very precisely. If available, one should certainly incorporate this measurement into the two-chain techniques to better determine the masses. This kind of "hybrid" approach has been studied here and in [21,22,24]. As summarized below, adding the

dilepton edge information improves the two-chain mass determinations obtained following the basic procedure developed in this paper.

In our procedure, we have applied a set of bias-reduction methods. In particular, to reduce the number of combinations, we have utilized the fact that correct combinations can pair with relatively more events than wrong combinations. Alternatively, one may try to reduce the number of wrong combinations before doing the event reconstruction. For example, one could try to group objects into two hemispheres [36] and assume that the objects in the same hemisphere come from the same decay chain. However, this only works well when the initial particles are substantially boosted, while the squarks in our case are produced mostly close to the threshold without large boosts. For small boosts, the quark and $\tilde{\chi}_2^0$ from squark decay actually belong to two opposite hemispheres instead of the same one. The directions of the subsequent $\tilde{\chi}_2^0$ and $\tilde{\ell}$ decay products are even more random. We have applied the hemisphere method on a set of ideal events from squark pair production, each containing 2 quarks and 4 leptons according to the decay chain in Fig. 1. Even without any complications from extra jets, experimental smearing and so forth, only about 12% of the events have the decay chains correctly identified (this does not account for the ambiguity of the two leptons in the same decay chain).

One could explore the effectiveness of imposing a cut which accepts only events with substantial thrust or small circularity before separating each event into two hemispheres. It is not clear to us that the gain from decreasing the combinatorics problem would outweigh the reduced statistics associated with the fact that such a cut would remove a large fraction of the available events.

However, we *have* shown that when a two-particle mass edge (such as the dilepton mass edge in the examples we have considered) can be identified, it is very useful to impose a cut whereby only solutions that give a dilepton mass within roughly ± 20 GeV of the edge location are retained. By applying this cut before proceeding with the rest of our analysis procedure the mass peak biases are essentially eliminated and the errors on the central mass values are very similar. In addition, we have shown that this cut is capable of essentially eliminating contamination of the mass determinations for the SUSY signal of interest by events coming from some other SUSY signal that does not share the same dilepton edge location. Presumably, any other recognizable kinematic edge could be exploited in similar fashion, but dilepton mass edges will typically be least impacted by detector momentum smearing.

In conclusion, we have proposed a kinematic technique for mass determination in events with two invisible dark matter particles. The technique seeks constraints on the mass space from measured momenta. In Sec. II, we have given general constraint counting and discuss the corresponding strategies for both the single decay chain case and the double decay chain case. In the former, one only uses the information from one of the two decay chains in each event; in the latter, one uses information associated with both decay chains. The constraints include the mass-shell constraints for the dark matter particle as well as all intermediate particles. In the double decay chain case, we obtain extra constraints from the measured missing transverse momenta. In both cases, more constraints are available when the decay chain is longer. In certain instances (this includes the single decay chain case with 3 visible particles per chain and the double identical decay chain case with 2 visible particles per chain), we obtain discrete solutions for the missing momenta by using trial masses for the unknown particles. Requiring that we obtain physical solutions for the momenta, the consistent mass region becomes more restricted when more events are included. The actual masses are obtained by studying the kinematic distributions of the consistent region. This is the strategy we adopted in Refs. [6,24]. When the decay chains presented in each event are longer, it is possible to obtain discrete solutions for the momenta (and therefore the masses) by combining the constraints from a few different events, *without* assuming any trial masses. This occurs when there are 4 or more visible particles per decay chain for the single chain case, or 3 or more visible particles per decay chain for the double chain case. In this article, we have focused on the double identical chain case with the number of visible particles per chain fixed to 3.

In our case, the constraints can be solved for discrete solutions of the unknown masses when two events are combined. However, because the system of equations contains quadratic equations, wrong solutions are introduced. Nonetheless, if the visible momenta could be measured without errors, it would take only three events to obtain the correct masses. This is because the wrong solutions would be different for the three different possible event pairings whereas the correct solution remains the same for every event pair. This remains true even when wrong combinations are included. However, in practice the nonzero experimental resolutions imply that the correct solution distribution becomes smeared which then overlaps with the distribution coming from wrong solutions, wrong combinations, and background events. Despite this, we have shown that when two of the visible particles in a decay chain are leptons, we obtain good precision ($\sim$ a few GeV) if a few hundred events are available. We have developed methods to reduce the number of wrong combinations and backgrounds. The resulting mass solution distributions are clearly peaked around the input masses with small systematic errors which can be eliminated either by comparing with Monte Carlo distributions around the estimated masses or by imposing an initial dilepton edge mass cut on the accepted solutions. An important assumption for doing this comparison is that the distributions are only sensitive to the masses instead of the

underlying theories. We have shown that this is indeed the case by comparing two distinct theories with different spin structure, minimal supersymmetric standard model and UED. We have set the spectra of the two models to be the same and examined the mass distributions, which show little difference. Correspondingly, the systematic errors introduced by model dependence are much smaller than the statistical errors.

Finally, we comment on possible improvements of our method. Given the precision of the purely kinematic results and availability of spin determination techniques [37], it would certainly be possible to figure out the underlying theory/spins and then apply model-dependent techniques to refine the mass determinations. For example, one could adopt a likelihood method similar to the one used in the top mass measurement with dilepton events at Tevatron [38]. In this method, a probability density, as a function of all unknown masses, is defined for each event by convoluting matrix elements and detector resolution functions. One then obtains the joint probability by taking the product of the probability densities from all events. The best-fit masses are given by the values that maximize the joint probability. Compared with top quark mass measurement, the event topology considered in this paper is more complicated since more final state particles are involved. One also needs to scan a four dimensional instead of one dimensional mass space to minimize the probability function. Therefore, another benefit of the purely kinematic method is that it significantly simplify the computation by reducing the candidate mass space to a very small region.

We could also consider simplified likelihood methods using the detector resolution functions, but without the knowledge of the matrix element. One such method is based on the constraints, Eqs. (6)–(10), which can be viewed as a generalization of the method discussed in this paper: for each event, we can eliminate the four-momenta of the two missing particles by using Eqs. (6)–(10). We are then left with two equations in the form

$$f(m_N, m_X, m_Y, m_Z; p_{\rm vis}) = 0, \tag{28}$$

$$g(m_N, m_X, m_Y, m_Z; p_{\rm vis}) = 0, \tag{29}$$

where f and g are functions of the masses and all visible momenta, $p_{\rm vis}$. If we require that the equalities hold exactly, then each event defines a two-dimensional surface in a four-dimensional mass space. The two-dimensional surfaces of two different events intersect at discrete points, which correspond to the mass solutions. This is equivalent to saying that in the mass space, for each event we assign a nonzero probability density for points on the surface and a zero probability density for points off the surface. Then by combining two events, the joint probability is nonzero only at discrete points. Obviously, a more sophisticated method is to assign a maximum probability density for points on the surface, and smaller but nonzero probability densities for points away from the surface. This can be done by calculating the χ^2 distribution for each point in the mass space,

$$\chi^2 = \frac{f^2}{\sigma_f^2} + \frac{g^2}{\sigma_g^2}, \tag{30}$$

where

$$\sigma_f^2 \equiv \sum_{p_{\rm vis}} \left(\frac{\partial f}{\partial p_{\rm vis}} \sigma_{p_{\rm vis}} \right)^2, \tag{31}$$

and a similar formula holds for σ_g. Here, we have assumed Gaussian distributions for the invisible momenta $p_{\rm vis}$ with errors given by $\sigma_{p_{\rm vis}}$. Of course, in practice Eq. (31) is complicated by correlations among the visible momenta, as well as the correlation between f and g. We then sum the χ^2 over all events and obtain the masses when the total χ^2 is minimized. In practice, f and g are very complicated functions of the visible momenta, which may result in large round-off errors when calculating the derivatives in Eq. (31). As in the matrix element method, one also needs to efficiently minimize the χ^2 over a four-dimensional space. Therefore, further studies are needed to determine whether this approach can improve the precision of the mass measurement over the simple method discussed in this paper.

In closing, we note that the program for determining the $m_{Z,Y,Z,N}$ solutions as a function of input visible momenta for two-chain events is available at Ref. [39] . Stand-alone programs that implement the methods of Refs. [6,24] are also available at this same website.

ACKNOWLEDGMENTS

This work was supported in part by the U.S. Department of Energy Grant No. DE-FG02-91ER40674.

APPENDIX A: SOLVING THE CONSTRAINT EQUATIONS

In this Appendix, we describe in detail our procedure to solve the system of the kinematic constraint equations

$$p_1^2 = p_2^2 = q_1^2 = q_2^2, \tag{A1}$$

$$\begin{aligned} 2p_1 \cdot p_3 + p_3^2 = 2p_2 \cdot p_4 + p_4^2 = 2q_1 \cdot q_3 + q_3^2 \\ = 2q_2 \cdot q_4 + q_4^2, \end{aligned} \tag{A2}$$

$$\begin{aligned} 2(p_1 + p_3) \cdot p_5 + p_5^2 &= 2(p_2 + p_4) \cdot p_6 + p_6^2 \\ &= 2(q_1 + q_3) \cdot q_5 + q_5^2 \\ &= 2(q_2 + q_4) \cdot q_6 + q_6^2, \end{aligned} \tag{A3}$$

$$\begin{aligned} 2(p_1 + p_3 + p_5) \cdot p_7 + p_7^2 &= 2(p_2 + p_4 + p_6) \cdot p_8 + p_8^2 \\ &= 2(q_1 + q_3 + q_5) \cdot q_7 + q_7^2 \\ &= 2(q_2 + q_4 + q_6) \cdot q_8 + q_8^2, \end{aligned} \tag{A4}$$

$$p_1^x + p_2^x = p_{\rm miss}^x, \qquad p_1^y + p_2^y = p_{\rm miss}^y, \tag{A5}$$

$$q_1^x + q_2^x = q_{\rm miss}^x, \qquad q_1^y + q_2^y = q_{\rm miss}^y. \tag{A6}$$

As mentioned earlier, it is straightforward to numerically solve the above equations using commercial software such as MATHEMATICA, but the speed is intolerably low. Instead, we solve the equation system using a programming language such as C^{++}. The idea is to reduce the system to a univariate polynomial equation whose coefficients are (fixed) functions of the original visible momenta. Note that it is convenient to obtain the coefficient functions with the assistance of MATHEMATICA. After that, the functions are hard coded in the C^{++} program and MATHEMATICA is no longer needed. The univariate equation can then be solved numerically using any available polynomial solver. The MATHEMATICA notebook and the C^{++} code are available from Ref. [39] or any of the authors. The key MATHEMATICA operations are also described in Appendix B for reference. The method can be easily generalized to solve other polynomial equations efficiently.

It is straightforward to eliminate 13 variables using the 13 linear equations and obtain 3 quadratic equations with 3 variables. Generically, 3 quadratic equations can be written as

$$\begin{aligned} z^2 + a_8 zy + a_7 zx + a_6 z + a_5 y^2 + a_4 yx \\ + a_3 y + a_2 x^2 + a_1 x + a_0 = 0, \end{aligned} \tag{A7}$$

$$\begin{aligned} zy + b_7 zx + b_6 z + b_5 y^2 + b_4 yx + b_3 y \\ + b_2 x^2 + b_1 x + b_0 = 0, \end{aligned} \tag{A8}$$

$$zx + c_6 z + c_5 y^2 + c_4 yx + c_3 y + c_2 x^2 + c_1 x + c_0 = 0, \tag{A9}$$

where x, y, z are variables, and a_i, b_i, c_i are coefficients as functions of the original visible momenta. We have ordered the left-hand terms lexicographically in the order $z > y > x$. We will eliminate variables also in this order and eventually obtain a univariate equation in x. In our implementation, we choose x, y, z to be p_1^0, p_2^0, q_1^0 or some permutation thereof. In fact, we solve several times for several different permutations just to make sure we do not miss any solutions.

First, by calculating (A7) $\times\, y -$ (A8) $\times\, z$, we cancel the term $z^2 y$ and obtain a polynomial equation with leading term $-b_7 x z^2$. We repeatedly use (A7)–(A9), to reduce the polynomial, i.e., to eliminate the leading term of the polynomial. For example, we eliminate the $-b_7 x z^2$ term by subtracting $-b_7 \times$ (A9). The next leading term is $\propto z^2$, which again can be eliminated by subtracting from it Eq. (A7) with the appropriate coefficient. Repeating this procedure until it cannot be reduced further, we obtain an equation in the form

$$\begin{aligned} z + y^3 + y^2 x + y^2 + yx^2 + yx + y + x^3 \\ + x^2 + x + 1 = 0, \end{aligned} \tag{A10}$$

where we have omitted all coefficients. Similarly, reducing (A7) $\times\, x -$ (A8) $\times\, z$, we obtain another equation also in the form of (A10). Canceling the leading term, we eliminate the variable z and obtain a cubic equation in (y, x). Reducing (A8) $\times\, x -$ (A9) $\times\, y$, we obtain another cubic equation in (y, x).

We are then left with 2 cubic equations in 2 variables,

$$\begin{aligned} &y^3 + d_8 y^2 x + d_7 y^2 + d_6 yx^2 + d_5 yx + d_4 y + d_3 x^3 + d_2 x^2 \\ &\quad + d_1 x + d_0 = 0, \\ &y^2 x + e_7 y^2 + e_6 yx^2 + e_5 yx + e_4 y + e_3 x^3 \\ &\quad + e_2 x^2 + e_1 x + e_0 = 0, \end{aligned} \tag{A11}$$

where the coefficients d_i and e_i are derived from a_i, b_i, and c_i in Eqs. (A7)–(A9). This system can be solved using the following method. The resultant [40] of two univariate polynomials is defined as follows. Given a polynomial [note that the a_i and b_i below are not those given in Eqs. (A7) and (A8)]

$$P(x) = a_n x^n + a_{n-1} x^{n-1} + \ldots + a_1 x + a_0, \tag{A12}$$

of degree n with roots α_i, $i = 1, \ldots, n$ and a polynomial

$$Q(x) = b_m x^m + b_{m-1} x^{m-1} + \ldots + b_1 x + b_0 \tag{A13}$$

of degree m with roots β_j, $j = 1, \ldots, m$, the resultant $\rho(P, Q)$, also denoted $R(P, Q)$ and also called the eliminant, is defined by

$$\rho(P, Q) = a_n^m b_m^n \prod_{i=1}^{n} \prod_{j=1}^{m} (\alpha_i - \beta_j). \tag{A14}$$

The resultant is also given by the determinant of the corresponding Sylvester matrix [41]. The Sylvester matrix associated to polynomials P and Q is the $(n + m) \times (n + m)$ matrix obtained as follows:

(1) The first row is

$$\begin{pmatrix} a_n & a_{n-1} & \cdots & a_1 & a_0 & 0 & \cdots & 0 \end{pmatrix}. \tag{A15}$$

(2) The second row is the first row, shifted one column to the right; the first element of the row is zero.

(3) The following $(m - 2)$ rows are obtained the same way, still filling the first column with a zero.

(4) The $(m+1)$th row is

$$\begin{pmatrix} b_m & b_{m-1} & \cdots & b_1 & b_0 & 0 & \cdots & 0 \end{pmatrix}. \quad \text{(A16)}$$

(5) The following rows are obtained the same way as before.

Taking the left-hand side polynomials of Eqs. (A11) as P and Q, respectively, we have $n = 3$ and $m = 2$ and the Sylvester matrix in this case is

$$S = \begin{pmatrix} a_3 & a_2 & a_1 & a_0 & 0 \\ 0 & a_3 & a_2 & a_1 & a_0 \\ b_2 & b_1 & b_0 & 0 & 0 \\ 0 & b_2 & b_1 & b_0 & 0 \\ 0 & 0 & b_2 & b_1 & b_0 \end{pmatrix}, \quad \text{(A17)}$$

where

$$a_3 = 1, \qquad a_2 = d_8x + d_7, \qquad a_1 = d_6x^2 + d_5x + d_4,$$
$$a_0 = d_3x^3 + d_2x^2 + d_1x + d_0, \qquad b_2 = x + e_7,$$
$$b_1 = e_6x^2 + e_5x + e_4, \qquad b_0 = e_3x^3 + e_2x^2 + e_1x + e_0. \quad \text{(A18)}$$

Therefore, $\rho(P, Q) = \det S$ is a 9th order polynomial in x. By Eq. (A14), if $\rho(P, Q) = 0$, the two equations in (A11) have at least one common root. In other words, for each root x_i of $\det S = 0$, we can find a y_i such that (x_i, y_i) is a solution of Eqs. (A11). Thus, the problem has been reduced to the solution of a 9th order polynomial equation with real coefficients. We solve it numerically using the algorithm TOMS/493 [30]. Note that one of the roots is fake. It can easily be identified and eliminated by substituting back each of the solutions into the original equations to identify the root that is not actually a solution of the original equations.

APPENDIX B: MATHEMATICA FILE

Many of the operations described in Appendix A can be conveniently done in MATHEMATICA, which we describe below.

First, suppose we have obtained Eqs. (A7)–(A9). As mentioned, the coefficients a_i, b_i, and c_i are functions of the visible momenta. These functions are usually complicated, therefore it is always desirable to use intermediate parameters such as a_i, b_i, and c_i, and normalize the coefficients of the leading terms to 1. To obtain Eqs. (A11), we use the MATHEMATICA function PolynomialReduce:

$$P1 = a0 + a1x + a2x^2 + a3y + a4yx + a5y^2 + a6z + a7zx + a8zy + z^2;$$
$$P2 = b0 + b1x + b2x^2 + b3y + b4yx + b5y^2 + b6z + b7zx + zy;$$
$$P3 = c0 + c1x + c2x^2 + c3y + c4yx + c5y^2 + c6z + zx;$$
$$P4 = \text{Expand}[\text{PolynomialReduce}[P1y - P2z, \{P1, P2, P3\}, \{z, y, x\}][[2]]];$$
$$P5 = \text{Expand}[\text{PolynomialReduce}[P1x - P3z, \{P1, P2, P3\}, \{z, y, x\}][[2]]];$$
$$P6 = \text{Expand}[\text{PolynomialReduce}[P2x - P3y, \{P1, P2, P3\}, \{z, y, x\}][[2]]];$$

where we obtain $P4$ and $P5$ as two polynomials in the form of Eq. (A10) and $P6$ a cubic polynomial in x and y. Then it is straightforward to cancel the variable z and cast the remaining two polynomials in the form of Eqs. (A11). The resultant is obtained from

$$P7 = d0 + d1x + d2x^2 + d3x^3 + d4y + d5yx + d6yx^2 + d7y^2 + d8y^2x + y^3;$$
$$P8 = e0 + e1x + e2x^2 + e3x^3 + e4y + e5yx + e6yx^2 + e7y^2 + y^2x;$$
$$\text{resultant} = \text{Resultant}[P7, P8, y];$$

where resultant is the 9th order polynomial in x we are seeking. After obtaining the roots of the resultant numerically, y and z can be uniquely determined as follows: Defining the polynomial $P9$ by

$$P9 = \text{Expand}[\text{PolynomialReduce} \times [P7x - P8y, \{P7, P8\}, \{y, x\}][[2]]],$$

we see that $P9$ can be written in the form

$$P9 = f0 + f1x + f2x^2 + f3x^3 + f4x^4 + f5y + f6yx + f7yx^2 + f8yx^3 + y^2.$$

It turns out that the polynomial $P10$ defined below is linear in y:

$$P10 = \text{Expand}[\text{PolynomialReduce} \times [P9x - P8, \{P8, P9\}, \{y, x\}][[2]]].$$

We then obtain y from the equation $P10 = 0$ by substituting in the solutions for x. Once this is done, we can obtain z from the equation $P4 = 0$ since it is linear in z. One of the 9 solutions is a fake solution to the original equation system, which can be easily identified by finding the solution which does not actually solve the system of equations when substituted back into the system of equations.

[1] T. Appelquist, H. C. Cheng, and B. A. Dobrescu, Phys. Rev. D **64**, 035002 (2001).
[2] H. C. Cheng, D. E. Kaplan, M. Schmaltz, and W. Skiba, Phys. Lett. B **515**, 395 (2001); Phys. Rev. D **66**, 056006 (2002).
[3] H. C. Cheng and I. Low, J. High Energy Phys. 09 (2003) 051; J. High Energy Phys. 08 (2004) 061.
[4] K. Agashe and G. Servant, Phys. Rev. Lett. **93**, 231805 (2004).
[5] E. A. Baltz, M. Battaglia, M. E. Peskin, and T. Wizansky, Phys. Rev. D **74**, 103521 (2006).
[6] H. C. Cheng *et al.*, J. High Energy Phys. 12 (2007) 076.
[7] H. C. Cheng *et al.*, Phys. Rev. Lett. **100**, 252001 (2008).
[8] I. Hinchliffe *et al.*, Phys. Rev. D **55**, 5520 (1997).
[9] C. G. Lester and D. J. Summers, Phys. Lett. B **463**, 99 (1999); A. Barr, C. Lester, and P. Stephens, J. Phys. G **29**, 2343 (2003).
[10] H. Bachacou, I. Hinchliffe, and F. E. Paige, Phys. Rev. D **62**, 015009 (2000).
[11] B. C. Allanach, C. G. Lester, M. A. Parker, and B. R. Webber, J. High Energy Phys. 09 (2000) 004.
[12] B. K. Gjelsten, D. J. Miller, and P. Osland, J. High Energy Phys. 12 (2004) 003.
[13] K. Kawagoe, M. M. Nojiri, and G. Polesello, Phys. Rev. D **71**, 035008 (2005).
[14] D. J. Miller, P. Osland, and A. R. Raklev, J. High Energy Phys. 03 (2006) 034.
[15] C. G. Lester, M. A. Parker, and M. J. White, J. High Energy Phys. 01 (2006) 080.
[16] N. Arkani-Hamed, G. L. Kane, J. Thaler, and L. T. Wang, J. High Energy Phys. 08 (2006) 070.
[17] J. M. Butterworth, J. R. Ellis, and A. R. Raklev, J. High Energy Phys. 05 (2007) 033.
[18] C. Lester and A. Barr, J. High Energy Phys. 12 (2007) 102.
[19] W. S. Cho, K. Choi, Y. G. Kim, and C. B. Park, Phys. Rev. Lett. **100**, 171801 (2008); J. High Energy Phys. 02 (2008) 035.
[20] A. J. Barr, B. Gripaios, and C. G. Lester, J. High Energy Phys. 02 (2008) 014.
[21] G. G. Ross and M. Serna, Phys. Lett. B **665**, 212 (2008); A. J. Barr, G. G. Ross, and M. Serna, Phys. Rev. D **78**, 056006 (2008); A. J. Barr, A. Pinder, and M. Serna, Phys. Rev. D **79**, 074005 (2009).
[22] M. M. Nojiri, G. Polesello, and D. R. Tovey, J. High Energy Phys. 05 (2008) 014.
[23] P. Huang, N. Kersting, and H. H. Yang, arXiv:0802.0022; M. Bisset, R. Lu, and N. Kersting, arXiv:0806.2492.
[24] H. C. Cheng and Z. Han, J. High Energy Phys. 12 (2008) 063.
[25] M. Burns, K. Kong, K. T. Matchev, and M. Park, J. High Energy Phys. 03 (2009) 143.
[26] P. Konar, K. Kong, and K. T. Matchev, J. High Energy Phys. 03 (2009) 085.
[27] A. Papaefstathiou and B. Webber, J. High Energy Phys. 06 (2009) 069.
[28] M. Burns, K. T. Matchev, and M. Park, J. High Energy Phys. 05 (2009) 094.
[29] http://mathworld.wolfram.com/GroebnerBasis.html.
[30] http://www.netlib.org/toms/493; see also http://www.akiti.ca/rpoly_ak1_Intro.html.
[31] B. C. Allanach *et al.*, in *Proceedings of the APS/DPF/DPB Summer Study on the Future of Particle Physics (Snowmass 2001)*, edited by N. Graf (Snowmass, Colorado), p. P125; B. C. Allanach *et al.*, Eur. Phys. J. C **25**, 113 (2002).
[32] T. Sjostrand, S. Mrenna, and P. Skands, J. High Energy Phys. 05 (2006) 026.
[33] J. Conway, http://www.physics.ucdavis.edu/~conway/research/software/pgs/pgs4-general.htm.
[34] M. Bahr *et al.*, Eur. Phys. J. C **58**, 639 (2008).
[35] J. M. Smillie and B. R. Webber, J. High Energy Phys. 10 (2005) 069.
[36] G. L. Bayatian *et al.* (CMS Collaboration), J. Phys. G **34**, 995 (2007); S. Matsumoto, M. M. Nojiri, and D. Nomura, Phys. Rev. D **75**, 055006 (2007).
[37] For a review, see L. T. Wang and I. Yavin, Int. J. Mod. Phys. A **23**, 4647 (2008).
[38] T. Aaltonen *et al.* (CDF Collaboration), Phys. Rev. Lett. **102**, 152001 (2009).
[39] Go to http://particle.physics.ucdavis.edu/hefti/projects/doku.php?id=wimpmass.
[40] http://mathworld.wolfram.com/Resultant.html.
[41] http://mathworld.wolfram.com/SylvesterMatrix.html.
[42] E. Richter-Was, D. Froidevaux, and L. Poggioli, ATLAS Internal Note ATL-PHYS-98-131.