FERMILAB-PUB-13-456-A-T, CALT 68-2866

Super-Razor and Searches for Sleptons and Charginos at the LHC

Matthew R. Buckley[1,2], Joseph D. Lykken[3], Christopher Rogan[4,5], and Maria Spiropulu[4]
[1] *Center for Particle Astrophysics, Fermi National Accelerator Laboratory, Batavia, IL 60510*
[2] *Department of Physics and Astronomy, Rutgers University, Piscataway, NJ 08954*
[3] *Theoretical Physics Department, Fermi National Accelerator Laboratory, Batavia, IL 60510*
[4] *Lauritsen Laboratory of Physics, California Institute of Technology, Pasadena, CA 91125 and*
[5] *Department of Physics, Harvard University, Cambridge, MA 02138*

(Dated: October 21, 2013)

Direct searches for electroweak pair production of new particles at the LHC are a difficult proposition, due to the large background and low signal cross sections. We demonstrate how these searches can be improved by a combination of new razor variables and shape analysis of signal and background kinematics. We assume that the pair-produced particles decay to charged leptons and missing energy, either directly or through a W boson. In both cases the final state is a pair of opposite sign leptons plus missing transverse energy. We estimate exclusion reach in terms of sleptons and charginos as realized in minimal supersymmetry. We compare this super-razor approach in detail to analyses based on other kinematic variables, showing how the super-razor uses more of the relevant kinematic information while achieving higher selection efficiency on signals, including cases with compressed spectra.

I. INTRODUCTION

Searches with the ATLAS and CMS detectors at the Large Hadron Collider have already placed strong lower bounds on the mass of pair-produced strongly-interacting gluinos or degenerate squarks decaying into final states with missing transverse energy [1–5]. A determination of the role of supersymmetry in electroweak symmetry breaking requires a much broader campaign of searches, many of which are already underway. Some of these searches present special challenges at a hadron machine, even when they involve the pair production of relatively light superpartners. Examples include light stops whose decays closely resemble those of top quarks [6–18], a variety of models with compressed spectra, R-parity violating models [15, 19–23], and relatively long-lived superpartners with displaced decays [24, 25].

Of particular importance to this program is the direct electroweak production of charginos, neutralinos, and sleptons at the LHC. Relatively light charginos and neutralinos have a possible connection to weakly-interacting dark matter in supersymmetry models with conserved R-parity. Light sleptons are motivated by the measured value of the anomalous magnetic moment $(g-2)$ of the muon [26, 27], providing a thermal annihilation cross section for bino-like neutralino dark matter [28], and the possibility that the branching fraction of the newly-discovered Higgs boson into two photons is enhanced over the Standard Model prediction [29]. Charginos, neutralinos, and sleptons could also appear in cascade decays of heavier colored superpartners, but this prospect merely emphasizes the importance of being able to produce these lighter superpartners directly.

We will focus on electroweak pair production of charged particles that decay to charged leptons and a stable (or long-lived) neutral particle, appearing in the detector only as missing transverse energy ($\vec{E}_T^{\rm miss}$). The decay to leptons can occur either directly or through the leptonic decay of a W boson. We will consider two canonical examples: sleptons of the first or second generation ($\tilde{e}^-\tilde{e}^+$ or $\tilde{\mu}^-\tilde{\mu}^+$) with 100% branching into leptons and the lightest supersymmetric particle (LSP) neutralino, and charginos ($\tilde{\chi}^+\tilde{\chi}^-$) decaying through an on- or off-shell W boson and the neutralino LSP. In the latter case, we require the W to decay leptonically. In both cases, we set all other superpartner masses heavy, including the other charginos and neutralinos. Though our study is performed assuming a supersymmetric model, it can easily be generalized to other scenarios that contain similar particles with the same broad characteristics. The pair production of tau partners (*e.g.* staus) has different backgrounds and will be considered in a later work.

Searches at LEP have already set lower bounds on the masses of new charged particles, ranging between 90 and 105 GeV assuming supersymmetric-like cross sections [30]. The ATLAS and CMS collaborations have performed model-independent dilepton searches for both the slepton and chargino pair production scenarios we consider in this paper. ATLAS, using 20.3 fb^{-1} of integrated luminosity at 8 TeV places an upper bound of 300 GeV on left-handed sleptons (assuming massless neutralinos), and an upper limit of 450 GeV on charginos assuming a 100% branching ratio to leptons and neutralinos [31, 32]. CMS places a 300 GeV bound on pair production of degenerate selectrons and smuons using 19.5 fb^{-1} at 8 TeV, and 550 GeV on chargino pair production decaying to neutralinos with 100% branching ratio [33, 34]. Both experiments [33–38] have also performed multilepton searches for production of heavier chargino/neutralino pairs (such as $\tilde{\chi}_2^0\tilde{\chi}_2^0$ or $\tilde{\chi}_2^0\tilde{\chi}_1^\pm$), followed by cascades of the form $\tilde{\chi}_2^0 \to W^\pm\tilde{\chi}_1^\mp \to \ell^\pm\ell^\mp\nu\bar{\nu}\tilde{\chi}_1^0$ to obtain three or more leptons in the final state.

We propose several techniques that can increase the sensitivity of the LHC experiments to electroweak pair production in the dilepton channel, using the data currently available from the completed 8 TeV run. Starting from the

CMS razor variables [39, 40] (see Refs. [41–43] for applications), we develop an improved version that more accurately approximates the production frame and center-of-mass (CM) energy scale of the pair production event, compared to the original razor formulation. This "super-razor" results in a set of mass variables, $\sqrt{\hat{s}_R}$ and M_Δ^R that contain information about the mass differences involved in the pair production and subsequent decay, allowing for discrimination between signal and background. In addition, the derivation of these mass variables involves constructing the approximate boost to the pair production frame, followed by a boost to an approximation of the decay frame. From this boost direction and the momenta of the visible particles, we construct angular variables $\Delta\phi_R^\beta$ and $|\cos\theta_{R+1}|$ that also distinguish between the signal events and background. Using these super-razor variables, we develop a new set of selection criteria and apply a multi-dimensional shape analysis to maximize the sensitivity to signal over the dominant backgrounds (primarily W^-W^+ and Drell-Yan + jets). Shape analyses have been implemented by experimental groups [39, 42] and have been used in theoretical proposals for new searches [9, 41]. As we show through direct comparison to ATLAS- and CMS-like searches, this technique is promising in difficult channels.

In the next section we review the construction of the standard razor variables, followed by a derivation of the improved super-razor and the associated angular variables of interest. The background and signal simulations are described in Section III, along with comparisons to the alternative searches by the ATLAS and CMS experiments that employed the kinematic variables M_{T2} [44, 45] and $M_{CT\perp}$ [46, 47]. The shape analysis techniques and statistical tools are described in Section IV. Our expected exclusion limits for 20 fb^{-1} of integrated luminosity at 8 TeV are presented in Section V.

II. KINEMATIC VARIABLES

We are interested in the pair production of particles that each decay either into a lepton and a massive undetected "invisible" particle, or into an invisible particle and a W boson, followed by leptonic decays of the W's. For specificity we consider slepton pair production and chargino pair production as in the minimal supersymmetric standard model (MSSM):

$$pp \to \tilde{\ell}^-\tilde{\ell}^+ \;\to\; (\ell^-\tilde{\chi}_1^0)(\ell^+\tilde{\chi}_1^0) \tag{1}$$

$$pp \to \tilde{\chi}_1^-\tilde{\chi}_1^+ \;\to\; (W^-\tilde{\chi}_1^0)(W^+\tilde{\chi}_1^0) \to (\ell^-\bar{\nu}\tilde{\chi}_1^0)(\ell^+\nu\tilde{\chi}_1^0). \tag{2}$$

In both cases, the observables at the LHC are the same: opposite-sign leptons (which may or may not be of the same flavor) and large missing transverse energy. Searching for these types of new particles is difficult for several reasons. The production cross sections are small, on the order of tens of femtobarns to a few picobarns before branching fractions. The background cross sections are large. The dilepton backgrounds (primarily W^-W^+ and Drell-Yan + jets production but also with contributions from WZ, ZZ, and top pair production) have kinematic distributions that are similar to the signal, since most of these backgrounds have two charged leptons and real missing transverse momentum from neutrinos. Kinematic variables sensitive to the mass (or mass squared) differences between the parent and invisible particles are less effective in regions of the mass plane when the parent/daughter mass difference is close to or smaller than the W mass. The $M_{CT\perp}$ [46, 47] and M_{T2} [44, 45] variables used by the CMS [34] and ATLAS [32] (see also Refs. [48, 49] for other experimental applications of M_{T2}) searches have this drawback, as does the original formulation of the razor variable, as we will show.

Our new work is motivated by the razor variables M_R and R, originally developed in Ref. [39, 40] to distinguish between new massive strongly interacting particles (*e.g.* squarks and gluinos) and QCD background, and implemented by CMS [42, 43] in various searches. Razor variables have also been demonstrated to be of use in distinguishing signal and background in electro-weak channels [41]. Here, we describe the motivating principles behind the razor (for a full description, see Ref. [50]), and then propose a series of improvements that more accurately capture the relevant mass differences in events with final states relevant to electroweak production. We then introduce new kinematic variables, motivated by the construction of the improved razor, which contain information about the ratio of mass scales of the particles in the event.

A. Principles of the razor

The razor variables are intended for use in a very generic new physics scenario. Two massive particles, S_1 and S_2, with a common mass m_S, are produced at the LHC. Each then decays into a set of visible particles (Q_1 and Q_2, respectively) and an invisible particle (χ_1 and χ_2) with common mass m_χ. For this paper, we will be assuming that the visible decays each consist of a single effectively massless particle (an electron or muon). In a more inclusive razor

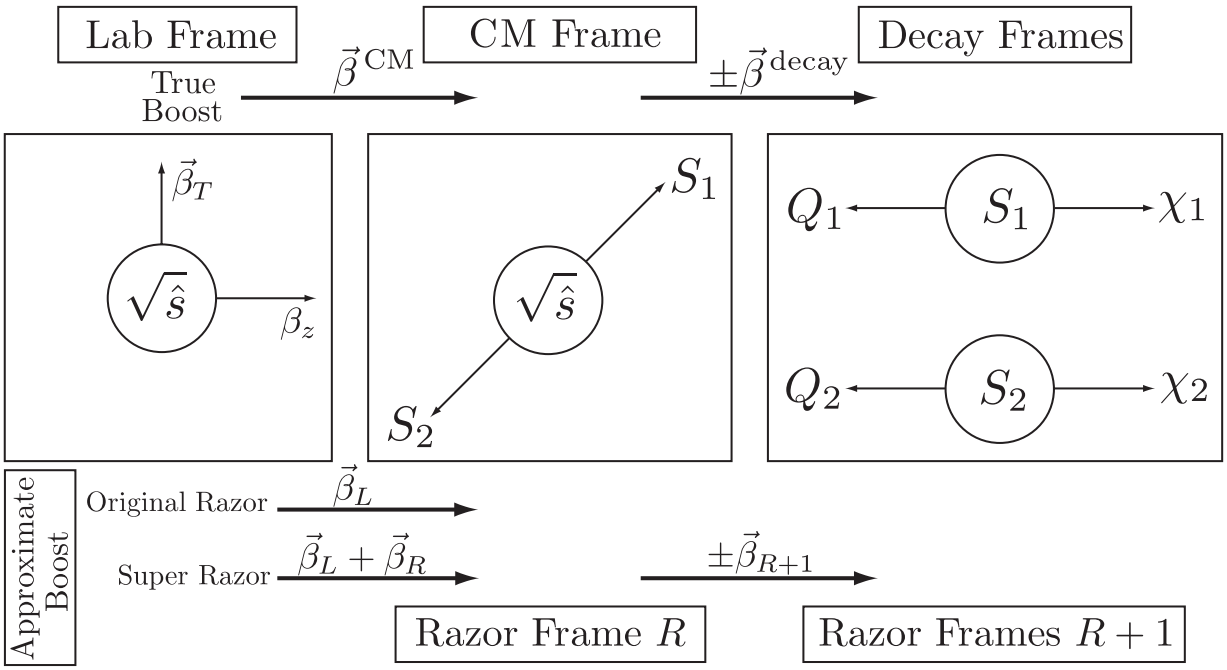

FIG. 1: Sketch of the three sets of frames relevant to the razor reconstruction: the lab frame, the pair production frame for S_1 and S_2, and the two decay frames of the particles S_i. The approximate razor frame identified with each physically relevant frame is also shown, along with the actual and approximate boosts from one frame to the next. By convention, we label each boost by the destination frame (*i.e.* boost $\vec{\beta}^{\,\mathrm{CM}}$ takes you *from* the lab from *to* the pair production center of mass frame).

analysis decays may include more than one visible particle, in which case their four-momenta are summed to create two visible objects known as megajets.

If we could identify the rest frames of the S_i decay, then in that frame the energies E_i of the visible Q_i would be

$$2E_1 = 2E_2 = \frac{m_S^2 - m_\chi^2}{m_S} \equiv M_\Delta. \tag{3}$$

If this frame could be identified using the available visible momenta and the E_T^{miss} of the invisible particles, then the momentum of the visible Q_i in signal events would be easily distinguished from background, which does not inherit information about this scale (save in cases where $M_\Delta \sim m_W$).

However, as is well understood in hadron colliders, with S_i both decaying into at least one invisible particle, we do not possess enough kinematic information to reconstruct the decay frames. The approach of the razor is to make a series of assumptions which, while not capable of reconstructing the precise decay frames event-by-event, approximate the relevant frames on average. In both simulations and data these approximations work well in the experimental environment of the LHC.

There are three kinds of frames relevant to pair production at the LHC: the lab frame, the pair production center-of-mass (CM) frame, and the two decay frames (see Figure 1). The initial assumption made by the original razor construction is that the heavy parent particles are generally produced near threshold, due to the fall-off of the parton distribution functions with CM energy $\sqrt{\hat{s}}$. If we could identify the boost $\vec{\beta}^{\,\mathrm{CM}}$ from the lab frame into the S_1 and S_2 production frame (the center of mass frame CM), then this could serve as an approximation to the decay frames. We approximate this frame by making a longitudinal boost $\vec{\beta}_L$ to the razor frame R, which is defined here as the frame where the two sets of visible decay products Q_1 and Q_2 have equal and opposite z-component of momentum. This boost has magnitude

$$\beta_L = \frac{q_1^z + q_2^z}{E_1 + E_2}. \tag{4}$$

Here, E_i is the energy of decay product Q_i and q_i^z is the z-component of the momentum.

In this razor frame, we expect $2E_{R1} \approx 2E_{R2} \approx M_\Delta$. Writing the boosted momenta in terms of lab-frame observables, we define a longitudinally boost-invariant mass

$$M_R^2 = (E_1 + E_2)^2 - (q_1^z + q_2^z)^2. \tag{5}$$

We expect that the distribution of M_R for signal events will have a peak near M_Δ, assuming that our approximations of near-threshold production and $q_1^z \approx -q_2^z$ are correct on a statistical basis. Background events will not, in general, have any special feature near M_Δ. For example, events consisting only of visible particles and $E_T^{\rm miss}$ from mismeasurement would be expected to have an M_R distribution proportional to the distribution of CM energy $\sqrt{\hat{s}}$, as in the case of QCD backgrounds.

We then define a second mass variable that inherits knowledge of the mass splitting M_Δ, using the visible and invisible transverse momentum in the event. Note that this information was not used in the definition of M_R. Motivated by the fact that backgrounds with no invisible particles must have Q_1 and Q_2 back-to-back (a fact that mismeasurement dos not tend to greatly change), we define a transverse mass in terms of the visible transverse momenta, q_{1T} and q_{2T}, and the missing transverse energy $E_T^{\rm miss}$:

$$(M_T^R)^2 = \frac{1}{2}\left[E_T^{\rm miss}(q_{1T}+q_{2T}) - \vec{E}_T^{\rm miss}\cdot(\vec{q}_{1T}+\vec{q}_{2T})\right]. \tag{6}$$

Assuming pair production at threshold, $M_T^R \leq M_\Delta$ for signal events. Introducing the dimensionless ratio

$$R^2 = \left(\frac{M_T^R}{M_R}\right)^2, \tag{7}$$

we expect $R^2 < 1$ for signal events, with a rough spread around $R^2 \sim \frac{1}{4}$, while for background without real $E_T^{\rm miss}$ we expect $R \sim 0$.

The razor variables M_R and R^2 were originally designed to separate QCD and other backgrounds from pair production of strongly-interacting heavy particles [39, 40, 42, 43].[1] When used in these studies, all visible particles are assumed to fall into one of the decay chains of the parent particles S_1 or S_2. Therefore, all visible particles are assigned to a megajet Q_1 or Q_2 by a simple algorithm, and their momenta summed. Calculation of M_R and R^2 then proceeds as if there were only two visible objects.

B. The super-razor

Consider events that have both visible and invisible particles. Rather than splitting the visible particles into two objects Q_1 and Q_2, suppose we can divide them into three classes: particles (or groupings of particles) Q_1 and Q_2 that are assumed to come from the decay of the new physics particles S_1 and S_2, and a third class of particles that come from initial state radiation or something else extraneous to the heavy particle decays. In electroweak production of non-colored particles, every jet in an event can be assigned to this third class. The sum of the momenta of all particles in this class is $\vec{J}$. By construction

$$\vec{J}_T = -\vec{E}_T^{\rm miss} - \vec{q}_{1T} - \vec{q}_{2T}. \tag{8}$$

The effect of $\vec{J}$ is to shift the production frame by an additional boost that was not taken into account by the original longitudinal razor boost of Eq. (4). To correct for this, we want to make an additional transverse boost which takes us to the frame in which is recoiling against the jet contamination. The direction of this transverse boost is trivial: we must boost in the direction opposite to $\vec{J}$. However, there is insufficient information in the events at the LHC to unambiguously determine the magnitude of the boost. The correct boost from the lab frame to the pair production frame is

$$\vec{\beta}^{\rm CM} = \frac{\{-\vec{J}_T, p_z^{\rm CM}\}}{\sqrt{|\vec{J}_T|^2 + (p_z^{\rm CM})^2 + \hat{s}}}, \tag{9}$$

where $p_z^{\rm CM}$ is the z-momentum of the center of mass frame relative to the lab frame. Neither $p_z^{\rm CM}$ or $\hat{s}$ can be determined from the available visible particle momenta at the LHC.

---

[1] The initial use of the razor was in squark and gluino searches, where a major background is QCD. As QCD is essentially scale-free at LHC energies, the QCD background in the razor variable M_R falls exponentially. Requiring a minimum value of R^2, the background falls off more and more steeply as the R^2 threshold is increased. This "slicing away" of the background is the origin of the name "razor."

We therefore must make new assumptions to build our approximate boost to the frame R, the razor frame that is our best guess to the pair production frame. To build this approximate boost $\vec{\beta}_R$, we make the longitudinal boost β_L, and then construct an additional boost $\vec{\beta}_R$ from approximate center of mass energy $\sqrt{\hat{s}}_R$, defining

$$\vec{\beta}_R = \frac{\{-\vec{J}_T, p_z^R\}}{\sqrt{|\vec{J}_T|^2 + |p_z^R|^2 + \hat{s}_R}}. \tag{10}$$

There are two necessary assumptions to build $\hat{s}_R$. The first assumption is that the invariant mass of the visible system is equal to the invariant mass of the invisible system. This guess will result in $\hat{s}_R$ systematically lower than the actual $\hat{s}$ when the weakly interacting particles in the event are massive. Conveniently, this will actually turn out to be useful in our construction of further discriminating variables, which will be discussed shortly. The second assumption we must make is that the constructed variables (such as $\hat{s}_R$) do not depend on the unknown p_z^R. Clearly, this is not correct on an event-by-event basis, but allows for a determination of $\hat{s}_R$ to be made (up to a two-fold ambiguity, which we resolve by taking the positive solution). By requiring $\partial\sqrt{\hat{s}}_R/\partial p_z^R = 0$, we find (in terms of the razor variable M_R of Eq. (5))

$$\frac{\hat{s}_R}{4} = \frac{1}{2}\left(M_R^2 + \vec{J}_T \cdot (\vec{q}_1 + \vec{q}_2) + M_R\sqrt{M_R^2 + |\vec{J}_T|^2 + 2\vec{J}_T \cdot (\vec{q}_1 + \vec{q}_2)}\right). \tag{11}$$

This new variable $\hat{s}_R$ can be thought of as a "jet-corrected" version of the original razor variable M_R^2 (up to a factor of four). Which is to say, it inherits information about the mass difference M_Δ and the overall pair-production energy scale $\sqrt{\hat{s}}$.

In Figure 2, we show the distributions of M_R and $\sqrt{\hat{s}}_R$ (normalized to $\sqrt{\hat{s}}$) versus the p_T of the CM frame, for representative slepton pair production decaying to leptons and neutralinos. As can be seen, while both M_R and $\sqrt{\hat{s}}_R$ peak at the expected value given by the actual energy scale of the pair production ($\sqrt{\hat{s}}/2$ or $\sqrt{\hat{s}}$, respectively), when the center of mass is boosted to high p_T, the M_R variable begins to show deviations from the smooth distribution. Boosting against the jets corrects for the high p_T of the center of mass, as is seen in the distribution of $\sqrt{\hat{s}}_R$. The signal distributions are simulated using `MadGraph5` [51], `Pythia 6.4` [52], and `PGS`; complete details of our simulations and cuts are discussed in the next section.

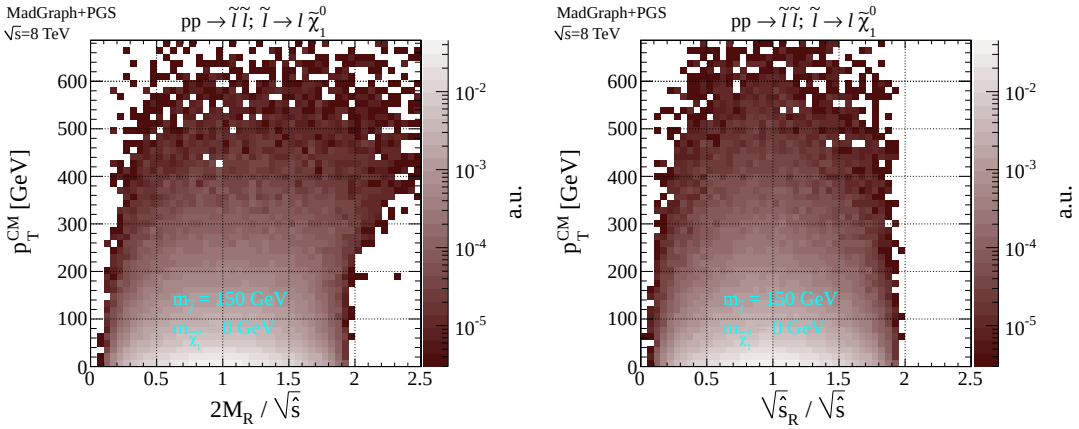

FIG. 2: Distributions of razor variable M_R normalized to $\sqrt{\hat{s}}$ (left) and $\sqrt{\hat{s}}_R$ normalized to $\sqrt{\hat{s}}/2$ (right) versus CM p_T compared for 150 GeV slepton pair production followed by decay to leptons and massless neutralinos,. See Section III for details of the simulation.

Interestingly, this variable $\hat{s}_R$ was constructed in Ref. [53], using a separate line of reasoning. In the razor framework, interpreting this invariant mass as the energy associated with a boost to an approximation of the pair production frame allows us to reconstruct that boost. As we will show, this leads to additional variables that add to our ability to distinguish signal and background.

Now that we are in the razor frame R, we can attempt to build boosts to approximations of the two decay frames of the parent particles S_i. Given the incomplete information available for the event, our choices for boosts are constrained. As there are two decay frames, which must have equal and opposite boosts from the pair production frame, we approximate the boost $\vec{\beta}^{\text{ decay}}$ by the boost

$$\vec{\beta}_{R+1} = \frac{\vec{q}_{R1} - \vec{q}_{R2}}{E_{R1} + E_{R2}}, \tag{12}$$

where q_{R1} and q_{R2} are the 4-momenta of the two visible particles Q_1 and Q_2 in the razor frame R. This boost has the correct symmetry property, in that the boost to the decay frame of S_2 is the negative of the boost to the decay frame of S_1.

If we correctly identified the boost $\vec{\beta}^{\,\text{decay}}$, then the invariant mass of the pair production frame would be related to the mass of the particles S_i by

$$\sqrt{\hat{s}} = 2\gamma^{\text{decay}} m_S. \tag{13}$$

We have constructed our boosts using information from the visible system Q_1 and Q_2, so our approximate boost $\vec{\beta}_{R+1}$ and approximate CM energy $\sqrt{\hat{s}}_R$ should be related not to the mass m_S, but the mass difference M_Δ. We therefore define a second razor variable M_Δ^R,

$$M_\Delta^R = \frac{\sqrt{\hat{s}}_R}{2\gamma_{R+1}}, \tag{14}$$

where γ_{R+1} is the Lorentz factor associated with the boosts $\vec{\beta}_{R+1}$. This variable should approximate M_Δ for signal events.

Clearly, building these razor frames requires many assumptions, approximations, and choices that may appear to be *ad hoc*. We take the attitude that this technique is justified if, in the end, we find variables that well-approximate the true values. In Figure 3, we plot the distributions of β_R for both the primary W^-W^+ background and slepton or chargino signal production, in all cases decaying to two charged leptons and missing energy. We also plot the boosts β_R normalized to the true transverse boost to the CM frame β_T^{CM}. The equivalent plots for β_{R+1} (including normalization to $\beta^{\,\text{decay}}$) are shown in Figure 4.

As expected for a proton-proton collider, the distributions of signal and background events all tend towards small boosts. For signal events, we see that we are systematically overestimating – albeit slightly – the magnitude of the boost β_R as compared to the true value β_T^{CM} . This effect is more pronounced when the splitting between the parent and daughter is small. We also mis-estimate the boost by a larger amount for charginos as compared to sleptons. This makes sense, as in constructing $\sqrt{\hat{s}}_R$, we made the assumption that the visible and invisible invariant masses are equal. This becomes increasingly incorrect as the invisible system's mass increases. The presence of extra invisible particles (neutrinos) in the chargino decays also will systematically skew that measurement. We will shortly take advantage of these systematic differences between the mass of the invisible system in the background and signal events to increase our discrimination power using a new set of variables.

In Figure 5 we plot the distributions of $\sqrt{\hat{s}}_R$ for a range of signal points and the W^-W^+ background, as well as the ratio of this razor variable to the quantity it is intended to estimate, $2\gamma^{\text{decay}} M_\Delta$ (for background, the splitting M_Δ is the W boson mass m_W, as the neutrino is massless for our purposes). As can be seen, the razor approximation is reasonably good, though systematically low for signal points with massive neutralinos, and less accurate for charginos than sleptons, for the reasons discussed previously. In Figure 6, we plot the distributions of the variable M_Δ^R, both by itself and normalized to the estimator value of M_Δ. The sharp edge at $M_\Delta^R = M_\Delta$ (seen most clearly in the slepton plot) indicates that this variable is useful in searches for new physics, especially in the regime where M_Δ is greater than the mass of the W boson.

Both $\hat{s}_R$ and M_Δ^R contain information about the mass splitting M_Δ for signal events. In Figure 7, we plot the two variables (normalized to the physical quantities they estimate). Two things can be seen from these plots. First, the variables $\hat{s}_R$ and M_Δ^R are not degenerate; though both estimate the same quantity (M_Δ), they contain independent kinematic information in that estimation. Secondly, we see that the scatter of the $\hat{s}_R$ around the true value is minimized near the edge structure of the M_Δ^R variable. This second piece of information will not be fully utilized in our analyses for computational simplicity, though it may provide a useful handle in the future.

As with the original razor variables M_R and M_T^R, one or both of these new razor variables could be used. However, we would ideally like a variables that encapsulated information not about M_Δ, but about the overall mass scale of the new particles in the event. This would help distinguish signal from background, especially in the cases where the mass difference is very small (*i.e.* parent and invisible daughter are nearly degenerate in mass), or when the mass difference approaches the mass of the W.

To try to capture more information about the event, we move beyond the mass variables already introduced and look at kinematic angles. In particular, we will be interested in the azimuthal angle between the razor boost $\vec{\beta}_R$ between the lab and R frames and the sum of the visible momenta $\vec{q}_1 + \vec{q}_2$, calculated in the razor frame R. An illustrative example of the relevant kinematics and angle definition is shown in Figure 8. We call this angle $\Delta\phi_R^\beta$, as it is the difference in azimuthal angle between the visible system and the boost $\vec{\beta}_R$, all defined in the razor frame R.

This angle is useful because it inherits information about ratio of masses of the pair produced particles and their invisible daughters, and so can be used in conjunction with a variable such as M_Δ^R or $\sqrt{\hat{s}}_R$, which have information

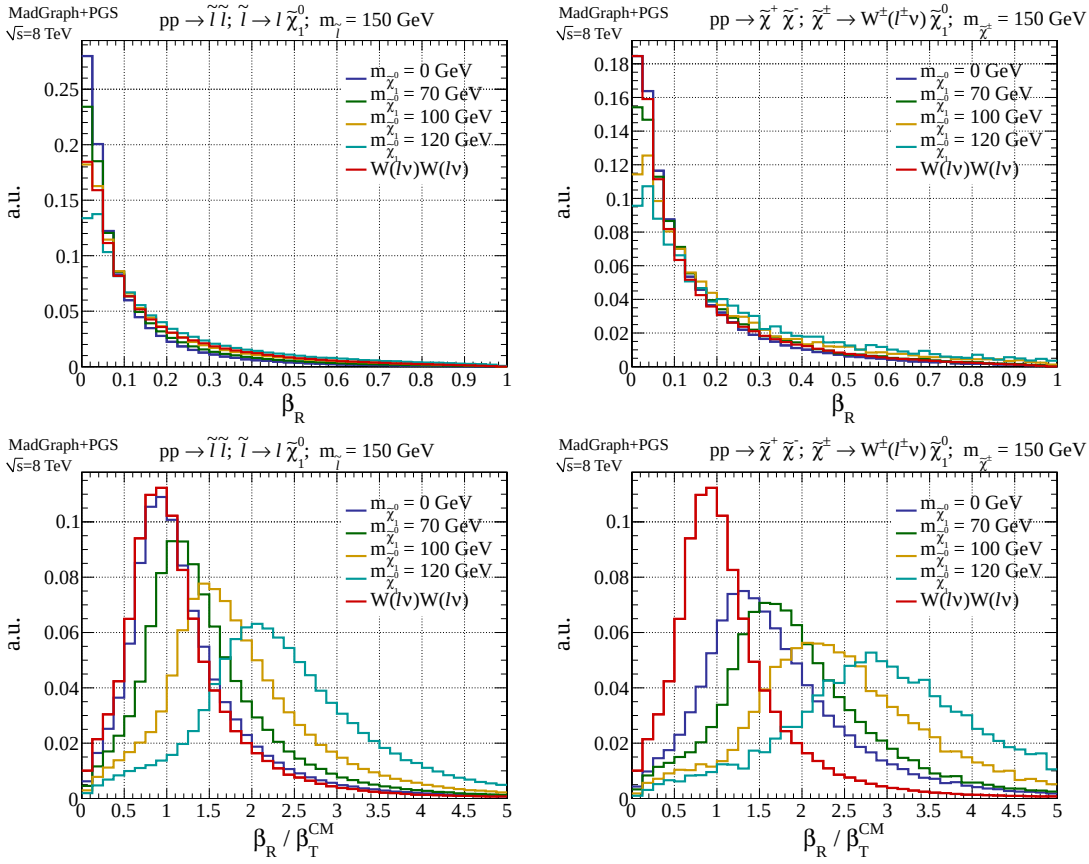

FIG. 3: Top Row: Distributions of β_R for 150 GeV selectrons (left) or charginos (right) decaying into neutralinos and electrons, for a range of neutralino masses. Also shown is the distribution of the W^-W^+ background. Bottom Row: Distributions of normalized $\beta_R/\beta_T^{\rm CM}$ (right) for 150 GeV selectrons (left) or charginos (right) decaying into neutralinos, again for a range of neutralino masses.

about the mass difference M_Δ, as previously discussed. The sensitivity of this angular variable to the ratio of masses actually comes from the previously discussed systematic shift of the variable $\sqrt{\hat{s}_R}$ relative to the mass difference M_Δ. As can be seen from Figures 3 and 5, our estimators of $\beta^{\rm CM}$ and $\hat{s}$ (β_R and $\hat{s}_R$), do not completely track the center of mass energy of the pair production. $\sqrt{\hat{s}_R}$, for example, is systematically smaller than $\hat{s}$, and β_R systematically larger than $\beta^{\rm CM}$. This behavior can be easily understood: it is due to the assumption that the energy of the event is evenly split between the visible and invisible systems. For events with invisible particles that are heavy compared to the parent, this assumption will underestimate the energy associated with the missing transverse momentum, and thus $\hat{s}_R$ is an underestimate of $\hat{s}$.

If $\hat{s}_R < \hat{s}$, then the boost $\vec{\beta}_R$ built using $\hat{s}_R$ will be systematically larger than the correct boost $\vec{\beta}^{\rm CM}$. In the CM frame, the distribution of the sum of the visible particles relative to the boost direction should be relatively flat. However, if we are "over-boosting" from the lab frame to the approximation of the CM frame, then the sum of the visible momenta will tend to be anti-aligned with the boost direction. That is, for systems where $m_\chi/m_S \ll 1$, we expect that the azimuthal angle between β_R and $\sum q_i$ will have a peak near $\Delta\phi_R^\beta \sim \pi$. In Figure 9, we show the distribution of this angle for a range of neutralino masses (for a fixed slepton or chargino mass). As can be seen, as the ratio m_χ/m_S approaches one, the peak of the distribution near π becomes more pronounced. Note the large drop in statistics for chargino events where the mass of the neutralino approaches that of the parent chargino. With such a mass spectrum, events have difficulty passing the selection criteria, which will be discussed in more detail in the next section.

Notice also from this figure that sleptons decaying to massless neutralinos are very similar to the $W^+W^- \to \ell^-\ell^+\nu\nu$ background. This is as expected, as the WW background is a case where the invisible particles (neutrinos) are massless, and so our estimate of $\hat{s}_R$ for this background will overboost to the R frame, just as with the massless signal case. Thus, we do not expect this angle to be of great use in the massless neutralino limit, however, it will be of significant help in distinguishing from background in the near-degenerate limit, where traditional mass variables sensitive to M_Δ

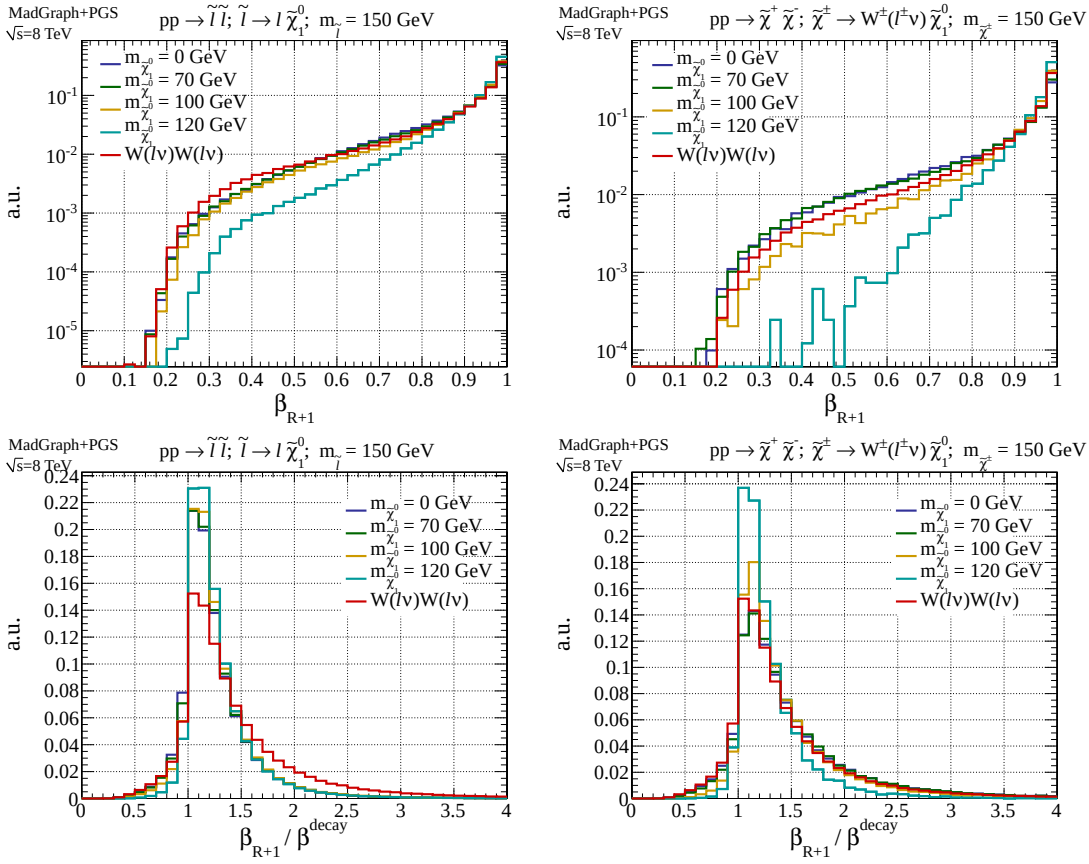

FIG. 4: Top Row: Distributions of β_{R+1} for 150 GeV selectrons (left) or charginos (right) decaying into neutralinos and electrons, for a range of neutralino masses. Also shown is the distribution of the W^-W^+ background. Bottom Row: Distributions of normalized $\beta_{R+1}/\beta^{\text{decay}}$ (right) for 150 GeV selectrons (left) or charginos (right) decaying into neutralinos, again for a range of neutralino masses.

are less effective. We also comment that the Drell-Yan $Z \to \ell\ell$ background, also shown in Figure 9, has a strong peak near $\Delta\phi_R^\beta \sim 0$. In this case, we are underboosting compared to the correct CM frame, as we are assuming that there is real missing transverse energy in an event that has no invisible particles.

In the R-frame, there is one final kinematic variable that we can construct. The variable $\sqrt{\hat{s}_R}$ is our estimate of the total energy available in the pair-production event. In the razor frame R, it can be divided up into three components:

$$\frac{\hat{s}_R}{4} = (M_\Delta^R)^2 + (q_{1R} + q_{2R})^2 + (E_{1R} - E_{2R})^2. \tag{15}$$

M_Δ^R and the invariant mass of the visible system $\sqrt{(q_{1R} + q_{2R})^2}$ have already been considered. However, the energy difference of the visible particles, $E_{1R} - E_{2R}$, has not been used. As with $\hat{s}_R$ and M_Δ^R, the overall mass scale of $E_{1R} - E_{2R}$ is sensitive to M_Δ, and is thus degenerate with our other mass variables. We therefore construct a new dimensionless variable

$$|\cos\theta_{R+1}|^2 = \frac{(E_{1R} - E_{2R})^2}{\hat{s}_R/4 - (M_\Delta^R)^2} = \frac{\hat{s}_R/4 - (M_\Delta^R)^2 - (q_{1R} + q_{2R})^2}{\hat{s}_R/4 - (M_\Delta^R)^2}. \tag{16}$$

This particular definition (and identification as a cosine of an angle) is because this variable can also be interpreted as the angle between the boost direction $\vec{\beta}_R$ and the direction of q_1 or q_2 in the frame $R+1$. However, it is more useful to think of this angle as a measure of the energy difference between the two visible particles.

A measure of the energy difference is useful in background rejection, especially in removing W^-W^+ events. The reasoning is as follows: for scalar particles, the decay of the parent into the visible and invisible daughters has a flat angular distribution in the parent's rest frame. In the production frame, we do not then expect a large correlation between the energy of the two visible particles. Though in their respective decay frames each has the same energy, the

orientation of their momentum relative to the momentum of the parent is uncorrelated, and so $|E_1 - E_2| \propto |\cos\theta_{R+1}|$ will not cluster at zero. The exception is for very large boosts of the parent particle; in this case, the direction of the visible daughter in the decay frame is effective erased by the very large boost. In such cases, both visible particles are colinear with their parent direction and have $E_1 \approx E_2$.

Now consider W^-W^+ background. Unlike scalar decay, the vector W boson decaying into fermions has a correlation in the direction of the visible lepton relative to the parent polarization. As the polarizations of the two W bosons in an event are themselves correlated, this means that, after the boost from the decay frames to the production frame (or to our approximation of that frame, the R frame), the two visible leptons will tend to have similar energies: $E_1 \approx E_2$. Therefore, the distribution of $|\cos\theta_{R+1}|$ for this background will be more highly peaked towards zero. The behavior of signal and background in this variable is shown in Figure 10, for representative signal points.

In Figure 11, we show the distributions of $|\cos\theta_{R+1}|$ with respect to M_Δ^R for a representative choice of slepton and neutralino masses, and compare with the distribution for the W^-W^+ background. As can be seen, when $M_\Delta^R \sim 0$, the signal events cluster near $|\cos\theta_{R+1}| = 0$. This makes sense, as $M_\Delta^R \sim 0$ corresponds to large boosts of the parent particles (as can be seen in Eq. (14)). As M_Δ^R approaches M_Δ, we recover the essentially flat distribution of $|\cos\theta_{R+1}|$ we expect from a scalar decay. The W^-W^+ background, on the other hand, does not have a flat distribution for $M_\Delta^R \sim m_W$. This, therefore, allows for discrimination of signal and background events even for signal events where $M_\Delta \sim m_W$.

These new angular variables demonstrates the utility of the razor boosts. The mass variables ($\hat{s}_R$ and M_Δ^R) are not completely unique to our work; they have been independently developed in different contexts in the past (see Ref. [53]). However, by associating these variables with a particular set of boosts, we can approximate the CM of the event. This allows us to build additional variables using this approximation, two of which ($\Delta\phi_R^\beta$ and $|\cos\theta_{R+1}|$) turn out to encode further information about the event. Furthermore, as the construction of the $|\cos\theta_{R+1}|$ variable relies on the spin of the new particles being searched for, it has the potential to be used as a measurement of spin if new physics is found. This possibility will be investigated in a future work.

In this study, we work primarily with the set of four variables $\hat{s}_R$, M_Δ^R, $\Delta\phi_R^\beta$, and $|\cos\theta_{R+1}|$. The two mass variables are somewhat degenerate, as both are estimators of the mass splitting between the parent and daughter particles. Though there may be some utility in using all four variables in a single analysis, here we will demonstrate the possible reach of our super-razor search by restricting ourselves to the M_Δ^R, $\Delta\phi_R^\beta$, and $|\cos\theta_{R+1}|$ combination only. We will also use γ_R and γ_{R+1} in our improved selection criteria to reduce background contamination We choose M_Δ^R over $\hat{s}_R$ for our analysis because, as Figure 12 shows, the variables M_Δ^R and $|\cos\theta_{R+1}|$ are approximately uncorrelated with $\Delta\phi_R^\beta$ for signal events. This simplifies the shape analysis we will discuss in the next section, as it allows us to decompose a 3D analysis into a 2D $\times$ 1D one.

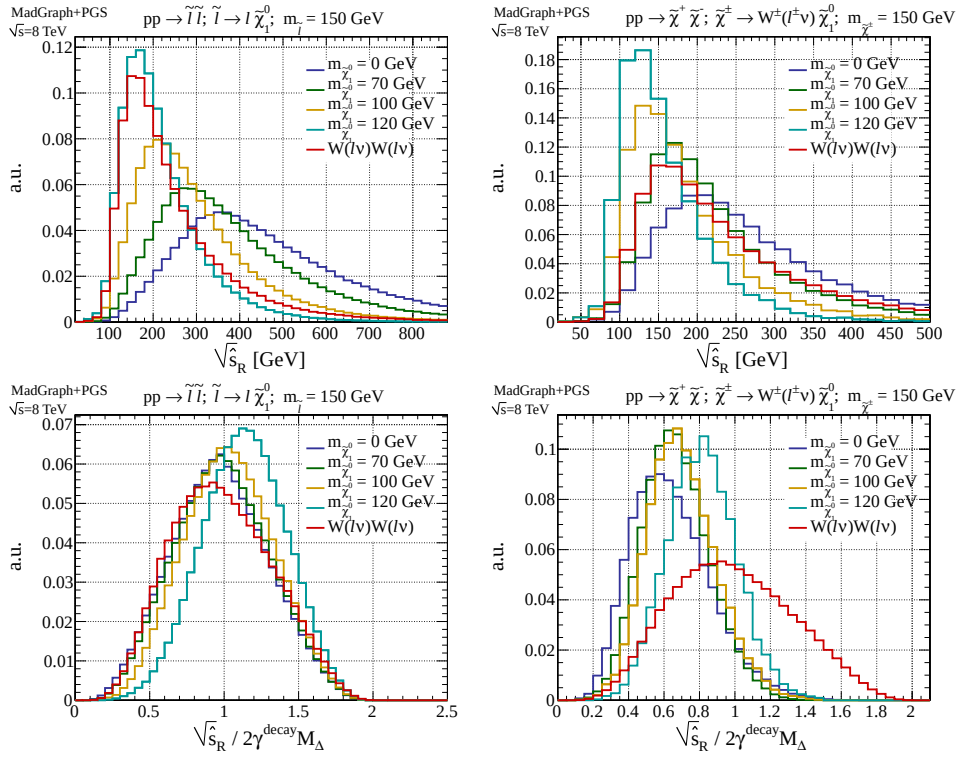

FIG. 5: Top Row: Distributions of $\sqrt{\hat{s}_R}$ for a 150 GeV slepton (left) or chargino (right) and a range of neutralino masses. Also shown is the distribution of the W^-W^+ background. Bottom row: Distributions of $\sqrt{\hat{s}_R}$ normalized to $2\gamma^{\rm decay} M_\Delta$ for selectrons (left) and charginos (right), again for a range of neutralino masses.

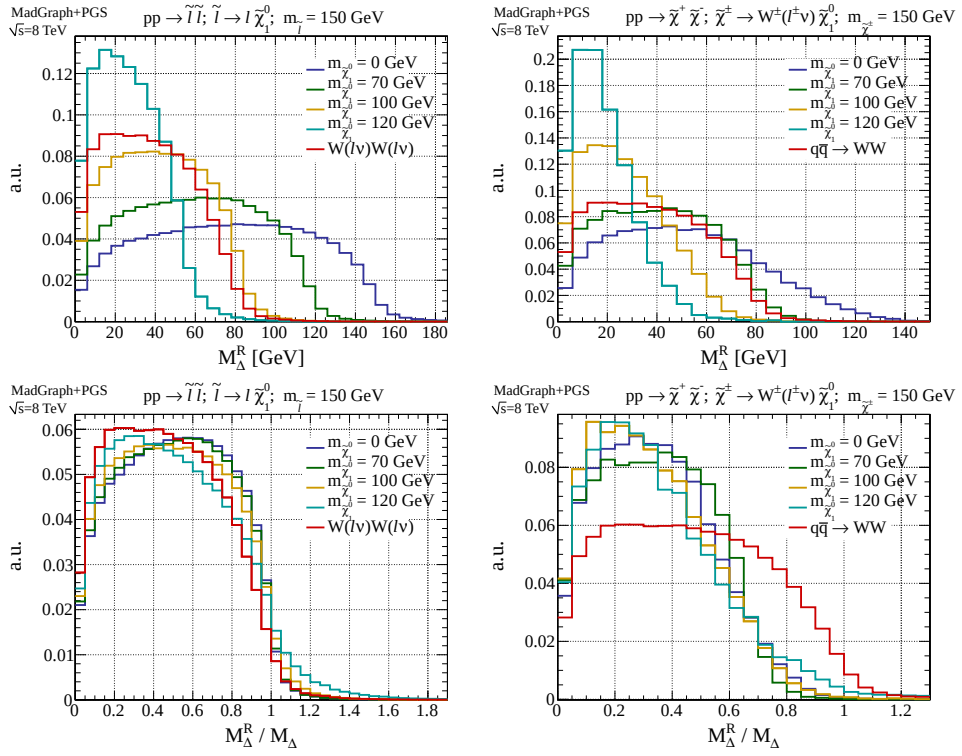

FIG. 6: Top Row: Distributions of M_Δ^R for a 150 GeV slepton (left) or chargino (right) and a range of neutralino masses. Also shown is the distribution of the W^-W^+ background. For the W background, $M_\Delta = m_W$. Bottom row: Distributions of M_Δ^R normalized to M_Δ for selectrons (left) and charginos (right), again for a range of neutralino masses

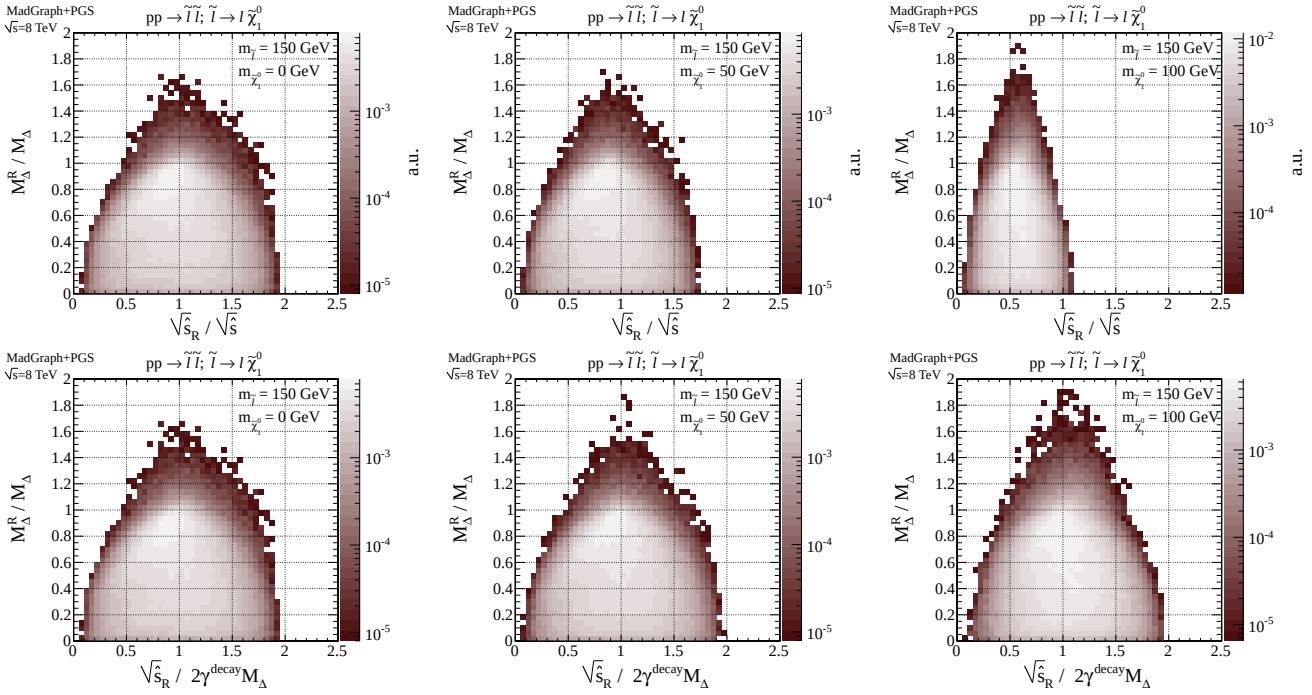

FIG. 7: Top Row: Distributions of $\sqrt{\hat{s}}_R/\sqrt{\hat{s}}$ vs. M_Δ^R/M_Δ for 150 GeV sleptons and a range of $m_{\tilde{\chi}}$) masses. Bottom Row: Distributions of $\sqrt{\hat{s}}_R/2\gamma^{\text{decay}}M_\Delta$ vs. M_Δ^R/M_Δ for 150 GeV sleptons and a range of $m_{\tilde{\chi}}$) masses.

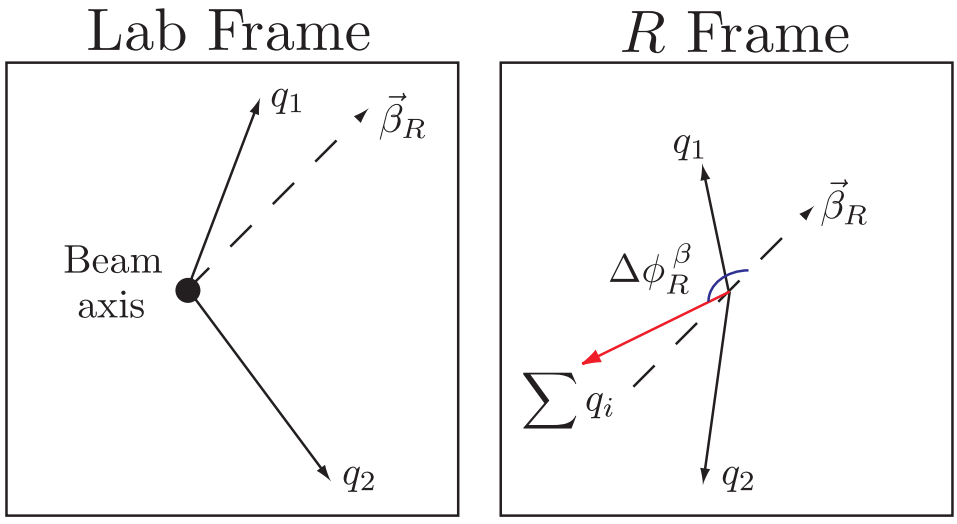

FIG. 8: Schematic example of the definition of the azimuthal angle $\Delta\phi_R^\beta$. The lab frame (seen here down the beam-line) contains two visible objects, q_1 and q_2. The direction of the boost $\vec{\beta}_R$ (defined in Eq. (10)), in the lab frame is also shown. In the frame R, arrived at by performing the boost $\vec{\beta}_R$, the visible momenta q_1 and q_2 are shown, along with their sum. The azimuthal angle between their sum $q_1 + q_2$ and the boost direction $\vec{\beta}_R$ in frame R defines $\Delta\phi_R^\beta$.

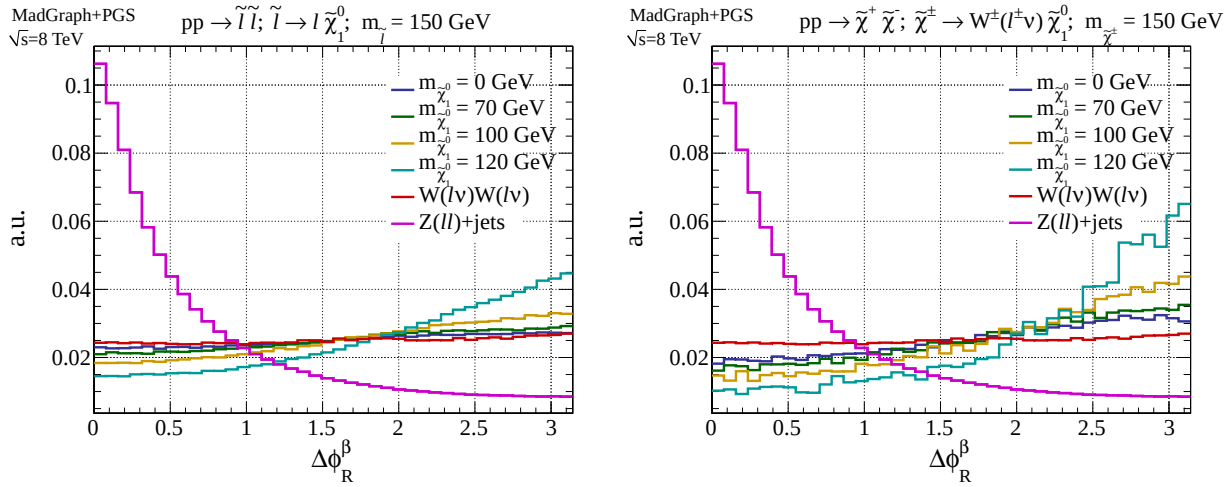

FIG. 9: Distributions of $\Delta\phi_R^\beta$ for a 150 GeV slepton (left) or chargino (right) and a range of neutralino masses. Also shown are the distributions of the W^-W^+ and Drell-Yan Z backgrounds.

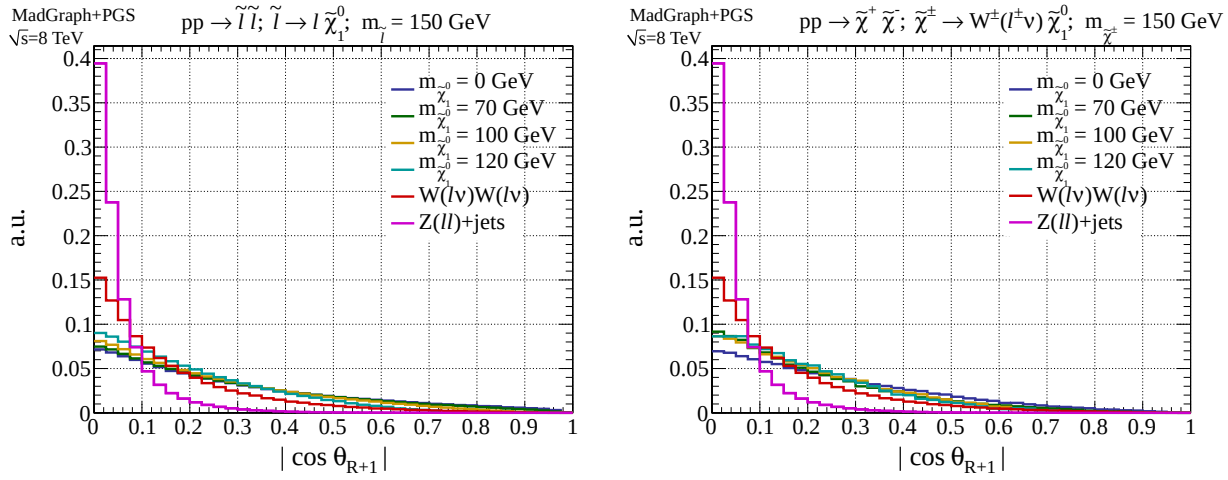

FIG. 10: Distribution of $|\cos\theta_{R+1}|$ for 150 GeV selectron (left) and chargino (right) pair production, decaying into a range of neutralino masses. Also shown are the W^-W^+ and Drell-Yan Z background distributions.

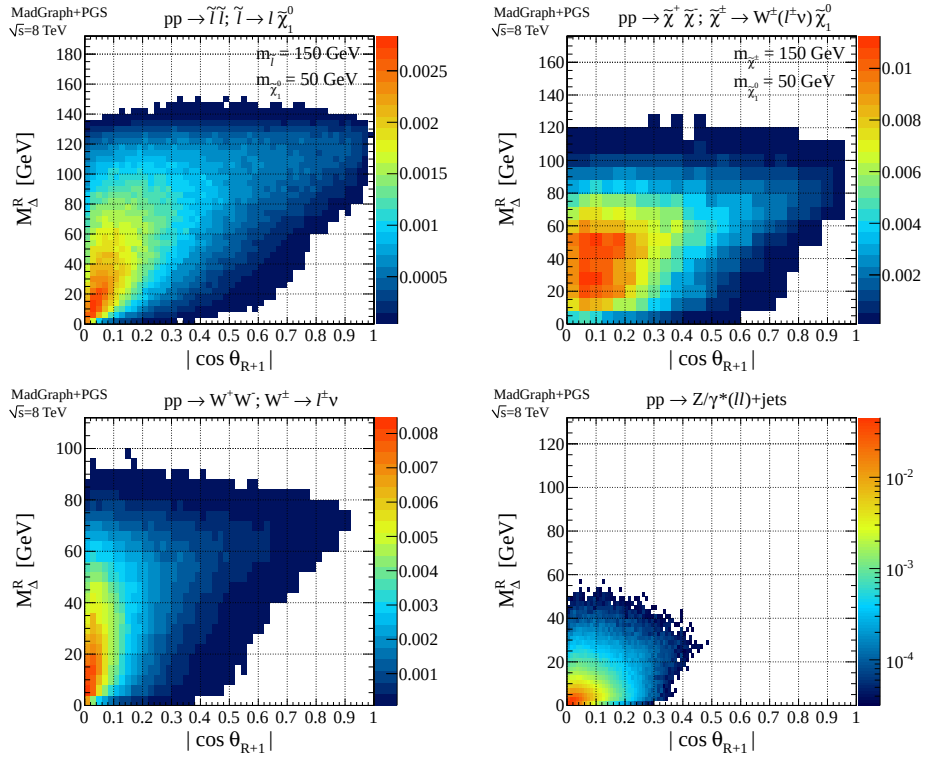

FIG. 11: Upper Row: Distribution of $|\cos\theta_{R+1}|$ versus M_Δ^R for 150 GeV selectron (left) or chargino (right) pair production, decaying into 50 GeV neutralinos. Lower Row: $|\cos\theta_{R+1}|$ versus M_Δ^R for W^-W^+ pair production (left) or Drell-Yan Z (right) backgrounds decaying into leptons.

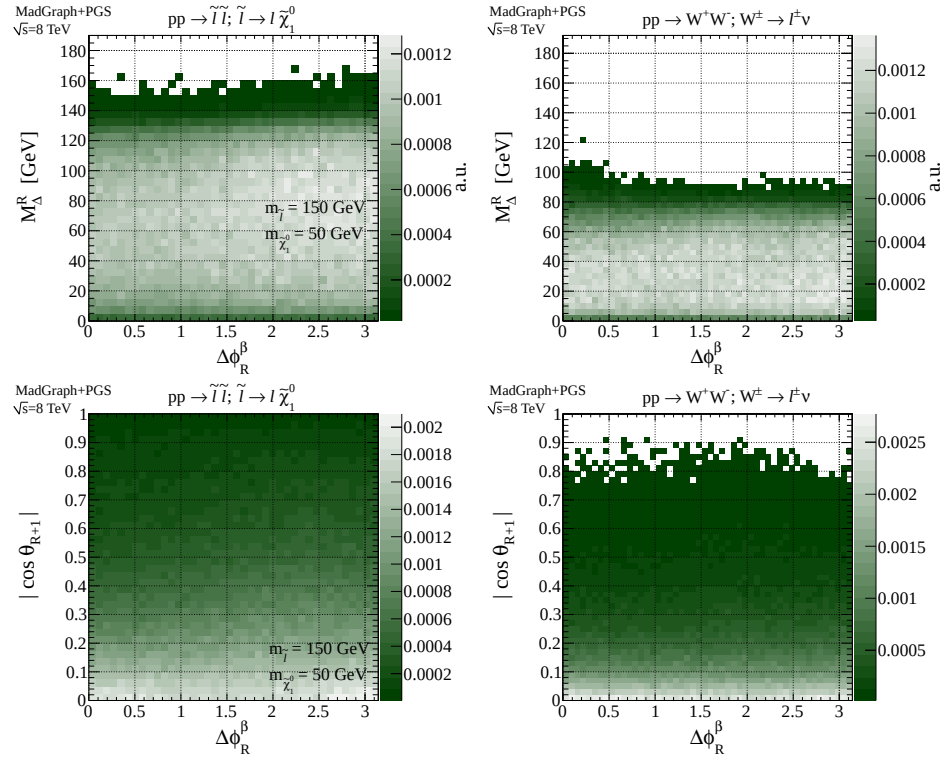

FIG. 12: Representative distribution of $\Delta\phi_R^\beta$ vs. M_Δ^R (top row) and $\Delta\phi_R^\beta$ vs. $|\cos\theta_{R+1}|$ (bottom row) for 150 GeV selectrons decaying to 50 GeV neutralinos (left) and W^-W^+ background (right).

III. EVENT SIMULATION AND SELECTION

A. Sample generation

We study the performance of the super-razor variables in the context of searches for new physics appearing in two different scenarios:

- Pair production of sleptons (selectrons $\tilde{e}^{\pm}$ or smuons $\tilde{\mu}^{\pm}$) decaying to electrons or muons, respectively, and neutralinos ($\tilde{\chi}_1^0$) with 100% branching ratio (BR) for both left- or right-handed sparticles.
- Pair production of the lightest chargino ($\tilde{\chi}_1^{\pm}$) decaying into $\tilde{\chi}_1^0$ and a W-boson with 100% BR. We require both W-bosons to decay leptonically while accounting for the SM W-boson branching ratio.

Event samples corresponding to these signal models were generated using `MadGraph5` [51] and `Pythia 6.4` [52], with up to two extra matched jets. We consider slepton and chargino masses between the LEP bound (~ 100 GeV) up to 450 GeV, with neutralino masses varying between zero and 20 GeV less than their respective parent sparticle masses. Production cross sections for these signal events were obtained for the LHC with $\sqrt{s} = 8$ TeV at next-to-leading order (NLO) using `Prospino` [54]. All superpartners except for the neutralino and active slepton flavor/chargino were decoupled by setting their mass to 2.5 TeV and the chargino was assumed to be wino-like. In the mass intervals considered, the resulting cross sections range from $\sim 100 - 1$ fb for both flavors of sleptons and $\sim 5000 - 100$ fb for the chargino, with the cross-sections as a function of sparticle mass illustrated in Figure 13.

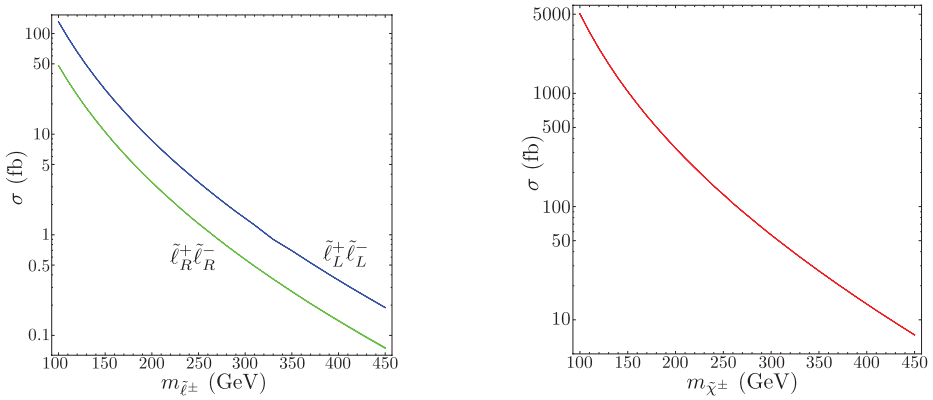

FIG. 13: Left: 1st or 2nd generation left-handed (blue) and right-handed (green) slepton pair production cross sections. Right: Chargino pair production cross section. Cross sections calculated using `Prospino` [54] at NLO for the LHC with $\sqrt{s} = 8$ TeV. Theoretical errors indicated by width of lines.

In order to estimate the sensitivity of the CMS and ATLAS experiments to these putative signals we also generate event samples corresponding to the primary SM backgrounds in the di-lepton final state: di-boson (W^+W^-, $W^{\pm}Z$, and ZZ) production, Drell-Yan ($Z/\gamma^* \to \ell\ell$)+jets, and top pair production. Event samples were generated for all channels in `MadGraph5`+`Pythia 6.4`, with up to two extra matched jets and cross-sections calculated from the same generator configuration.

B. Detector simulation and baseline selection

All of the event samples, for both signal and background processes, are analyzed using the PGS toy detector simulation, from which reconstructed leptons (electrons and muons) are identified and jets are clustered. For all the kinematic distributions and results presented in this work, simulated events are included only if they satisfy baseline selection requirements.

Each event is required to have exactly two reconstructed leptons with $p_T > 20$ GeV and $|\eta| < 2.5$. Events are discarded which have more than two leptons satisfying this requirement. Furthermore, these leptons are required to have opposite charge. The combination of these requirements reduces the yields of selected events corresponding to di-boson backgrounds such as WZ and ZZ where there are either more than two leptons reconstructed or the two leptons arise from the decays of different bosons. Events are assigned to one of three flavor categories corresponding to same flavor (SF) where there are either two reconstructed electrons (e^-e^+) or muons ($\mu^-\mu^+$) and opposite flavor

(OF), containing $e\mu$ events. An additional requirement of $m(\ell\ell) > 15$ GeV is applied to events falling in the SF categories in order to reject backgrounds with low-mass di-lepton resonances.

Jets are clustered from simulated calorimeter cells using FastJet [55] and the anti-$k(t)$ algorithm [56]. Events containing at least one jet with $p_T > 25$ GeV and $|\eta| < 2.5$ which is identified as b-tagged are discarded from the event sample in order to reduce the contribution from events containing top quarks. The number of reconstructed jets is used to classify events into one of three jet multiplicity categories: 0 jet, 1 jet and ≥ 2 jet. This jet counting scheme is based on jets with with $p_T > 30$ GeV and $|\eta| < 3$. Furthermore, events are discarded if either of the two reconstructed leptons falls within a cone of $\Delta R \equiv \sqrt{\Delta\eta^2 + \Delta\phi^2} = 0.4$ around any of the reconstructed jets in the event. Unless otherwise indicated, kinematic distributions include the sum of all three flavor and jet multiplicity categories.

C. Comparison of different kinematic variables

We evaluate the potential for the variable M_Δ^R to be used in a search for di-slepton and di-chargino production signals by comparing it with similar variables used in CMS and ATLAS searches. The CMS search for slepton production in the di-lepton final state [34] utilizes the variable $M_{CT\perp}$ [46, 47] while the analogous ATLAS analysis [32] includes requirements on the variable M_{T2} [44, 45] in definitions of signal regions sensitive to the presence of di-leptons following from slepton decays. The distributions of each of these kinematic variables, M_Δ^R, $M_{CT\perp}$, and M_{T2}, are shown in Figure 14 for slepton and chargino signals with various sparticle mass combinations.

The behavior of each of the three variables is similar. Each is sensitive to the quantity M_Δ for signal events, with a sharp edge or endpoint at the true value. The shape of each distribution is largely insensitive to the absolute value of M_Δ, such that distributions are nearly identical when scaled by M_Δ (differences are observed when the parent sparticle and the neutralino approach degeneracy). The similarities between these M_Δ sensitive variables are indicative of the fact that they are highly correlated and represent largely redundant information about events.

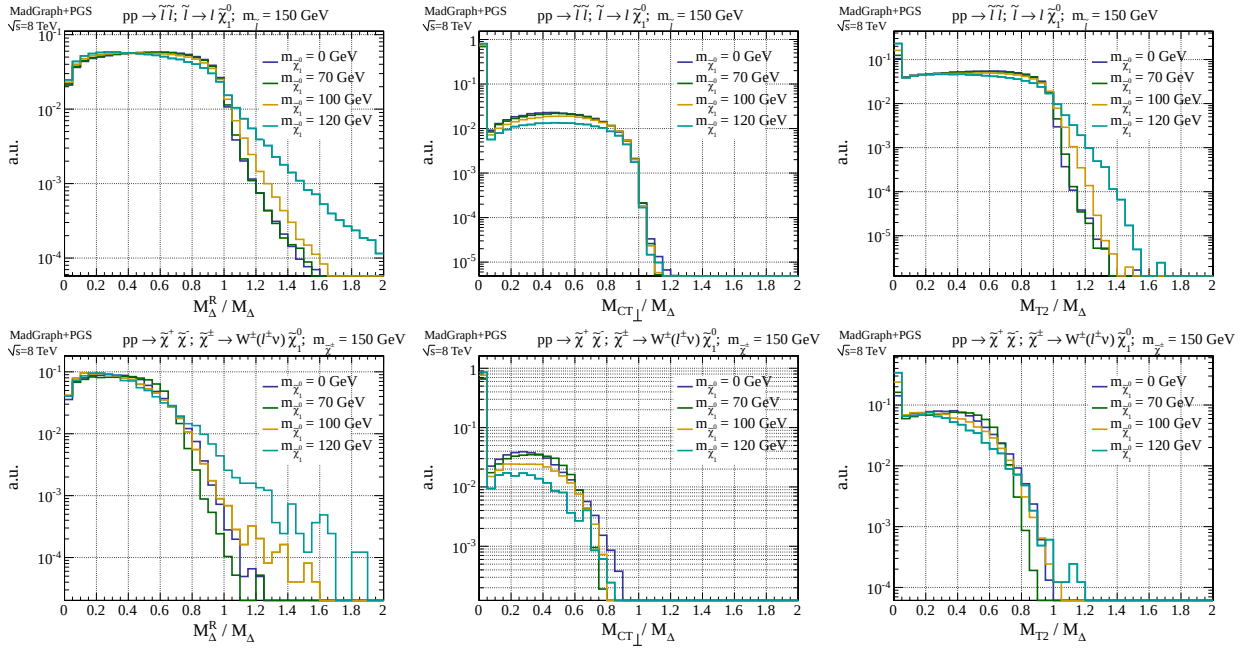

FIG. 14: Distributions of M_Δ estimating variables for sleptons (top row) and charginos (bottom row) with mass 150 GeV decaying into neutralinos and leptons, for a range of neutralino masses. Variables include M_Δ^R (left), $M_{CT\perp}$ (center) and M_{T2} (right), all normalized to the true value of M_Δ for each sample.

An important property of $M_{CT\perp}$ and M_{T2} is their almost complete insensitivity to the transverse momenta of the di-sparticle CM frame ($p_T^{\rm CM}$) in these events. Regardless of the velocity of the sparticles in the laboratory frame, the position of the M_Δ endpoint in these distributions remains largely unchanged. This property is convenient for interpretation of the putative signal distributions and essential in the construction of these searches, since it also guarantees the invariance of the same kinematic feature for backgrounds like WW and $t\bar{t}$, even for large $p_T^{\rm CM}$. For M_{T2}, under-constrained kinematic degrees of freedom are assigned through minimization which removes the $p_T^{\rm CM}$

dependence. Meanwhile, $M_{CT\perp}$ considers only the lepton kinematics along the transverse axis perpendicular to $\vec{p}_T^{\,\mathrm{CM}}$, largely ignoring variations which are sensitive to its magnitude.

For M_Δ^R the same behavior is achieved by explicitly correcting for non-zero p_T^{CM}, transforming the di-lepton system from the laboratory frame to an approximation of the CM frame. By using only Lorentz invariant information in the determination of this transformation, the definition of the resulting reference frame is stable under variations of p_T^{CM}, as are kinematic variables (such as M_Δ^R) evaluated in it. From Figure 14 is clear that the endpoint behavior of these M_Δ estimators has only mild sensitivity to the choice of strategy for removing p_T^{CM} dependence.

However, this choice does effect which events can be used to gain sensitivity to M_Δ, as can be seen in the presence or absence of an accumulation of events with a value of zero for each of these discriminants. By only considering information along one transverse axis in the event, $M_{CT\perp}$ is exactly zero around 50% of the time, corresponding to cases where the two leptons are moving in opposite directions along that axis. M_{T2} exhibits similar behavior, though with fewer events having $M_{T2} = 0$. This fraction of events are observed to vary with different signal mass combinations, and also with the magnitude of p_T^{CM} (see Figure 14). This latter dependence can be seen by comparing the M_{T2} distribution between different jet multiplicity categories, as shown in Figure 15, for each of the three M_Δ estimators, whereby larger jet multiplicity is generally correlated with larger p_T^{CM}. We observe that M_Δ^R does not exhibit this behavior.

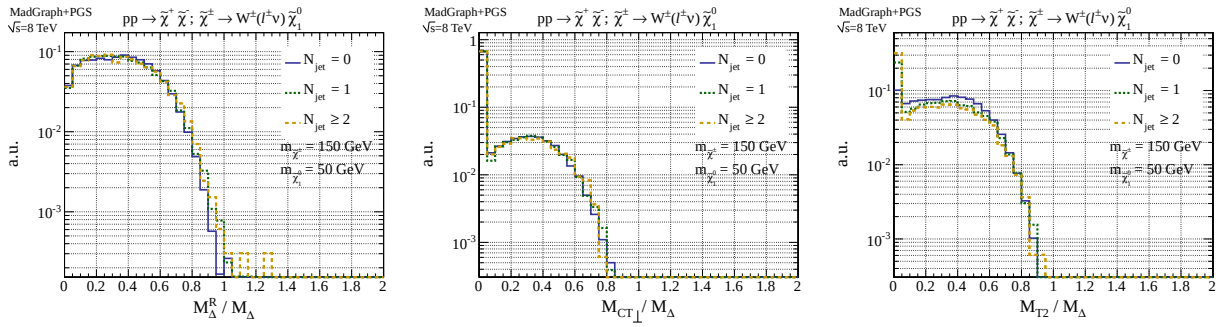

FIG. 15: Distributions of M_Δ estimating variables for charginos with mass 150 GeV decaying into 50 GeV neutralinos and leptonically decaying W bosons, as a function of reconstructed jet multiplicity. Variables include M_Δ^R (left), $M_{CT\perp}$ (center) and M_{T2} (right), all normalized to the true value of M_Δ for each sample.

The accumulation of events at zero in any of these kinematic variables has subtle effects on analysis that use them. Since the behavior is similar for both signal and background events, the ratio of the expected yields is largely insensitive to this effect. What does change is the size of the effective dataset that contains information about M_Δ. The 50% of events with $M_{CT\perp} = 0$ means that the integrated luminosity used in a search is effectively halved, while the increasing number of $M_{T2} = 0$ events at larger p_T^{CM} results in a similar effect for higher jet multiplicities and/or boosted topologies. The quantitative implications of these dependencies on both selection efficiency and ultimately expected sensitivity of searches are discussed in the following sections.

D. CMS and ATLAS-like event selections

At this stage we have only considered kinematic variables in the context of quite inclusive event selections. In practice, searches for new physics in di-lepton final states include multiple kinematic requirements, each designed to suppress particular backgrounds. These additional requirements often involve other discriminating variables, like E_T^{miss}, which can be highly correlated with the M_Δ estimators described in the previous section. Understanding the efficacy of any of these variables depends on the context of where in kinematic phase-space an analysis is searching. In order to evaluate whether using the variable M_Δ^R would yield an improvement in the sensitivity of the CMS and ATLAS searches we define CMS- and ATLAS-like event selections through which we attempt to capture the relevant qualitative features of the kinematic requirements enforced in these analyses.

In addition to the baseline selection requirements described in Section III B we consider criteria used by CMS and ATLAS (largely designed to reject Drell-Yan backgrounds):

CMS selection	**ATLAS selection**
$\lvert m(\ell\ell) - m_Z\rvert > 15$ GeV (SF channels)	$\lvert m(\ell\ell) - m_Z\rvert > 10$ GeV (SF channels)
$E_T^{\mathrm{miss}} > 60$ GeV	$E_T^{\mathrm{miss,rel}} > 40$ GeV

where

$$E_T^{\text{miss,rel}} = \begin{cases} E_T^{\text{miss}} & \text{if } \sin\Delta\phi_{\ell,j} \geq \pi/2 \ , \\ E_T^{\text{miss}} \times \sin\Delta\phi_{\ell,j} & \text{if } \sin\Delta\phi_{\ell,j} < \pi/2 \ , \end{cases} \tag{17}$$

and $\Delta\phi_{\ell,j}$ is the azimuthal angle between the lepton or jet closest in the transverse plane to $\vec{E}_T^{\text{miss}}$. Topologically, backgrounds like WW and $t\bar{t}$ are similar to the slepton and chargino signals, with two massive W bosons each decaying to a lepton and neutrino. As a result, the searches' ability to distinguish signal events from these backgrounds is dependent mostly on the difference between m_W and M_Δ of each signal, and is accomplished primarily through kinematic variables estimating M_Δ. The CMS and ATLAS selection requirements listed above are added specifically to reject Z/γ^*+jets events, using a Z mass window veto and E_T^{miss} related selections to eliminate events with mismeasured jets or leptons which result in spurious E_T^{miss}.

The CMS and ATLAS selection efficiencies for sleptons and charginos, as a function of sparticle masses, are summarized in Figure 16. Without explicit requirements placed on M_Δ-estimating variables, the selection efficiencies are relatively flat throughout most of the sparticle mass parameter-space, dropping quickly as parent sparticles and neutrinos approach mass degeneracy. This efficiency drop is a result of portions of the event selection involving energy scale, in this case minimum lepton p_T requirements and $E_T^{\text{miss}}/E_T^{\text{miss,rel}}$ cuts. For di-slepton production, the leptons are produced in two-body decays of the parent sparticle resulting in the lepton and neutralino momentum distributions scaling closely with M_Δ and a large efficiency gradient once energy scale requirements become comparable to the sparticle mass difference. For chargino models this gradient is less dramatic; with leptons produced in subsequent decays of W bosons (rather than in two-body decays of sparticles) the momentum distributions are broader, with weaker M_Δ correspondence. Furthermore, neutralinos are not the only weakly interacting particles in the final state and the total momentum of the entire system of weakly interacting particles is smaller as more energy is contained in the mass rather than the momentum. The result is that E_T^{miss} requirements can be especially inefficient for these signals relative to sleptons, with large drops in efficiency extending further from the mass degeneracy diagonal.

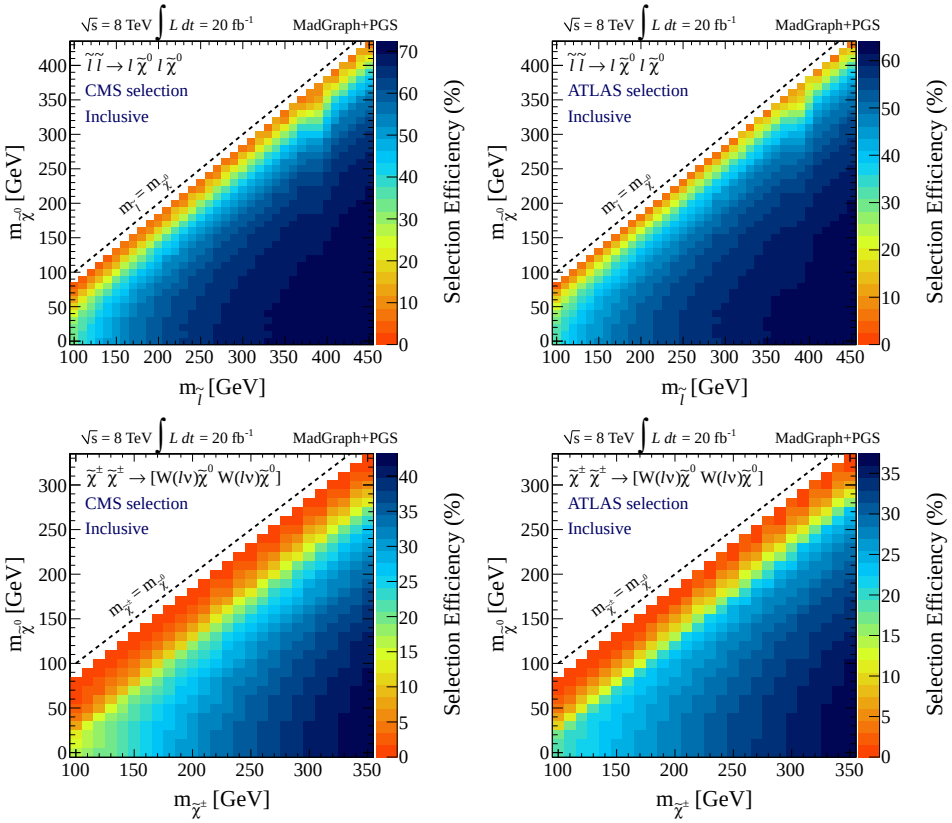

FIG. 16: Efficiency of the CMS (left) and ATLAS (right) selections for the slepton (top) and chargino (bottom) signal models. Selection efficiencies are calculated as a function of parent sparticle and neutralino mass and include all final state categories.

In practice, the sensitivity of search analyses in this final state will not scale exactly with this inclusive efficiency, due to large differences in background yields as a function of M_Δ-sensitive variables. Each of the largest backgrounds

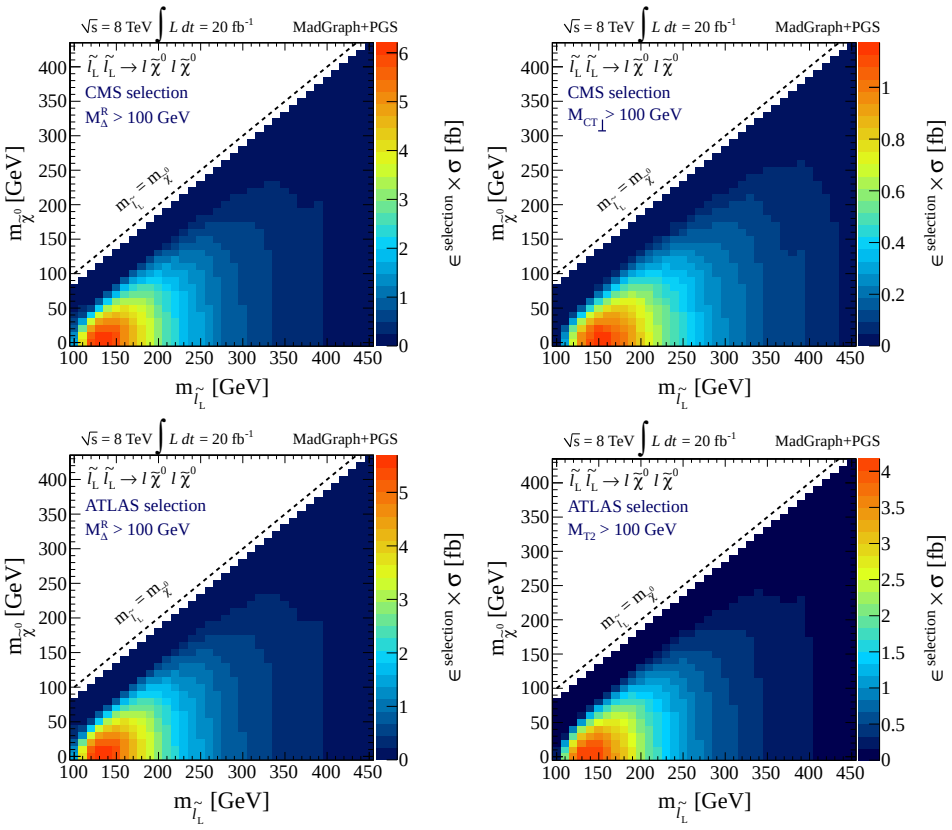

FIG. 17: Efficiency times cross section for slepton signal samples, as a function of neutralino mass, for the CMS (top) and ATLAS (bottom) selections with additional requirement that the mass sensitive variable (M_Δ - left, $M_{CT\perp}$ - top right, M_{T2} - bottom right) is in excess of 100 GeV.

after the CMS and ATLAS selections (WW and $t\bar{t}$), have M_Δ-sensitive variable distributions that have inherited information of the scale m_W. Therefore the sensitivity scales strongly with M_Δ, with significant experimental reach appearing only once M_Δ is in excess of the W mass. The effective cross-sections for signal models after the additional requirement that the M_Δ-sensitive variable used in each analysis is in excess of 100 GeV are shown in Figures 17 and 18 for sleptons and charginos, respectively. The expected sensitivity of the hypothetical searches described in the following sections closely follows these yields. Efficiencies and cross-sections for the SM backgrounds considered in these analyses are summarized in Table I.

E. Super-Razor selection without an $E_T^{\rm miss}$ cut

Kinematic variables sensitive to M_Δ can be powerful discriminants between slepton and chargino signals and SM backgrounds when M_Δ is much larger than the W mass, while heavy sparticle production with relatively compressed spectra can more easily remain hidden under large SM backgrounds. The angular variables introduced in Section II, $\Delta\phi_R^\beta$ and $|\cos\theta_{R+1}|$, are designed to address this deficiency. They are sensitive to quantities in events other than M_Δ: the ratio of daughter to parent mass and the spin correlations of decaying particles in the event. Thus they can be used to further discriminate between signal and background.

Each of the super-razor variables, M_Δ^R, $\sqrt{\hat{s}}_R$, $\vec{\beta}_R$, $\vec{\beta}_{R+1}$, $\Delta\phi_R^\beta$, and $|\cos\theta_{R+1}|$, represents a different piece of information about an event, and the collection can be thought of as a kinematic basis. Here we explore a new kinematic selection based on this basis, attempting to increase sensitivity to models with smaller values of M_Δ. In particular, we consider how one can remove explicit requirements on $E_T^{\rm miss}$. Included primarily to reject Drell-Yan background, such a requirement is inefficient for signal events at low M_Δ. Rather than attempting to determine an optimized set of cuts on the super-razor variables, we demonstrate how a selection criteria can be designed through simple choices for each variable based on the backgrounds we are attempting to reject.

We first consider the triplet of variables M_Δ^R, $\sqrt{\hat{s}}_R$, and γ_{R+1}, which for di-slepton production are meant to estimate

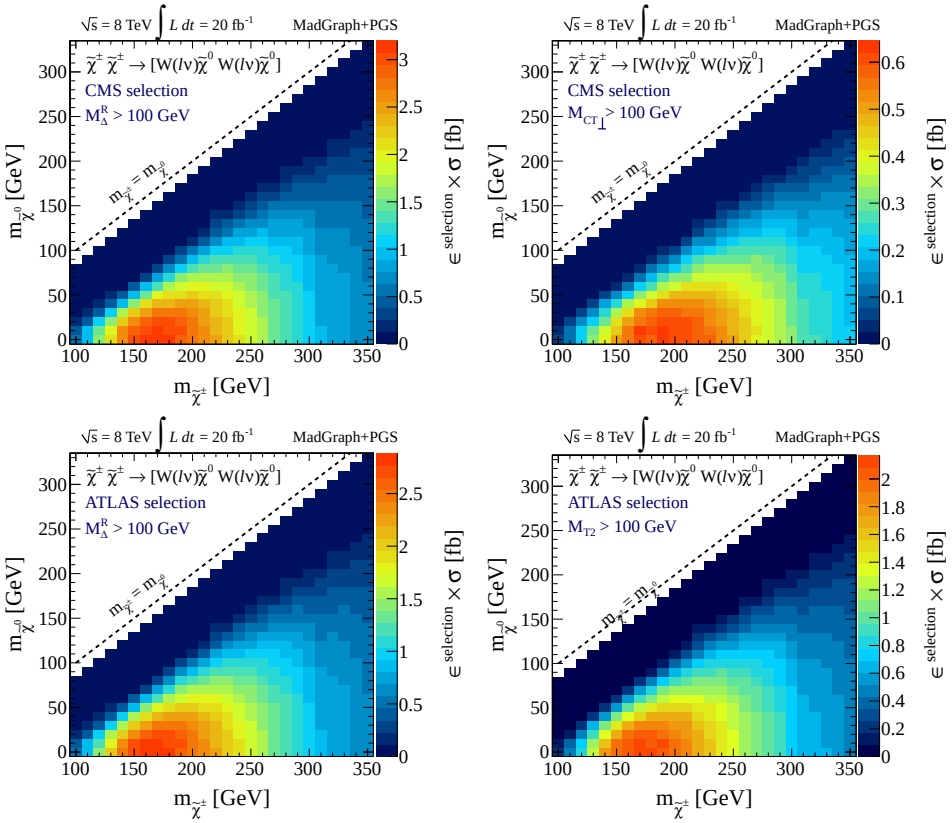

FIG. 18: Efficiency times cross section for chargino signal samples, as a function of neutralino mass, for the CMS (top) and ATLAS (bottom) selections with additional requirement that the mass sensitive variable (M_{Δ} - left, $M_{CT\perp}$ - top right, M_{T2} - bottom right) is in excess of 100 GeV.

$m_{\tilde{\ell}}$, $\sqrt{\hat{s}}$, and γ^{decay}, respectively. For both the true and reconstructed quantities the three variables represent only two unique pieces of information, since they are related by $\sqrt{\hat{s}}_R = 2\gamma_{R+1}M_{\Delta}^{R}$ and $\sqrt{\hat{s}} = 2\gamma^{\text{decay}}m_{\tilde{\ell}}$. Which two variables to consider depends on which signal and backgrounds are being investigated. For example, if searching for $H \to W(\ell\nu)W(\ell\nu)$ the variable $\sqrt{\hat{s}}_R$ will be resonant at the Higgs mass. If the Higgs is too light to accommodate two on-shell W's then the variables M_{Δ}^{R} and γ_{R+1} are more difficult to interpret, representing a combination of two W bosons with different masses. For the case of interest in this paper, non-resonant slepton pair production, off-shell sleptons are kinematically suppressed, and so M_{Δ}^{R}, $\sqrt{\hat{s}}_R$ and γ_{R+1} are all meaningful. However, they are not equally useful for discriminating between signal and background.

For di-slepton pair production the quantity that M_{Δ}^{R} is attempting to measure is effectively constant event by event, and is useful for discriminating against backgrounds. On the other hand, γ^{decay} varies between events, characteristic of non-resonant production, as does WW and $t\bar{t}$ backgrounds. As a result, $\sqrt{\hat{s}}_R$ provides information largely redundant with M_{Δ}^{R} while γ_{R+1} does not strongly discriminate against these large backgrounds.

The distributions of M_{Δ}^{R} as a function of $1/\gamma_{R+1}$, for simulated signal and background events, are shown in Figure 19. We observe that, for large $1/\gamma_{R+1}$, the two variables are largely uncorrelated for signal events. For small values of this variable, the signal M_{Δ}^{R} distribution collapses to zero, corresponding to the case where γ^{decay} is large and the decay leptons are largely back-to-back in the CM frame. For backgrounds, smaller values of $1/\gamma_{R+1}$ correspond to larger M_{Δ}^{R} on average. This is especially true for Z/γ^*+jets events; fixed di-lepton invariant mass corresponds to positively correlated lines in the M_{Δ}^{R} versus $1/\gamma_{R+1}$ plane, as seen in Figure 19. We observe that requiring $1/\gamma_{R+1}$ to be above some small value can remove many background events, especially Z/γ^*+jets, while only removing signal events that fall in an unremarkable region of phase-space (away from their M_{Δ}^{R} edge).

Resonant di-lepton production from Z/γ^*+jets is a particularly pernicious background. Associated jets produced in the event can boost the di-leptons to topologies that mimic those of di-slepton signals. Spurious E_T^{miss}, often from mis-measurements of jets, prevents this effect from being correctly accounted for. While an absolute requirement on E_T^{miss} can be used to remove many of these events, this also sets a lower bound M_{Δ} to which an analysis will be sensitive. Furthermore, Z/γ^*+jets with large E_T^{miss} that survive such a cut will tend to have large M_{Δ}^{R}, appearing in

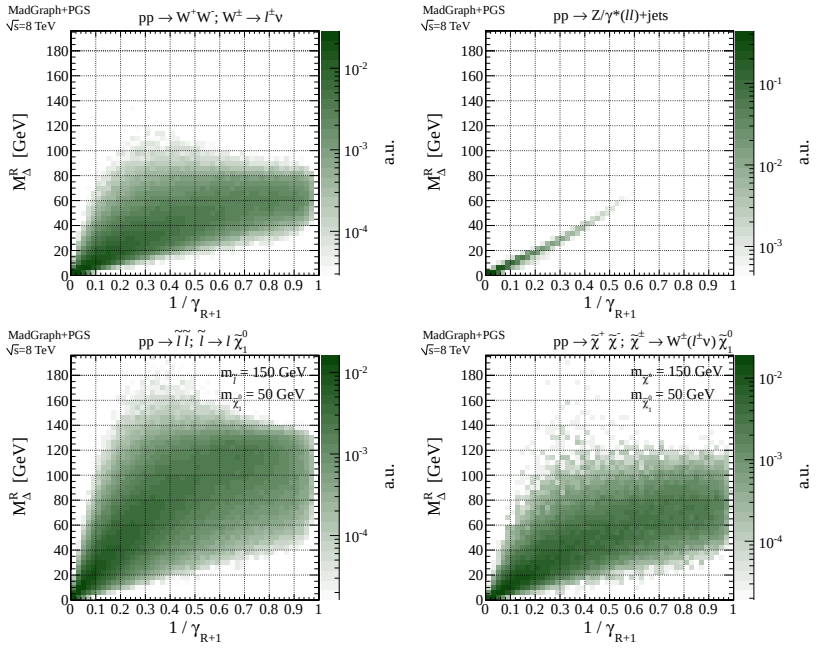

FIG. 19: Distributions of M_Δ^R vs. $1/\gamma_{R+1}$ for simulated events samples with baseline selection applied. Top left: WW. Top right: $Z/\gamma*$+jets. Bottom left: di-slepton production with $m_{\tilde{\ell}} = 150$ GeV and $m_{\tilde{\chi}_1^0} = 50$ GeV. Bottom right: di-chargino production with $m_{\tilde{\chi}_1^\pm} = 150$ GeV and $m_{\tilde{\chi}_1^0} = 50$ GeV.

the region of phase-space we had hoped to query for signal events. To replace a $E_T^{\rm miss}$ cut, we consider cuts based on scale-less variables, such that background events looking most like signal are also removed and we retain sensitivity to lower values of M_Δ^R.

The variable $\vec{\beta}_R$ is highly sensitive to the mis-measurements which make Z/γ^*+jets a difficult background to remove with $E_T^{\rm miss}$ cuts. This boost approximates the transverse portion of the Lorentz transformation from the lab frame to CM frame. Mis-calculations in the reconstruction of the direction and magnitude of this boost leave di-leptons from Z/γ^* decays in imbalanced configurations while they would be back-to-back in their true CM frame. The magnitude of $\vec{\beta}_R$ is not a strong discriminant (it is related to the ratio of CM system p_T and its reconstructed mass) while its direction is used in the calculation of $\Delta\phi_R^\beta$, as discussed in Section II.

Another angle that is useful to consider in diagnosing mis-measured Z/γ^*+jets events is the azimuthal angle between the $E_T^{\rm miss}$ and di-lepton system in the lab frame, $|\Delta\phi(\vec{p}_{\ell\ell}^{\rm lab}, \vec{E}_T^{\rm miss})|$, in particular for its correlation with $\Delta\phi_R^\beta$. The distribution of $|\Delta\phi(\vec{p}_{\ell\ell}^{\rm lab}, \vec{E}_T^{\rm miss})|$ is shown as a function of $\Delta\phi_R^\beta$ in Figure 20. For signal events the distribution of $|\Delta\phi(\vec{p}_{\ell\ell}^{\rm lab}, \vec{E}_T^{\rm miss})|$ is concentrated at π; the weakly interacting and di-lepton systems are back-to-back in the CM frame and will remain so in the lab frame without a large CM system transverse momentum. The distribution is more dispersed for $t\bar{t}$ events, as the W bosons are not only recoiling against each other in the CM frame, but also against two b-quarks. For Z/γ^*+jets the direction of the $\vec{E}_T^{\rm miss}$ is largely uncorrelated with the di-lepton system. The strength of the information contained in this two-dimensional plane can be seen when considering only events with $\gamma_{R+1} < 4$, *i.e.* those events which tend towards larger M_Δ^R values and are therefore of more significance in an analysis. We observe in the bottom part of Figure 20 that while the di-slepton and $t\bar{t}$ samples retain a similar shape after the $\gamma_{R+1} < 4$ requirement, the remaining Z/γ^*+jets exhibit a very particular correlation between $|\Delta\phi(\vec{p}_{\ell\ell}^{\rm lab}, \vec{E}_T^{\rm miss})|$ and $\Delta\phi_R^\beta$. The most difficult Z/γ^*+jets events, while still having a relatively flat $|\Delta\phi(\vec{p}_{\ell\ell}^{\rm lab}, \vec{E}_T^{\rm miss})|$ distribution tend to gather at low $\Delta\phi_R^\beta$. The correlation is such that a cut of $\Delta\phi_R^\beta + |\Delta\phi(\vec{p}_{\ell\ell}^{\rm lab}, \vec{E}_T^{\rm miss})| > \pi$ removes the majority of Z/γ^*+jets while keeping almost all the significant di-slepton events. This cut is indicated by the dotted red line in the bottom part of Figure 20, events being rejected if they fall below it.

As discussed in Section II, the variable $|\cos\theta_{R+1}|$ is also useful for rejecting Drell-Yan background, independently of the $\Delta\phi_R^\beta + |\Delta\phi(\vec{p}_{\ell\ell}^{\rm lab}, \vec{E}_T^{\rm miss})| > \pi$ requirement. Rather than include requirements on $|\cos\theta_{R+1}|$ in the selection we will use the full distribution as a descriminating variable in an analysis described in the following section. We define the Razor event selection as:

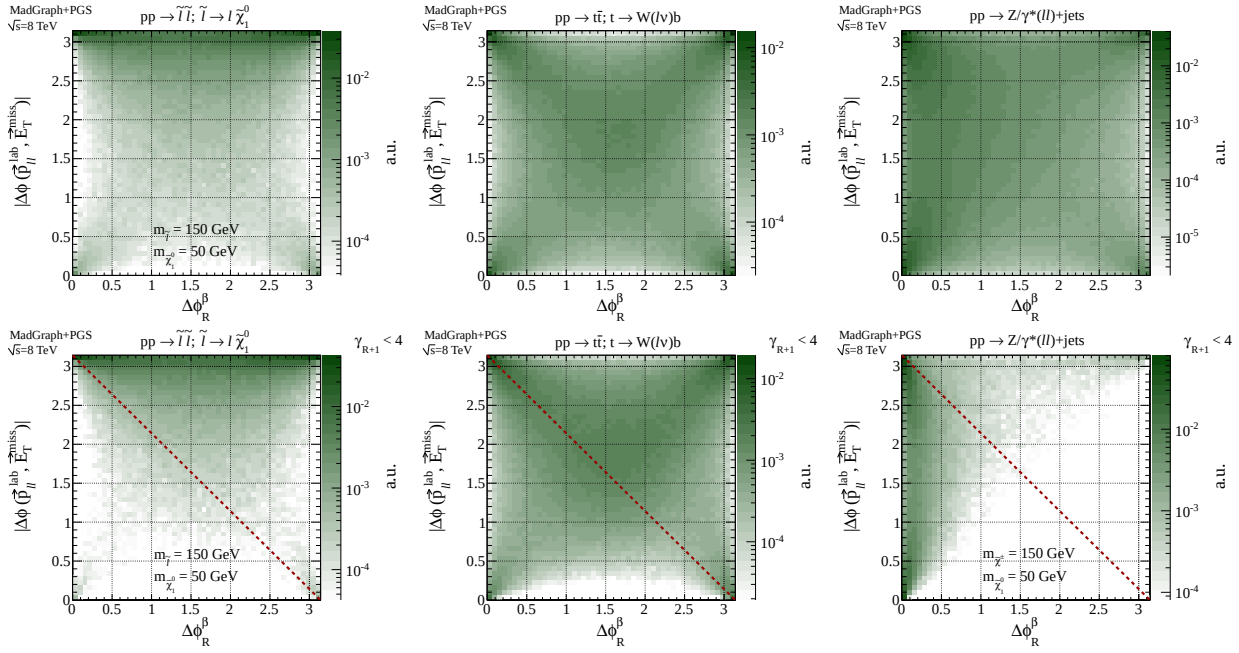

FIG. 20: Distributions of $|\Delta\phi(\vec{p}^{\,\rm lab}_{\ell\ell}, \vec{E}^{\rm miss}_T)|$ vs. $\Delta\phi^{\beta}_{R}$ for simulated events samples. Top row: inclusive baseline event selection. Bottom row: additional $\gamma_{R+1} > 4$ requirement. Samples correspond to di-slepton production (left), di-leptonic $t\bar{t}$ (center), and Z/γ^*+jets.

Razor selection:

SF Channels (ee,$\mu\mu$)	OF channels ($e\mu$)
$\gamma_{R+1} < 10$	
$\lvert m(\ell\ell) - m_Z\rvert > 10$ GeV	None
$\Delta\phi^{\beta}_{R} + \lvert\Delta\phi(\vec{p}^{\,\rm lab}_{\ell\ell}, \vec{E}^{\rm miss}_T)\rvert > \pi$	

Used in conjunction with $|\cos\theta_{R+1}|$, this simple event selection sufficiently reduces the Z/γ^*+jets background to a manageable level, without appealing to $E_T^{\rm miss}$ cuts that decrease selection efficiency for signals with lower M_Δ. There is likely room for optimization in these cuts, but this combination is sufficient for demonstrating that gains in sensitivity are possible for more compressed spectra, as we will see in the following sections. The efficiencies and expected cross-sections of event yields with the Razor selection applied for the slepton and chargino signal models considered are summarized in Figure 21. Analogous values for simulated background processes are provided in Table I. We observe that the efficiency for selecting low M_Δ events is improved over the CMS and ATLAS selections, while the number of Z/γ^*+jets at high M_Δ^R are reduced.

$\sigma \times \epsilon$ [fb]		CMS selection		ATLAS selection		Razor selection	
		Inclusive ($M_{CT\perp} > 100$ GeV)		Inclusive ($M_{T2} > 100$ GeV)		Inclusive ($M_\Delta^R > 100$ GeV)	
Process	Jet mult.	ee+$\mu\mu$	$e\mu$	ee+$\mu\mu$	$e\mu$	ee+$\mu\mu$	$e\mu$
Di-Bosons	0 jets	230 (0.067)	280 (0.068)	430 (0.15)	500 (0.15)	910 (0.55)	1300 (0.54)
	1 jet	120 (0.084)	150 (0.085)	120 (0.13)	142 (0.13)	170 (0.54)	380 (0.86)
	$\geq$ 2 jets	55 (0.044)	70 (0.061)	37 (0.056)	44 (0.059)	47 (0.36)	130 (0.78)
$t\bar{t}$	0 jets	38 (0.12)	46 (0.052)	31 (0.14)	34 (0.15)	65 (0.29)	95 (0.29)
	1 jet	140 (0.20)	180 (0.28)	110 (0.29)	120 (0.34)	180 (0.72)	340 (0.81)
	$\geq$ 2 jets	290 (0.46)	360 (0.57)	170 (0.50)	200 (0.58)	310 (1.4)	650 (1.8)
$Z/\gamma^*(\ell\ell)$	0 jets	70 (0.92)	4.8 (0.037)	160 (1.5)	3.3 (0.046)	1800 (1.6)	730 ($<$ 0.001)
	1 jet	85 (1.3)	37 (0.010)	70 (1.5)	5.2 (0.078)	110 (0.94)	110 ($<$ 0.001)
	$\geq$ 2 jets	44 (0.55)	27 (0.001)	13 (0.50)	2.4 ($<$ 0.001)	46 (0.24)	37 ($<$ 0.001)

TABLE I: Effective cross sections for di-lepton backgrounds at the LHC with $\sqrt{s} = 8$ TeV, after selection requirements (efficiency [ϵ] times cross section [σ]). Cross sections are listed for each of the event selections (CMS, ATLAS and Razor) as a function of jet multiplicity and lepton flavor, with and without selection requirements on mass sensitive variables.

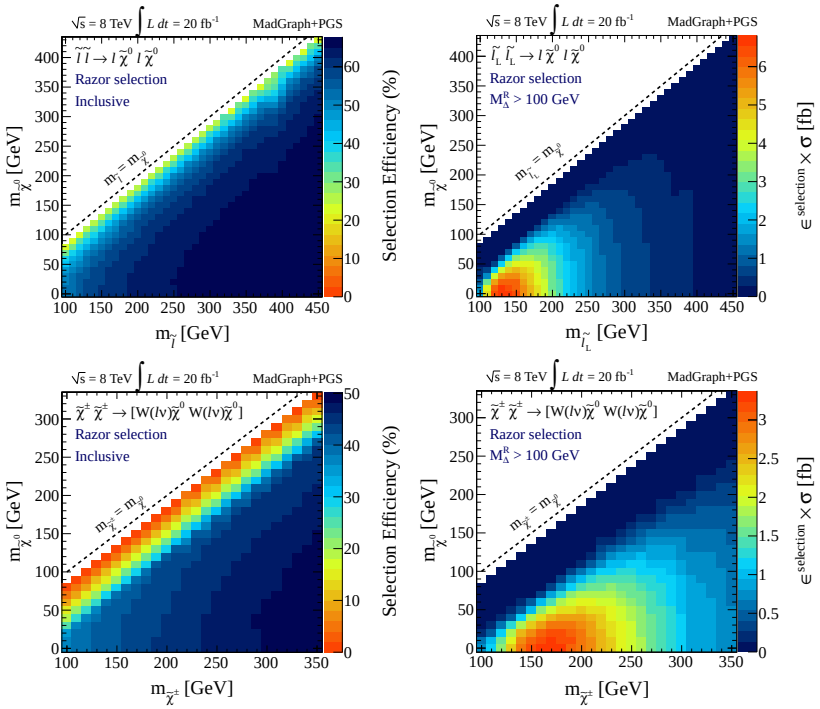

FIG. 21: Selection efficiencies (left) and efficiency times cross section (right) for left-handed selectrons (upper row) and chargino (lower row) signal samples, as a function of neutralino mass for the Razor selection criteria, described in the text.

IV. SHAPE ANALYSIS AND STATISTICAL APPROACH

In order to test the utility of the super-razor kinematic variables in the context of a search for slepton and chargino production we consider toy experimental analyses. Each of these analyses is a shape analysis, using multiple bins over the range of a kinematic variable of interest, and exploiting differences in changing signal and background expectations over the bins. This approach is used to increase the information being gleaned from these kinematic variables, allowing us to quantify what the maximal performance could look like, irrespective of changing optimized cuts associated with coarser binning. The predictions of these toy shape analyses are potentially optimistic relative to CMS and ATLAS results, due to both increased complexity of the analyses and the shortcomings of the detector simulation utilized here. To account for these differences, large systematic uncertainties are included in the procedure to represent potential experimental uncertainty in the relevant parameters that dictate the shape and yield of signal and background events.

A. Analysis strategy

For each toy analysis there are one or more kinematic variables identified as the discriminating variable, and the binned distribution of event yields in this variable are the observables in the toy experiment. The expected shape of both signal and background in the variable(s) of interest are required input for this procedure for each process. In our case, these shapes come from simulated event samples of each process. For an actual experimental analysis some can be measured or constrained from control regions. Regardless of their provenance, the uncertainties corresponding to these shapes are as important as the central values as we try to reflect in these toy analyses.

For the CMS and ATLAS analyses, control regions are identified using both object ID and kinematic information in order to isolate particular backgrounds. Z mass windows are used to select $(Z/\gamma^* \to \ell\ell)$+jets backgrounds for normalizing Z mass veto signal regions. Similarly, high jet multiplicity or b-tagged jet-enriched selections are used to constrain backgrounds with top quarks. In order to qualitatively capture these control region background constraints we consider multiple lepton flavor (ee, $e\mu$, $\mu\mu$) and jet multiplicity (0, 1, ≥ 2) categories simultaneously in a fit to data, with binned kinematic discriminants for each category. In each fit, high jet multiplicity categories effectively constrain top contributions while di-boson and Z/γ^* events at low jet multiplicity are disentangled using relative lepton flavor category yields.

We first consider one-dimensional analyses, where the kinematic discriminant is chosen to be M_Δ^R, $M_{CT\perp}$ or M_{T2}.

The distribution of the variable of interest is binned in 10 GeV steps from zero to 500 GeV. Only events satisfying the baseline selection and the relevant CMS (for $M_{CT\perp}$) or ATLAS (for M_{T2}) selection are included. The expected M_Δ^R and M_{T2} distributions in the ee final state for sample di-slepton signals and backgrounds are shown in Figure 22. We observe the changing background compositions and diminishing expected signal yield with increasing jet multiplicity. Distributions for each of the M_Δ-sensitive variables and selections considered are shown for the $e\mu$, $N_{jet} = 0$ final state in Figure 23.

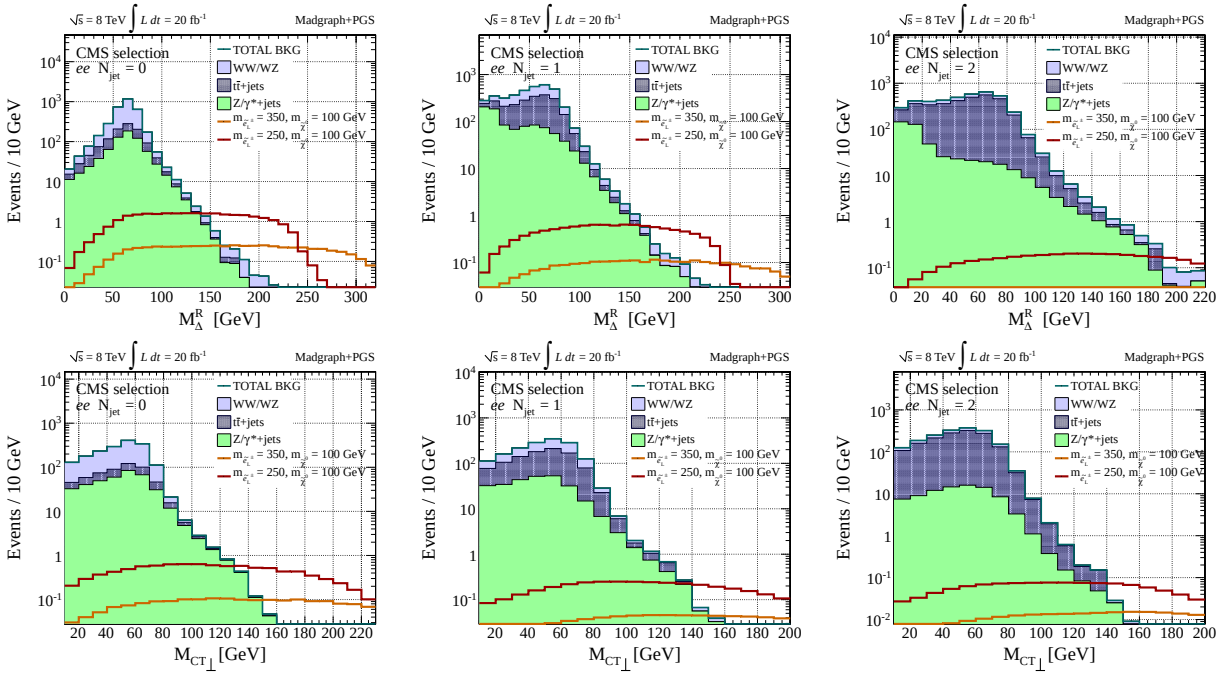

FIG. 22: Expected background yields in the ee final state passing the CMS selection, normalized to 20 fb^{-1} of data, for different jet multiplicities. Sample left-handed di-selectron signals are included with ($m_{\tilde{\ell}_L} = 350$, $m_{\tilde{\chi}_1^0} = 100$) and ($m_{\tilde{\ell}_L} = 250$, $m_{\tilde{\chi}_1^0} = 100$) GeV. Top: M_Δ^R distribution. Bottom: $M_{CT\perp}$ distribution. Left: $N_{jet} = 0$. Center: $N_{jet} = 1$. Right: $N_{jet} \geq 2$.

In addition to one dimensional shape analyses using the variables M_Δ^R, $M_{CT\perp}$, M_{T2} we also consider a three-dimensional analysis based on M_Δ^R, $|\cos\theta_{R+1}|$, and $\Delta\phi_R^\beta$. The two angular variables add complementary information to M_Δ^R; $\Delta\phi^\beta$ introduces sensitivity to the ratio of neutralino and parent sparticle masses while $|\cos\theta_{R+1}|$ helps further resolve the scale M_Δ of a particular sample while also adding discrimination against WW and $t\bar{t}$ using spin correlations (or lack thereof). Both of these angular variables are also useful in rejecting remaining Drell-Yan background events.

The three dimensional $M_\Delta^R \times \Delta\phi_R^\beta \times |\cos\theta_{R+1}|$ analysis uses the razor selection described in the previous section, and represents each of the kinematic discriminants as binned histograms. For both signal and background events we find that the variable $\Delta\phi_R^\beta$ has only weak correlations with the other two variables. We neglect any residual correlations such that $\Delta\phi_R^\beta$ distributions are modeled as a one dimensional histogram with five equal-width bins ranging from between zero and π. Examples of expected event yields for SM backgrounds and representative signal models are shown in Figure 24.

Strong correlations between M_Δ^R and $|\cos\theta_{R+1}|$ mean that these two variables cannot be factorized into one dimensional histograms. Rather, the two variables are modeled as two dimensional histograms with 10 GeV bins ranging from zero to 500 GeV for M_Δ^R (as for the one dimensional analysis) and 5 bins between zero and one for $|\cos\theta_{R+1}|$. The expected event yields for the sum of the SM backgrounds and sample signal models in this two dimensional binning are shown in Fig. 25

B. Fit to toy data and statistical analysis

For each dataset a fit is performed over all final state categories and bins of the kinematic discriminants simultaneously, measuring the yields of different background contributions. The fit proceeds by maximizing the binned

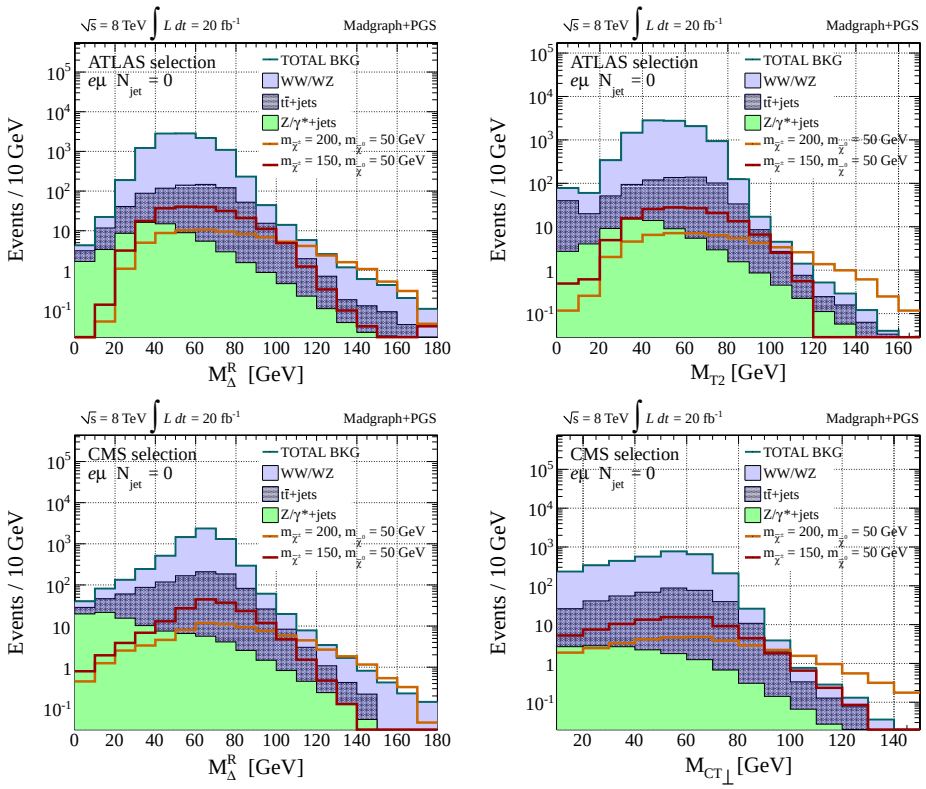

FIG. 23: Expected background yields in the $e\mu$, $N_{jet} = 0$ final state passing the CMS or ATLAS selections, normalized to 20 fb^{-1} of data. Top left: M_Δ^R with the ATLAS selection applied. Sample di-chargino signals are included with ($m_{\tilde{\chi}_1^\pm} = 200$, $m_{\tilde{\chi}_1^0} = 50$) and ($m_{\tilde{\chi}_1^\pm} = 150$, $m_{\tilde{\chi}_1^0} = 50$) GeV.Top right: M_{T2} with ATLAS selection. Bottom left: M_Δ^R with CMS selection. Bottom right: $M_{CT\perp}$ with CMS selection.

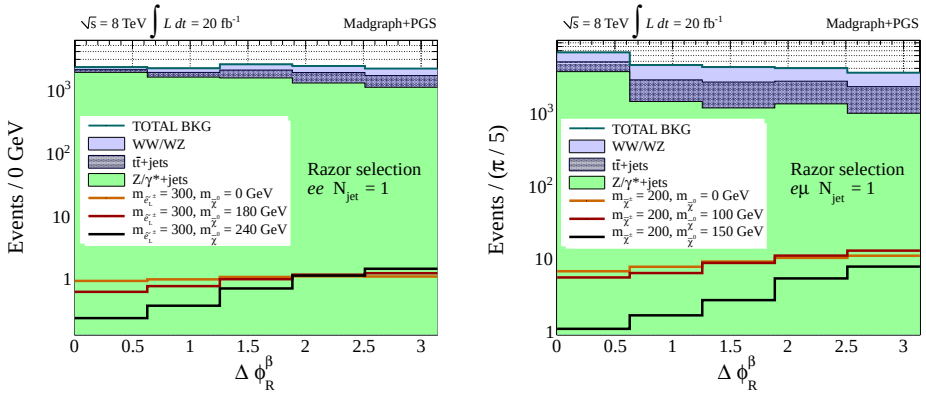

FIG. 24: Expected background yields in the $N_{jet} = 0$ final state passing the razor selection for $\Delta\phi_R^\beta$, normalized to 20 fb^{-1} of data. Left: ee final state including sample left-handed di-selectron signals with $m_{\tilde{\ell}} = 300$ GeV and varying neutralino masses. Right: $e\mu$ final state including sample di-chargino signals with $m_{\tilde{\chi}^\pm} = 200$ GeV and varying neutralino masses.

likelihood for the dataset being examined, which can be written as

$$\log \mathcal{L} = \sum_i \log \left(\frac{b_i^{n_i} e^{-b_i}}{n_i!} \right) , \tag{18}$$

where i runs over all of the bins and b_i and n_i are the expected and observed number of events in that bin, respectively. For each toy analysis fit the likelihood is maximized over the yields of each of the backgrounds, subject to constraints

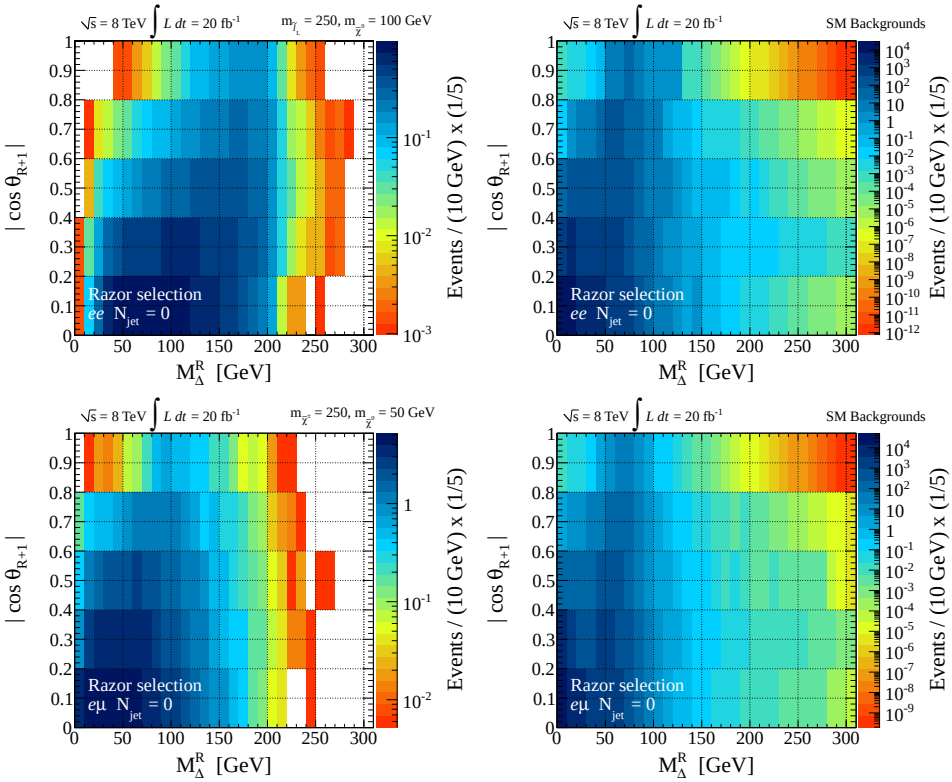

FIG. 25: Expected background yields in $M_\Delta^R \times |\cos\theta_{R+1}|$ plane, normalized to 20 fb^{-1} of data. Top: Expected event yields for the ee final state with $N_{jet} = 0$ and the Razor selection applied. Bottom: Analogous figures for the $e\mu$ final state. Left: Kinematic distributions for sample signal models including di-selectron production with ($m_{\tilde{l}_L} = 250$, $m_{\tilde{\chi}_1^0} = 100$) GeV (top left) and di-chargino production with ($m_{\tilde{\chi}_1^\pm} = 250$, $m_{\tilde{\chi}_1^0} = 50$) GeV (bottom left). Right: Total of expected SM background yields.

between bins so that the full likelihood can be written

$$\log \mathcal{L}[b_0, \cdots, b_{N_p}] = \sum_c \sum_k \left[n_{ck} \log \left(\sum_p b_p \hat{b}_{pck} \right) - \sum_p b_p \hat{b}_{pck} \right] , \tag{19}$$

where bins are now indexed by category (c) and kinematic discriminant bin (k). The total expected number of events for a single process p is b_p while $\hat{b}_{pck}$ is the fraction of events from process p expected to fall into bin ck, such that the number of expected events in a bin i from Equation (18), b_i, has become $\sum_p b_p \hat{b}_{pck}$. While the total normalization of each process is independent from the others, the probability distribution function (pdf) of each process, $\hat{b}_{pck}$, provides constraints between different categories and bins of the kinematic discriminant.

Two fits are performed on each dataset, one corresponding to the background-only hypothesis and the other to the signal plus background hypothesis, where the signal corresponds to whichever model is being tested. The background only fit can be represented as

$$\log \mathcal{L}_b = \max_{b_{\rm DB}, b_{t\bar{t}}, b_{\rm DY}} \mathcal{L}[b_{\rm DB}, b_{t\bar{t}}, b_{\rm DY}] \tag{20}$$

where $b_{\rm DB}$, $b_{t\bar{t}}$, and $b_{\rm DY}$ represent the normalizations for di-boson, $t\bar{t}$ and Drell-Yan backgrounds, respectively. Similarly, the signal plus background fit maximizes the likelihood

$$\log \mathcal{L}_{s+b} = \max_{b_{\rm DB}, b_{t\bar{t}}, b_{\rm DY}} \mathcal{L}[b_s = \hat{N}_S, b_{\rm DB}, b_{t\bar{t}}, b_{\rm DY}] \tag{21}$$

which differs from $\log \mathcal{L}_b$ in Equation (20) by the addition of a signal contribution with total yield b_s. This yield is not floated in the fit; rather, it is fixed to the expected number of signal events for a given model, $\hat{N}_S$. The two

maximized likelihoods, $\mathcal{L}_b$ and $\mathcal{L}_{s+b}$, are combined to form the test-statistic used to quantify the separation between the two hypotheses for a given model and dataset, the log-likelihood ratio λ

$$\lambda = \log\left(\mathcal{L}_{s+b}/\mathcal{L}_b\right) \tag{22}$$

Systematic uncertainties are included in this procedure through marginalization. In this scheme, the kinematic discriminant pdf shapes and normalizations used in the likelihood evaluation remain fixed at their nominal values. During the toy dataset generation process these same shapes and normalizations are systematically varied according to expected uncertainties. We consider several sources and qualitative types of systematic uncertainties. A 10% uncertainty is applied independently to each SM background process cross-section. The effect of this uncertainty is largely mitigated during the maximization of the likelihoods (where normalizations are floated). For backgrounds with multiple sub-contributions, like di-boson production, the relative sub-process yields are fixed in the likelihood evaluation resulting in an effective shape uncertainty. Each of the expected signal yields is also varied based on a calculation of the theoretical cross-section uncertainty.

In addition to overall normalization uncertainties there are also a collection of variations which change the shape of background pdfs, both by varying the relative yields in different final state categories and by altering the shapes of the kinematic discriminants themselves. A 2% uncertainty is assigned for the reconstruction and identification of each lepton, uncorrelated between lepton flavors. This uncertainty is assumed to be correlated between different processes. Similarly, a 10% uncertainty is assigned for the reconstruction and identification of each additional jet, effectively varying the relative yields between different jet multiplicity categories, independently for each process. This is meant to account for not only experimental effects relevant to jet counting, such as jet energy scale (JES) and resolution but also theoretical uncertainties in the production of strong emissions. To introduce uncertainty in the shape of kinematic discriminants we propagate the effects of potential JES uncertainties to the E_T^{miss} and kinematic variable calculation. For each simulated event all of the reconstructed jets, without a p_T threshold, are varied in p_T either up or down by 10%. The different between the original and new jet momenta is added vectorially to the $\vec{E}_T^{\text{miss}}$ and the kinematic variables of interest (for both selection requirements and kinematic discriminants) are recalculated using both the up and down variations separately. Each of the datasets corresponding to these variations are used to re-derive pdfs for the kinematic variables of interest such that the pdf shape for a given toy experiment is taken from a linear combination of the up, down and nominal templates.

For each signal model a series of toy pseudo-experiments are performed. For each pseudo-experiment, the parameters describing each of the systemic uncertainties are varied, yielding a new set of pdfs and normalizations for each process. These are then used to generate two toy data samples for the pseudo-experiment, with one set including the expected contribution of the signal in the data sample and another without the signal. Each dataset in the pseudo-experiment is then fit to each of the hypotheses (signal or no signal), yielding the maximized likelihoods $\mathcal{L}_b$ and $\mathcal{L}_{s+b}$ and the test-statistic λ. Repeating this procedure for many toys allows us to estimate the expected distribution of λ in the case that there is only background in the data sample, $P(\lambda|b \text{ only})$ and when there is also signal, $P(\lambda|s+b)$. Example distributions of λ for pseudo-experiments corresponding to representative signal models and analyses are shown in Figure 26.

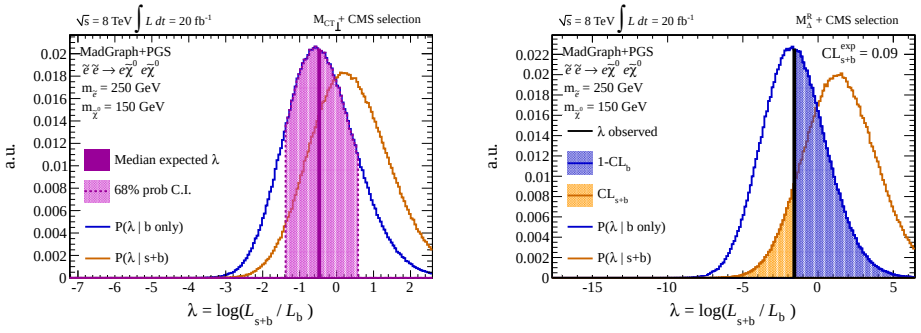

FIG. 26: Distributions of λ, assuming both signal and background-only scenarios. Left: Test-statistic distributions for the one-dimensional $M_{CT\perp}$ analysis searching for di-selectron production with ($m_{\tilde{\ell}_L} = 250$ GeV, $m_{\tilde{\chi}_1^0} = 150$ GeV). Right: λ distributions for the one-dimensional M_Δ^R analysis, using the same mass point.

The expected sensitivity of a given search is calculated from these test-statistic distributions. In order evaluate the probability of observing a given λ value in an experiment (λ^{exp}) the expected pdfs of λ are used to calculate the

quantities CL_b an CL_{s+b} as

$$\mathrm{CL}_b = \int_{-\infty}^{\lambda^{\mathrm{exp}}} P(\lambda|b \text{ only}) \ ,$$
$$\mathrm{CL}_{s+b} = \int_{-\infty}^{\lambda^{\mathrm{exp}}} P(\lambda|s+b) \ . \tag{23}$$

CL_b is the probability of observing a λ at least as background-like as λ^{exp} assuming that there is no signal contribution, while CL_{s+b} is the same probability assuming there is signal injected. In order to quantify the expected sensitivity of an analysis we choose λ^{exp} to be the median expected λ assuming it is distributed as $P(\lambda|b \text{ only})$. The resulting CL_{s+b} is then the median expected p-value for a given signal hypothesis, with lower values indicating that the model would be excluded at higher significance. These expected p-values are converted into a number of σ, corresponding to a normally distributed set of outcomes, as

$$N\ \sigma = \sqrt{2}\ \mathrm{erf}^{-1}(\mathrm{CL}_{s+b}) \ . \tag{24}$$

A particular model is expected to be excluded at 95% confidence level (C.L.) if the median expected CL_{s+b} is less than 0.05, and $N\sigma \geq 1.96$. The CMS and ATLAS experiments choose to quote results in the context of the CL_s convention [57, 58], where $\mathrm{CL}_s = \mathrm{CL}_{s+b}/\mathrm{CL}_b$. For median expectations, CL_b is exactly 1/2, implying that a $\mathrm{CL}_s \leq 0.05$ threshold for excluding a given hypothesis corresponds to a 97.5% C.L. exclusion, or $N\sigma \geq 2.24$.

The expected exclusions for di-slepton signals at 97.5% C.L. for analyses performed with 20 fb^{-1} of integrated luminosity at $\sqrt{s} = 8$ TeV, evaluated using this statistical approach, are shown in Figure 27. Comparing the excluded models from these toy analyses with those from the actual CMS [34] and ATLAS [32] searches we observe that the results are in reasonable agreement. The expectations from toy experiments tend to be more optimistic than the actual experimental results, which is expected given that kinematic discriminants are being used in a shape analysis and deficiencies in detector simulation likely correspond to underestimated resolution effects, particularly for E_T^{miss}. Regardless, this toy analysis framework allows for a quantitative comparison of different kinematic discriminants in the context of an analysis with realistic experimental effects at least partially accounted for.

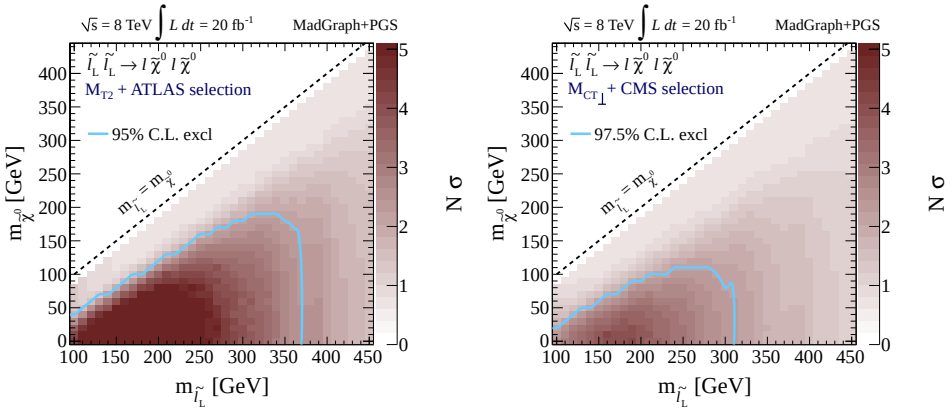

FIG. 27: Median expected number of σ for excluding the presence of different left-handed di-slepton signals, as a function of slepton and neutralino masses. Left: Expected results using M_{T2} with the ATLAS selection. Right: Expectations when using $M_{CT\perp}$ in conjunction with the CMS selection.

V. RESULTS AND DISCUSSION

Our 1D shape analyses using the mass variables M_Δ^R, $M_{CT\perp}$, and M_{T2} allow a fair and realistic comparison of their discriminating power. We begin by plotting the expected exclusion sensitivity for left-handed selectrons or charginos decaying to neutralinos, as a function of selectron/chargino and neutralino masses, assuming 20 fb^{-1} of data from a single experiment at the 8 TeV LHC. Charginos are assumed to decay into W bosons and an invisible neutralino, followed by Standard Model decays of the W bosons into leptons. Results for left-handed smuons would be similar to those for the selectron, but we assume only a single species of slepton for our analysis. In Figures 28 and 29, we show the expected exclusion reach (at 95% confidence level) of the ATLAS M_{T2} and CMS $M_{CT\perp}$ analyses compared to the new technique using M_Δ^R. In making the comparisions we use the same sets of ATLAS or CMS cuts as the existing experimental searches, which are not optimized for our analysis. Even with this disadvantage the expected exclusion limits using the super-razor variable M_Δ^R outperform the $M_{CT\perp}$ searches in terms of both absolute slepton or chargino mass and near the degenerate limit (when the mass of the parent is close to the mass of the invisible daughter). We show selected slices of these analyses in Figure 30, fixing either the selectron or neutralino mass, and varying the other. This allows a more direct comparison of our new variable M_Δ^R to the alternative techniques. Again the sensitivity using M_Δ^R outperforms that obtained from $M_{CT\perp}$. For these 1D analyses the performance using M_{T2} is only slightly worse than that obtained with M_Δ^R.

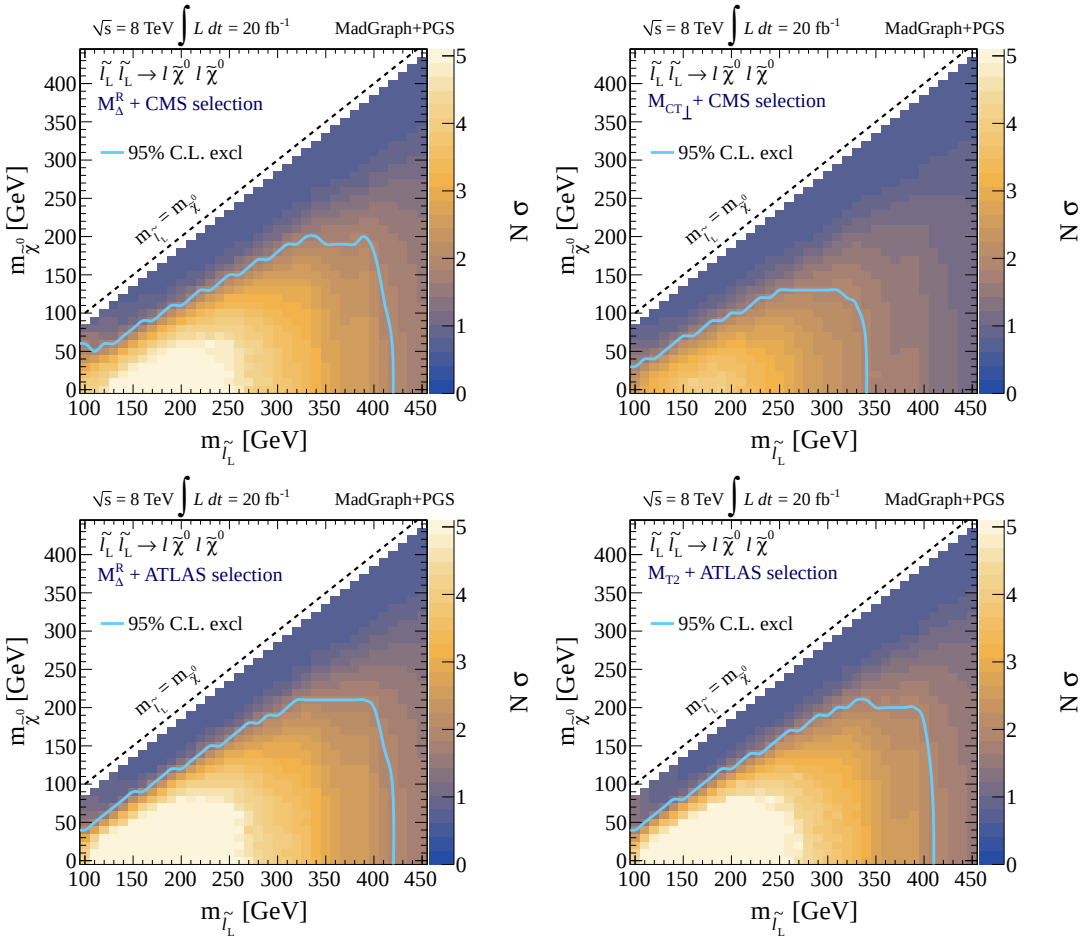

FIG. 28: Expected exclusion limits (in units of σ) for left-handed selectrons decaying to leptons and neutralinos using 20 fb^{-1} of 8 TeV data, as a function of both selectron and neutralino masses. Expected limits are shown for our 1D M_Δ^R analysis using CMS (upper left) and ATLAS (lower left) selection cuts, and directly compared to our expected exclusions using our simulated CMS $M_{CT\perp}$ (upper right) and ATLAS M_{T2} (lower right) analyses.

We can understand these 1D results by again consulting the kinematic distributions shown in Figure 14 of Section III. The fact that approximately 50% of signal events end up in the zero bin for $M_{CT\perp}$ gives a loss in statistics that is not compensated by the clean kinematic edge. For M_{T2} the corresponding effect is much smaller, resulting in

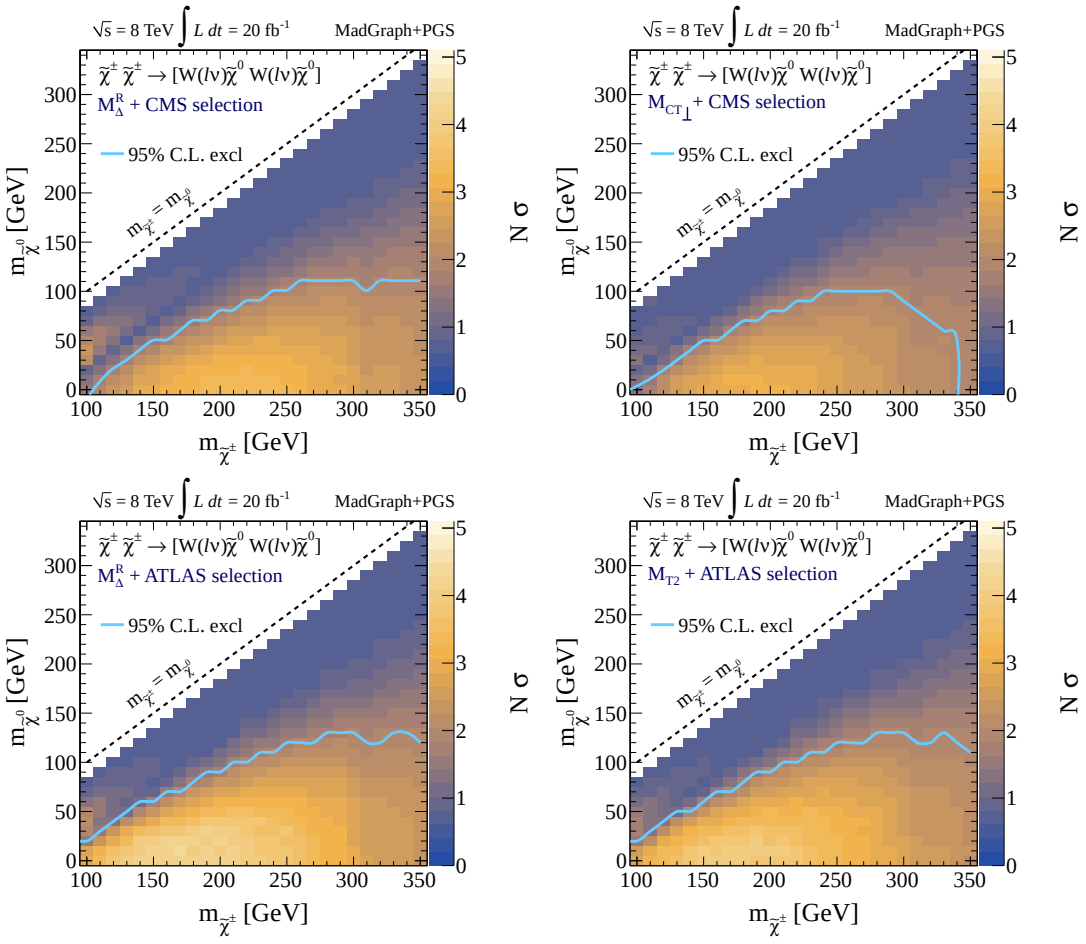

FIG. 29: Expected exclusion limits (in units of σ) for charginos decaying to neutralinos and leptonic W bosons using 20 fb^{-1} of 8 TeV data, as a function of both selectron and neutralino masses. Expected limits are shown for our 1D M_Δ^R analysis using CMS (upper left) and ATLAS (lower left) selection cuts, and directly compared to our expected exclusions using our simulated CMS $M_{CT\perp}$ (upper right) and ATLAS M_{T2} (lower right) analyses.

performance very similar to that achieved with M_Δ^R.

In Figure 31, we show the exclusion reach of the full super-razor analyses, using our multi-dimensional shape analysis which employs M_Δ^R, $\Delta\phi_R^\beta$ and $|\cos\theta_{R+1}|$, and the new super-razor selection described in Section IV in order to maximize the sensitivity over background. Exclusions are shown for both left- and right-handed selectrons, as well as charginos decaying to W bosons and neutralinos. The exclusion sensitivities include the effects from systematic errors on kinematic shapes, and on reconstruction of jets and leptons, as described in Section IV. Again we emphasize that the super-razor selection has no $E_T^{\rm miss}$ cut.

Moderate improvements over the M_Δ^R analysis are visible for the selectrons, while the chargino sensitivity is greatly increased in the low-mass degeneracy regime. The relative improvements can be more clearly seen in the Figures 32 and 33, where we show the exclusion reach for fixed values of selectron/chargino or neutralino masses.

The super-razor improvements in the sensitivity to compressed spectra can be understood from the additional kinematic information provided by the angles $\Delta\phi_R^\beta$ and $|\cos\theta_{R+1}|$. Recall that the magnitude of the approximate razor boost $\vec{\beta}_R$ is systematically larger than the correct boost $\vec{\beta}^{\rm CM}$, because of the the assumption that the energy of the event is evenly split between the visible and invisible systems. This causes a peaking of $\Delta\phi_R^\beta$ at π, since the sum of the visible momenta tends to be anti-aligned with the boost direction. As the spectrum becomes more and more compressed, this effect is magnified, as seen in Figure 20 of Section II. Thus for compressed spectra $\Delta\phi_R^\beta$ is a particularly good disciminator to appeal to in future searches.

As described in Section II, $|\cos\theta_{R+1}|$ is related to the energy difference of the leptons in the razor frame R, the approximation to the CM frame. This difference is expected to be small for the Drell-Yans + jets background, and is also peaked at zero for the W^-W^+ background, because of polarization effects. For signal events the distributions

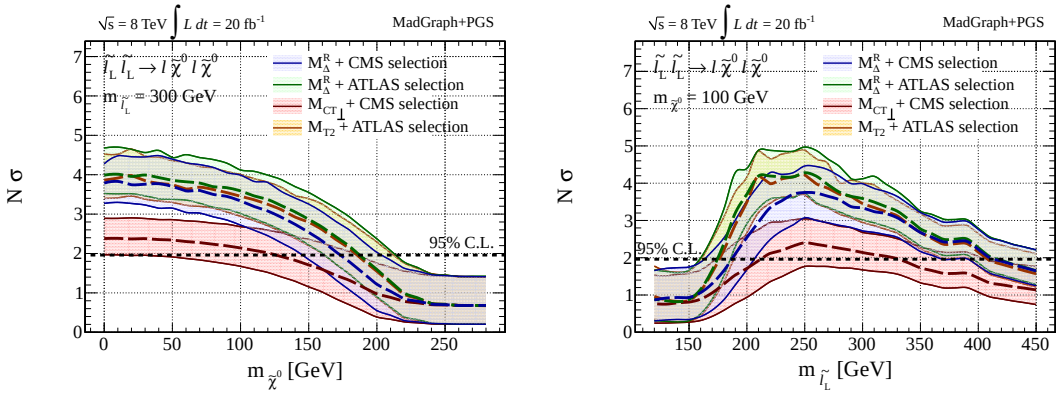

FIG. 30: Expected exclusion limits (in units of σ) for left-handed selectrons decaying to leptons and neutralinos using 20 fb^{-1} of 8 TeV data, as a function of neutralino mass with 300 GeV selectrons (left) or as a function of selectron mass with 100 GeV neutralinos (right). Expected limits are shown for our 1D M_Δ^R analysis using CMS (blue) and ATLAS (green) selection cuts, and directly compared to our expected exclusions using our simulated CMS $M_{CT\perp}$ (red) and ATLAS M_{T2} (orange) analyses.

in $|\cos\theta_{R+1}|$ are much flatter; the polarization effects are absent either because the parent particles are spin zero (sleptons) or because we have two-step decays (charginos).

Each of the super-razor variables, M_Δ^R, $\sqrt{\hat{s}}_R$, $\vec{\beta}_R$, $\vec{\beta}_{R+1}$, $\Delta\phi_R^\beta$, $|\cos\theta_{R+1}|$, and the angle $|\Delta\phi(\vec{p}_{\ell\ell}^{\,\rm lab}, \vec{E}_T^{\rm miss})|$ used in the super-razor selection, represents a different piece of information about an event. The collection can be thought of as a kinematic basis, which raises the question of whether one can identify an optimal kinematic basis for a particular type of search, e.g. searches for sleptons with compressed spectra. The answer to this question depends not just on the kinematic properties of the signal, but also on kinematics of the major backgrounds and especially on the detector effects that dominate the systematic uncertainties. It seems plausible that in some cases there may be a family of approximately equivalent kinematic bases, such that more or less the same kinematic information is exploited in different ways but resulting in approximately equivalent sensitivity.

Acknowledgments

We acknowledge helpful discussions with Paul Jackson, Maurizio Pierini, Chiu-Tien Yu, Javier Duarte and Avi Yagil. JL acknowledges the hospitality and support of the Theoretical Physics Group at SLAC. Fermilab is operated by Fermi Research Alliance, LLC, under contract DE-AC02-07CH11359 with the United States Department of Energy. MS and CR are funded by the United States Department of Energy under Grant DE-FG02-92-ER40701 and acknowledge the support of the Weston Havens Foundation.

[1] G. Aad et al. (ATLAS Collaboration), Eur.Phys.J. **C73**, 2362 (2013), 1212.6149.
[2] S. Chatrchyan et al. (CMS Collaboration), Phys.Rev.Lett. **111**, 081802 (2013), 1212.6961.
[3] S. Chatrchyan et al. (CMS Collaboration) (2013), 1301.2175.
[4] G. Aad et al. (ATLAS Collaboration) (2013), 1308.1841.
[5] S. Chatrchyan et al. (CMS Collaboration), Tech. Rep. CMS-PAS-SUS-13-004, CERN, Geneva (2013).
[6] M. Carena, A. Freitas, and C. Wagner, JHEP **0810**, 109 (2008), 0808.2298.
[7] X.-J. Bi, Q.-S. Yan, and P.-F. Yin, Phys.Rev. **D85**, 035005 (2012), 1111.2250.
[8] Y. Bai, H.-C. Cheng, J. Gallicchio, and J. Gu, JHEP **1207**, 110 (2012), 1203.4813.
[9] D. S. Alves, M. R. Buckley, P. J. Fox, J. D. Lykken, and C.-T. Yu, Phys.Rev. **D87**, 035016 (2013), 1205.5805.
[10] Z. Han, A. Katz, D. Krohn, and M. Reece, JHEP **1208**, 083 (2012), 1205.5808.
[11] B. Bhattacherjee and K. Ghosh (2012), 1207.6289.
[12] M. Carena, S. Gori, N. R. Shah, C. E. Wagner, and L.-T. Wang, JHEP **1308**, 087 (2013), 1303.4414.
[13] A. Delgado, G. F. Giudice, G. Isidori, M. Pierini, and A. Strumia, Eur.Phys.J. **C73**, 2370 (2013), 1212.6847.
[14] B. Dutta, T. Kamon, N. Kolev, K. Sinha, and K. Wang, Phys.Rev. **D86**, 075004 (2012), 1207.1873.
[15] J. A. Evans and Y. Kats, JHEP **1304**, 028 (2013), 1209.0764.
[16] C. Kilic and B. Tweedie, JHEP **1304**, 110 (2013), 1211.6106.
[17] M. R. Buckley, T. Plehn, and M. Takeuchi, JHEP **1308**, 086 (2013), 1302.6238.

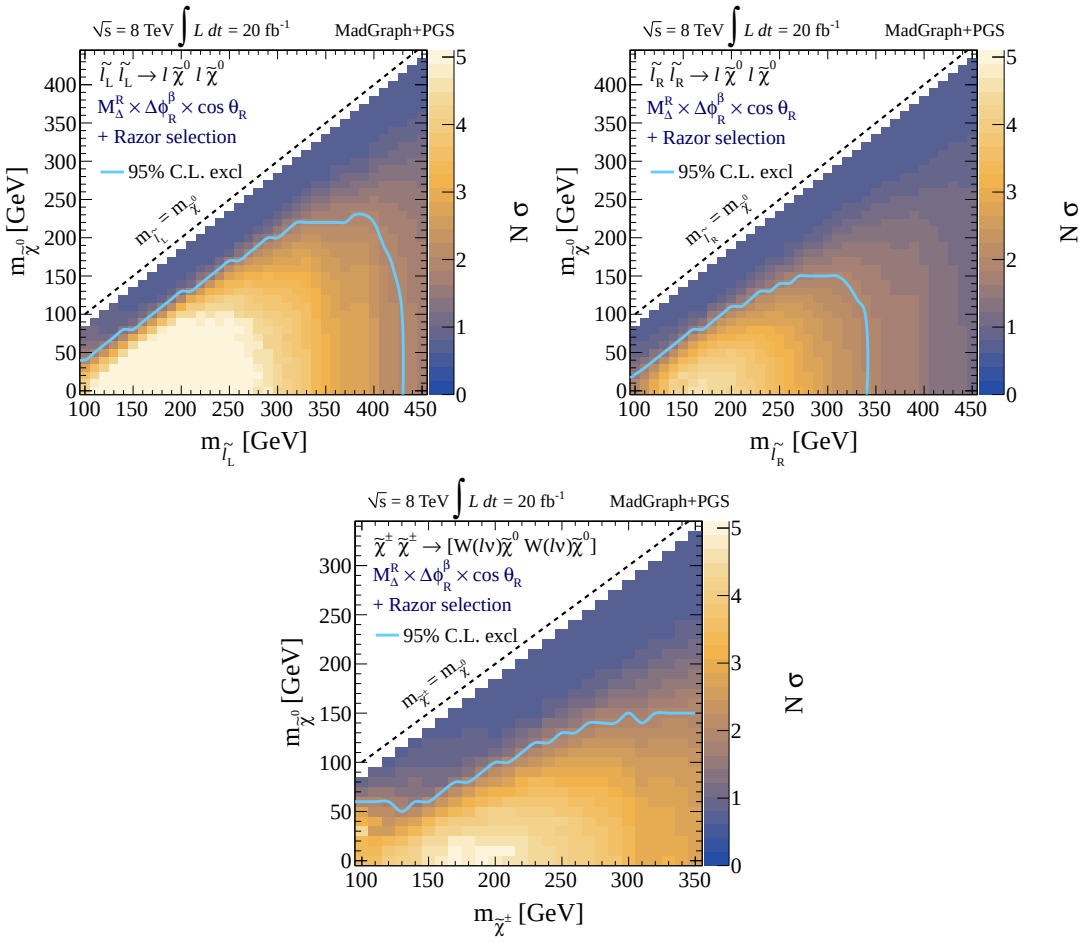

FIG. 31: Expected exclusion limits (in units of σ) for left-handed selectrons (upper left) right-handed selectrons (upper right), and charginos decaying to neutralinos and leptonic W bosons (bottom center) using 20 fb^{-1} of 8 TeV data, as a function of both selectron/chargino and neutralino masses. Expected limits are derived using our multi-dimensional M_Δ^R, $\Delta\phi_R^\beta$ and $|\cos\theta_{R+1}|$ analysis super-razor analyses with the razor selection cuts described in the previous section.

[18] Y. Bai, H.-C. Cheng, J. Gallicchio, and J. Gu, JHEP **1308**, 085 (2013), 1304.3148.
[19] R. Barbier, C. Berat, M. Besancon, M. Chemtob, A. Deandrea, et al., Phys.Rept. **420**, 1 (2005), hep-ph/0406039.
[20] C. Csaki, Y. Grossman, and B. Heidenreich, Phys.Rev. **D85**, 095009 (2012), 1111.1239.
[21] J. Berger, M. Perelstein, M. Saelim, and P. Tanedo, JHEP **1304**, 077 (2013), 1302.2146.
[22] S. Chatrchyan et al. (CMS Collaboration), Tech. Rep. CMS-PAS-SUS-12-027, CERN (2012).
[23] S. Chatrchyan et al. (CMS Collaboration), Tech. Rep. CMS-PAS-SUS-13-003, CERN (2013).
[24] G. Aad et al. (ATLAS Collaboration), JINST **1307**, P07015 (2013), 1305.2284.
[25] S. Chatrchyan et al. (CMS Collaboration), Phys.Lett. **B722**, 273 (2013), 1212.1838.
[26] F. Jegerlehner and A. Nyffeler, Phys.Rept. **477**, 1 (2009), 0902.3360.
[27] J. P. Miller, E. de Rafael, and B. L. Roberts, Rept.Prog.Phys. **70**, 795 (2007), hep-ph/0703049.
[28] M. R. Buckley, D. Hooper, and J. Kumar, Phys.Rev. **D88**, 063532 (2013), 1307.3561.
[29] M. Carena, S. Gori, N. R. Shah, C. E. Wagner, and L.-T. Wang, JHEP **1207**, 175 (2012), 1205.5842.
[30] J. Beringer et al. (Particle Data Group), Phys.Rev. **D86**, 010001 (2012).
[31] G. Aad et al. (ATLAS Collaboration) (2012), 1208.2884.
[32] ATLAS (ATLAS Collaboration), Tech. Rep. ATLAS-CONF-2013-049, CERN, Geneva (2013).
[33] S. Chatrchyan et al. (CMS Collaboration), Tech. Rep. CMS-PAS-SUS-12-022, CERN (2012).
[34] S. Chatrchyan et al. (CMS Collaboration), Tech. Rep. CMS-PAS-SUS-13-006, CERN, Geneva (2013).
[35] G. Aad et al. (ATLAS Collaboration) (2012), 1208.3144.
[36] G. Aad et al. (ATLAS Collaboration), Phys.Lett. **B718**, 841 (2013), 1208.3144.
[37] G. Aad et al. (ATLAS Collaboration), Tech. Rep. ATLAS-CONF-2013-036, CERN, Geneva (2013).
[38] ATLAS (ATLAS Collaboration), Tech. Rep. ATLAS-CONF-2013-035, CERN, Geneva (2013).
[39] C. Rogan (2010), 1006.2727.
[40] S. Chatrchyan et al. (CMS Collaboration), Phys.Rev. **D85**, 012004 (2012), 1107.1279.

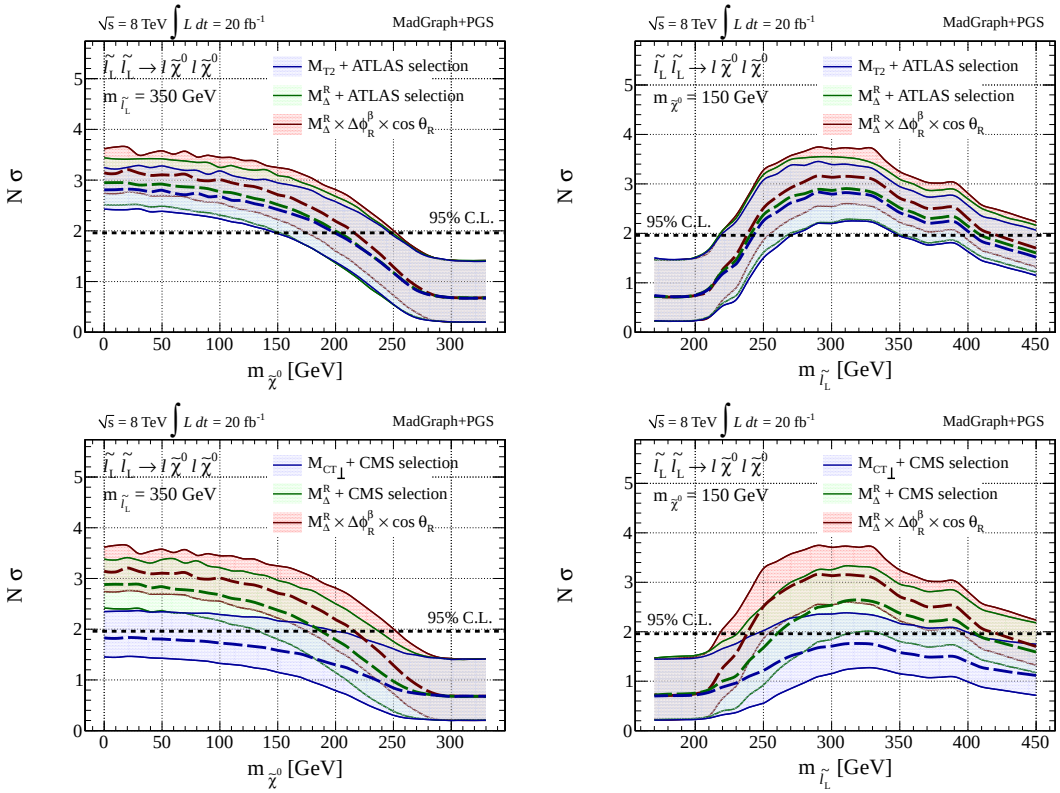

FIG. 32: Expected exclusion limits (in units of σ) for left-handed selectrons decaying to leptons and neutralinos using 20 fb^{-1} of 8 TeV data, as a function of neutralino mass with 350 GeV selectrons (upper and lower left) or as a function of selectron mass with 150 GeV neutralinos (upper and lower right). Expected limits are shown for our multi-dimensional razor analysis (red), and compared to either ATLAS (upper plots) or CMS (lower plots) mass variables and selection criteria.

[41] P. J. Fox, R. Harnik, R. Primulando, and C.-T. Yu, Phys.Rev. **D86**, 015010 (2012), 1203.1662.
[42] S. Chatrchyan et al. (CMS Collaboration), Tech. Rep. CMS-PAS-SUS-12-009, CERN, Geneva (2012).
[43] S. Chatrchyan et al. (CMS Collaboration) (2012), 1212.6961.
[44] C. Lester and D. Summers, Phys.Lett. **B463**, 99 (1999), hep-ph/9906349.
[45] A. Barr, C. Lester, and P. Stephens, J.Phys. **G29**, 2343 (2003), hep-ph/0304226.
[46] K. T. Matchev and M. Park, Phys.Rev.Lett. **107**, 061801 (2011), 0910.1584.
[47] D. R. Tovey, JHEP **0804**, 034 (2008), 0802.2879.
[48] S. Chatrchyan et al. (CMS Collaboration), JHEP **1210**, 018 (2012), 1207.1798.
[49] T. Aaltonen et al. (CDF Collaboration), Phys.Rev. **D81**, 031102 (2010), 0911.2956.
[50] C. Rogan, Ph.D. thesis, California Institute of Technology (2013).
[51] J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer, and T. Stelzer, JHEP **1106**, 128 (2011), 1106.0522.
[52] T. Sjostrand, S. Mrenna, and P. Z. Skands, JHEP **0605**, 026 (2006), hep-ph/0603175.
[53] D. L. Rainwater and D. Zeppenfeld, Phys.Rev. **D60**, 113004 (1999), hep-ph/9906218.
[54] W. Beenakker, M. Klasen, M. Kramer, T. Plehn, M. Spira, et al., Phys.Rev.Lett. **83**, 3780 (1999), hep-ph/9906298.
[55] M. Cacciari, G. P. Salam, and G. Soyez, Eur.Phys.J. **C72**, 1896 (2012), 1111.6097.
[56] M. Cacciari, G. P. Salam, and G. Soyez, JHEP **0804**, 063 (2008), 0802.1189.
[57] T. Junk, Nucl.Instrum.Meth. **A434**, 435 (1999), hep-ex/9902006.
[58] A. L. Read, J.Phys. **G28**, 2693 (2002).

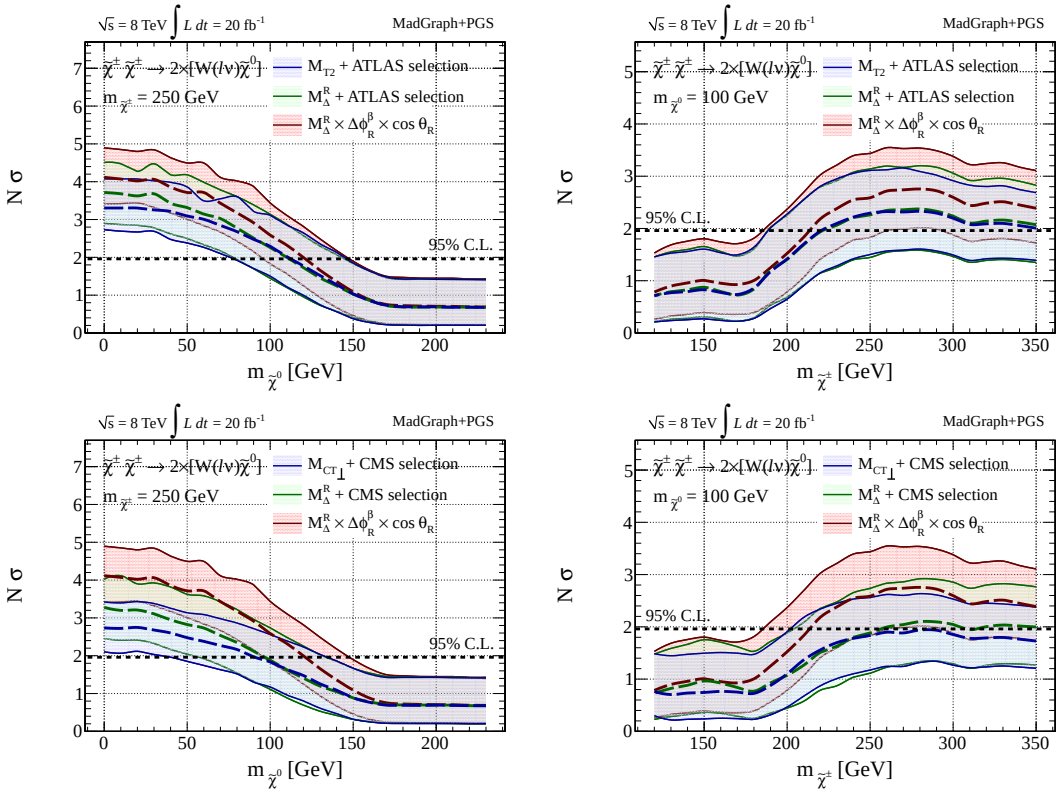

FIG. 33: Expected exclusion limits (in units of σ) for charginos decaying to neutralinos and leptonic W bosons using 20 fb^{-1} of 8 TeV data, as a function of neutralino mass with 250 GeV charginos (upper and lower left) or as a function of selectron mass with 100 GeV neutralinos (upper and lower right). Expected limits are shown for our multi-dimensional razor analysis (red), and compared to either ATLAS (upper plots) or CMS (lower plots) mass variables and selection criteria.