Challenges of Profile Likelihood Evaluation in Multi–Dimensional SUSY Scans

Farhan Feroz

Cavendish Laboratory, JJ Thomson Avenue, Cambridge CB3 0HE, UK
E-mail: f.feroz@mrao.cam.ac.uk

Kyle Cranmer

Center for Cosmology and Particle Physics, New York University, Washington Place, New York, NY 10003, USA
E-mail: cranmer@cern.ch

Mike Hobson

Cavendish Laboratory, JJ Thomson Avenue, Cambridge CB3 0HE, UK
E-mail: mph@mrao.cam.ac.uk

Roberto Ruiz de Austri

Instituto de Física Corpuscular, IFIC-UV/CSIC, Valencia, Spain
E-mail: rruiz@ific.uv.es

Roberto Trotta

Imperial College London, Blackett Laboratory, Prince Consort Road, London, SW7 2AZ, UK
E-mail: r.trotta@imperial.ac.uk

ABSTRACT: Statistical inference of the fundamental parameters of supersymmetric theories is a challenging and active endeavor. Several sophisticated algorithms have been employed to this end. While Markov-Chain Monte Carlo (MCMC) and nested sampling techniques are geared towards Bayesian inference, they have also been used to estimate frequentist confidence intervals based on the profile likelihood ratio. We investigate the performance and appropriate configuration of MULTINEST, a nested sampling based algorithm, when used for profile likelihood-based analyses both on toy models and on the parameter space of the Constrained MSSM. We find that while the standard configuration previously used in the literarture is appropriate for an accurate reconstruction of the Bayesian posterior, the profile likelihood is poorly approximated. We identify a more appropriate MULTINEST configuration for profile likelihood analyses, which gives an excellent exploration of the profile likelihood (albeit at a larger computational cost), including the identification of the global maximum likelihood value. We conclude that with the appropriate configuration MULTINEST is a suitable tool for profile likelihood studies, indicating previous claims to the contrary are not well founded.

KEYWORDS: Statistical Methods, Supersymmetry, phenomenology, Large Hadron Collider.

arXiv:1101.3296v2 [hep-ph] 25 May 2011

Contents

1. Introduction

The results of the first searches for supersymmetry (SUSY) at the Large Hadron Collider (LHC) have recently been released [1]. The field is eager to switch from a mode of excluding model parameters to an exciting and challenging era in which we are estimating the underlying parameters of a new fundamental theory. Within the context of SUSY, the technology has evolved from likelihood scans over low-dimensional subspaces of SUSY models [2–5] to higher-dimensional Bayesian analysis using Markov Chain Monte Carlo (MCMC) [6–8]. More recently, the nested sampling [9, 10] algorithm has been used by several authors for the study of SUSY models [11–20]. Although Bayesian techniques like MCMC and MULTINEST have been designed specifically to explore the Bayesian posterior distributions, they have also been used to obtain (approximate) profile likelihoods [14, 21, 22].

MULTINEST [23, 24], a publicly available implementation of the nested sampling algorithm, has been shown to reduce the computational cost of performing Bayesian analysis typically by two orders of magnitude as compared with basic MCMC techniques. MULTINEST has been integrated in the SuperBayeS code[1] for fast and efficient exploration of SUSY models. Recently, it has been demonstrated (at least for a limited region of the CMSSM parameter space around a specific benchmark point) that neural networks techniques can be used to reduce the computational efforts for these parameter scans by an additional factor of $\sim 10^4$ [25].

For highly non-Gaussian problems like supersymmetric parameter determination, inference can depend strongly on whether one chooses to work with the posterior distribution (Bayesian) or

---

[1]Available from: www.superbayes.org

profile likelihood (frequentist) [14, 21, 22]. There is a growing consensus that both the posterior and the profile likelihood ought to be explored in order to obtain a fuller picture of the statistical constraints from present-day and future data. This begs the question, which we address in this paper, of the algorithmic solutions available to reliably explore both the posterior and the profile likelihood in the context of SUSY phenomenology.

Recently, a genetic algorithm (GA) based method was developed in [26] specifically to obtain the profile likelihoods for CMSSM parameters and the resultant distributions were then compared with the profile likelihoods obtained with MULTINEST, run in the standard configuration commonly used in the literature. The authors found that MULTINEST missed several high likelihood regions in the CMSSM parameter space making the resultant profile likelihood functions highly inaccurate. They further questioned the accuracy of posterior distributions obtained with MULTINEST. They went on to argue that MULTINEST might not be able to find these high likelihood regions even if its termination criterion is adjusted for finding the profile likelihoods. One of the aims of this paper is to show that these shortcomings can be avoided through proper configuration of the MULTINEST.

The outline of this paper is as follows. In Sec. 2 we first give an introduction to the Bayesian and frequentist frameworks for statistical analysis. We briefly describe the MULTINEST algorithm and how it can be tuned for evaluating profile likelihoods in Sec. 3. In Sec. 4 we use two relevant toy problems to study the ability of MULTINEST to reconstruct both the posterior and profile likelihood. We then obtain the profile likelihood function of CMSSM parameters using MULTINEST in Sec. 5 and compare our results with the GA method adopted in [26]. We also discuss implications for direct and indirect detection prospects. Finally we present our conclusions in Sec. 6.

2. Statistical Framework

2.1 Bayesian Modelling

We briefly recall some basics about Bayesian inference, referring the reader to e.g. [27] for further details. Bayesian inference is based on Bayes' theorem, stating that

$$P(\boldsymbol{\Theta}|\mathbf{D}) = \frac{P(\mathbf{D}|\boldsymbol{\Theta})P(\boldsymbol{\Theta})}{P(\mathbf{D})}, \tag{2.1}$$

where $P(\boldsymbol{\Theta}|\mathbf{D})$ is the posterior probability distribution of the parameters $\boldsymbol{\Theta}$ given data $\mathbf{D}$, $P(\mathbf{D}|\boldsymbol{\Theta}) \equiv \mathcal{L}(\boldsymbol{\Theta})$ is the likelihood function, $P(\boldsymbol{\Theta})$ is the prior, and $P(\mathbf{D}) \equiv \mathcal{Z}$ is the Bayesian evidence (or model likelihood). The latter is a normalization constant, obtained by averaging the likelihood over the prior:

$$\mathcal{Z} = \int \mathcal{L}(\boldsymbol{\Theta})P(\boldsymbol{\Theta})d\boldsymbol{\Theta}. \tag{2.2}$$

In Bayesian statistics, the N–dimensional posterior distribution described in Eq. (2.1) constitutes the complete Bayesian inference of the parameter values. The inference about an individual parameter θ_i is then given by the marginalised posterior distribution $P(\theta_i|\mathbf{D})$ which can be obtained by integrating (marginalizing) the N–dimensional posterior distribution over all the parameters apart from θ_i. If samples from the full posterior are available (having been generated for example via Markov Chain Monte Carlo sampling), then the above integral is replaced by a simple counting

procedure: since posterior samples are distributed according to the posterior, the marginal posterior for θ_i above is obtained from the full posterior by dividing the range of θ_i in a number of bins and counting how many posterior samples fall in each bin. A 1-D marginal posterior interval corresponding to a symmetric credible region containing $1 - \alpha$ of posterior probability for θ_i is delimited by an interval $[\theta_i^-, \theta_i^+]$ such that:

$$\int_{-\infty}^{\theta_i^-} P(\theta_i|\mathbf{D})d\theta_i = \alpha/2 \quad \text{and} \quad \int_{\theta_i^+}^{\infty} P(\theta_i|\mathbf{D})d\theta_i = \alpha/2. \tag{2.3}$$

2.2 Profile Likelihood-Based Frequentist Analysis

In the classical or frequentist school of statistics, one defines the probability of an event as the limit of its frequency in a large number of trials. Classical confidence intervals based on the Neyman construction are defined as the set of parameter points in which some real-valued function, or *test statistic*, t evaluated on the data falls in an acceptance region $W_\mathbf{\Theta} = [t_-, t_+]$. If the boundaries of the acceptance regions are such that $P(t \in W_\mathbf{\Theta} | \mathbf{\Theta}) \geq 1 - \alpha$ is satisfied, then the resulting intervals will cover the true value of $\mathbf{\Theta}$ with a probability of at least $1 - \alpha$. Likelihood ratios are often chosen as the test statistic on which frequentist intervals are based. When $\mathbf{\Theta}$ is composed of parameters of interest, θ, and nuisance parameters, ψ, a common choice of test statistic is the profile likelihood ratio

$$\lambda(\theta) \equiv \frac{\mathcal{L}(\theta, \hat{\hat{\psi}})}{\mathcal{L}(\hat{\theta}, \hat{\psi})}. \tag{2.4}$$

where $\hat{\hat{\psi}}$ is the conditional maximum likelihood estimate (MLE) of ψ with θ fixed and $\hat{\theta}, \hat{\psi}$ are the unconditional MLEs. Under certain regularity conditions, Wilks showed that the distribution of $-2\ln\lambda(\theta)$ converges to a chi-square distribution with a number of degrees of freedom given by the dimensionality of θ [28]. Thus, by assuming that the asymptotic distribution is a good approximation of the finite sample case, one can trivially define the acceptance region $W_\mathbf{\Theta}$ from standard lookup tables. One must keep in mind that these intervals based on the asymptotic properties of the profile likelihood ratio require certain regularity conditions to be fulfilled and are not guaranteed to have strict coverage. In particular, in cases with complex multimodal likelihoods one might reasonably suspect that the distribution of $-2\ln\lambda(\theta)$ is still far from converging to its asymptotic form (see [25] for an illustration). While there exist modifications to the profile likelihood ratio that converge more rapidly [29–31], within particle physics the profile likelihood ratio is a standard technique[2].

One can approximate the $\lambda(\theta)$ using an arbitrary sampling of the likelihood function as long as one has access to the value of the likelihood function at the sampled points. The procedure is simple: one groups the samples in bins of θ and for each bin searches for the maximum likelihood sample in that bin, which corresponds to $\mathcal{L}(\theta, \hat{\hat{\psi}})$.

It can be seen from Eq. (2.4) that the profile likelihood function requires the conditional MLEs $\hat{\hat{\psi}}(\theta)$ to be found for each value of θ as well as the unconditional MLE $(\hat{\theta}, \hat{\psi})$, which represents the global maximum likelihood point. If the unconditional MLE is not properly resolved, this will

---

[2]In the following, we will refer to the profile likelihood ratio as "profile likelihood" or "likelihood function" for brevity.

effectively result in a different threshold on $-2\ln\lambda(\theta)$ corresponding to a wrong confidence level. Even if the unconditional MLE is found, we might also worry how well the conditional MLEs $\hat{\hat{\psi}}(\theta)$ are found, particularly when $\mathcal{L}(\theta,\hat{\hat{\psi}}) \ll \mathcal{L}(\hat{\theta},\hat{\psi})$.

3. Using MULTINEST for Profile Likelihood Exploration

Nested sampling [9] is a Monte Carlo method whose primary aim is the efficient calculation of the Bayesian evidence given by Eq. (2.2) As a by-product, the algorithm also produces posterior samples which can be used to map out the posterior distribution of Eq. (2.1). Those same samples have also been used to estimate the profile likelihood, a point to which we return below in greater detail. Nested sampling calculates the evidence by transforming the multi–dimensional evidence integral into a one–dimensional integral that is easy to evaluate numerically. This is accomplished by defining the prior volume X as $dX = P(\mathbf{\Theta})d\mathbf{\Theta}$, so that

$$X(\Lambda) = \int_{\mathcal{L}(\mathbf{\Theta})>\Lambda} P(\mathbf{\Theta})d\mathbf{\Theta}, \qquad (3.1)$$

where the integral extends over the region(s) of parameter space contained within the iso-likelihood contour $\mathcal{L}(\mathbf{\Theta}) = \Lambda$. The evidence integral, Eq. (2.2), can then be written as:

$$\mathcal{Z} = \int_0^1 \mathcal{L}(X)dX, \qquad (3.2)$$

where $\mathcal{L}(X)$, the inverse of Eq. (3.1), is a monotonically decreasing function of X. Thus, if one can evaluate the likelihoods $\mathcal{L}_i = \mathcal{L}(X_i)$, where X_i is a sequence of decreasing values,

$$0 < X_M < \cdots < X_2 < X_1 < X_0 = 1, \qquad (3.3)$$

the evidence can be approximated numerically using standard quadrature methods as a weighted sum

$$\mathcal{Z} = \sum_{i=1}^{M} \mathcal{L}_i w_i, \qquad (3.4)$$

where the weights w_i for the simple trapezium rule are given by $w_i = \frac{1}{2}(X_{i-1} - X_{i+1})$.

The summation in Eq. (3.4) is performed as follows. The iteration counter is first set to $i = 0$ and n_{live} 'active' (or 'live') samples are drawn from the full prior $P(\mathbf{\Theta})$, so the initial prior volume is $X_0 = 1$. The samples are then sorted in order of their likelihood and the smallest (with likelihood $\mathcal{L}_0$) is removed from the active set (hence becoming 'inactive') and replaced by a point drawn from the prior subject to the constraint that the point has a likelihood $\mathcal{L} > \mathcal{L}_0$. The corresponding prior volume contained within the iso-likelihood contour associated with the new live point will be a random variable given by $X_1 = t_1 X_0$, where t_1 follows the distribution $P(t) = Nt^{N-1}$ (i.e., the probability distribution for the largest of N samples drawn uniformly from the interval $[0,1]$). At each subsequent iteration i, the removal of the lowest likelihood point $\mathcal{L}_i$ in the active set, the drawing of a replacement with $\mathcal{L} > \mathcal{L}_i$ and the reduction of the corresponding prior volume $X_i = t_i X_{i-1}$ are repeated, until the entire prior volume has been traversed. The algorithm thus travels through nested shells of likelihood as the prior volume is reduced.

The nested sampling algorithm is terminated when the evidence has been computed to a pre-specified precision. The evidence that could be contributed by the remaining live points is estimated as $\Delta\mathcal{Z}_\mathrm{i} = \mathcal{L}_\mathrm{max}X_\mathrm{i}$, where $\mathcal{L}_\mathrm{max}$ is the maximum-likelihood value of the remaining live points, and X_i is the remaining prior volume. The algorithm terminates when $\Delta\mathcal{Z}_\mathrm{i}$ is less than a user-defined value.

The most challenging task in implementing nested sampling is to draw samples from the prior within the hard constraint $\mathcal{L} > \mathcal{L}_i$ at each iteration i. The MULTINEST algorithm [23, 24] tackles this problem through an ellipsoidal rejection sampling scheme. The live point set is enclosed within a set of (possibly overlapping) ellipsoids and a new point is then drawn uniformly from the region enclosed by these ellipsoids. The ellipsoidal decomposition of the live point set is chosen to minimize the sum of volumes of the ellipsoids. The ellipsoidal decomposition is well suited to dealing with posteriors that have curving degeneracies, and allows mode identification in multi-modal posteriors. If there are subsets of the ellipsoid set that do not overlap with the remaining ellipsoids, these are identified as a distinct mode and subsequently evolved independently.

The two most important parameters that control the parameter space exploration in nested sampling are the number of live points n_live – which determines the resolution at which the parameter space is explored – and a tolerance parameter tol, which defines the termination criterion based on the accuracy of the evidence. As discussed in the previous section, evaluating profile likelihoods is much more challenging than evaluating posterior distributions. Therefore, one should not expect that a vanilla setup for MULTINEST (which is adequate for an accurate exploration of the posterior distribution) will automatically be optimal for profile likelihoods evaluation. Generally, a larger number of live points is necessary to explore profile likelihoods accurately. Moreover, setting tol to a smaller value results in MULTINEST gathering a larger number of samples in the high likelihood regions (as termination is delayed). This is usually not necessary for the posterior distributions, as the prior volume occupied by high likelihood regions is usually very small and therefore these regions have relatively small probability mass. For profile likelihoods, however, getting as close to the true global maximum is crucial and therefore one should set tol to a relatively smaller value. We now investigate these issues in the context of two relevant toy problems.

4. Application to Toy Models

In order to demonstrate the differences between profile likelihoods and the posterior distributions reconstruction with MULTINEST, we consider two 8-D toy problems in this section. The dimensionality of these toy problems is chosen to correspond to the dimensionality of CMSSM analyses in which 4 CMSSM parameters are varied along with 4 Standard Model (SM) parameters (see Sec. 5).

The first problem we consider is a uni-modal 8-D isotropic Gaussian with the following likelihood function:

$$\mathcal{L}(\Theta) = \frac{1}{(2\pi)^4\sigma^8} \exp\left[-\frac{1}{2}\sum_{\mathrm{j}=1}^{8} \frac{(\theta_\mathrm{j} - \mu_\mathrm{j})^2}{\sigma^2} \right] \tag{4.1}$$

with $\sigma = 0.05$ and $\mu_\mathrm{j} = 0.5$ $(j = 1,\ldots,8)$. We assume uniform priors $\mathcal{U}(0, 1)$ for all 8 parameters. The analytical posterior distributions and profile likelihood functions are shown in Fig. 1 for a
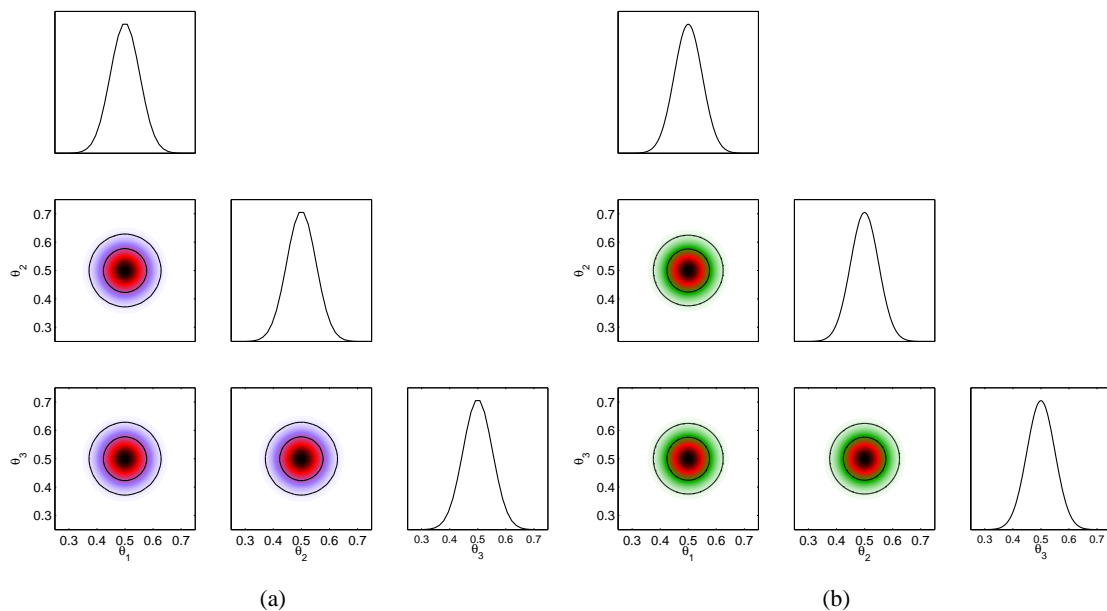
Figure 1: Analytical 1-D and 2-D (a) marginalized posterior distributions and (b) profile likelihood functions for the unimodal 8-D Gaussian problem. The contours represent the 68% and 95% Bayesian credible regions in (a) and profile likelihood confidence intervals in (b) respectively.

subset of the parameters. As expected, the posterior and profile likelihood distributions are identical in this case.

We now reconstruct both profile likelihood and posterior distribution with MULTINEST, using $n_{\text{live}} = 1000$ with $\text{tol} = 0.5$ (which is expected to be adequate for posterior reconstruction) and 1×10^{-4} (a much lower value targeted at profile likelihood reconstruction). The posterior distributions and profile likelihoods recovered by MULTINEST for the first three parameters are shown in Figs. 2 and 3 respectively. It is evident from these plots that recovered posterior distributions for both $\text{tol} = 0.5$ and 1×10^{-4} are almost identical to the analytical posterior more than 3σ into the tails. However, the recovered profile likelihood for $\text{tol} = 0.5$ is quite noisy and inaccurate, especially around the peak. This is because MULTINEST has not explored the high likelihood region in great detail, as the algorithm is terminated after about 163,000 likelihood evaluations, which is adequate to calculate the evidence to sufficient accuracy. We can circumvent this problem by setting $\text{tol} = 1 \times 10^{-4}$ (dotted blue curves), which results in a larger number of likelihood evaluations, around 261,000. The situation around the peak is now greatly improved, although the tails of the profile likelihood are still not explored to very high accuracy beyond the $\sim 3\sigma$ region. Nevertheless, the overall profile likelihood has been recovered with sufficient accuracy to allow for a reasonably correct parameter estimation.

The second problem we consider is a multi-modal 8-D Gaussian mixture model with the following likelihood function:

$$\mathcal{L}(\Theta) = \sum_{i=1}^{3} w_i \frac{1}{(2\pi)^4 \prod_{j=1}^{8} \sigma_{ij}} \exp\left[-\frac{1}{2} \sum_{j=1}^{8} \frac{(\theta_i - \mu_{ij})^2}{\sigma_{ij}^2}\right] \tag{4.2}$$
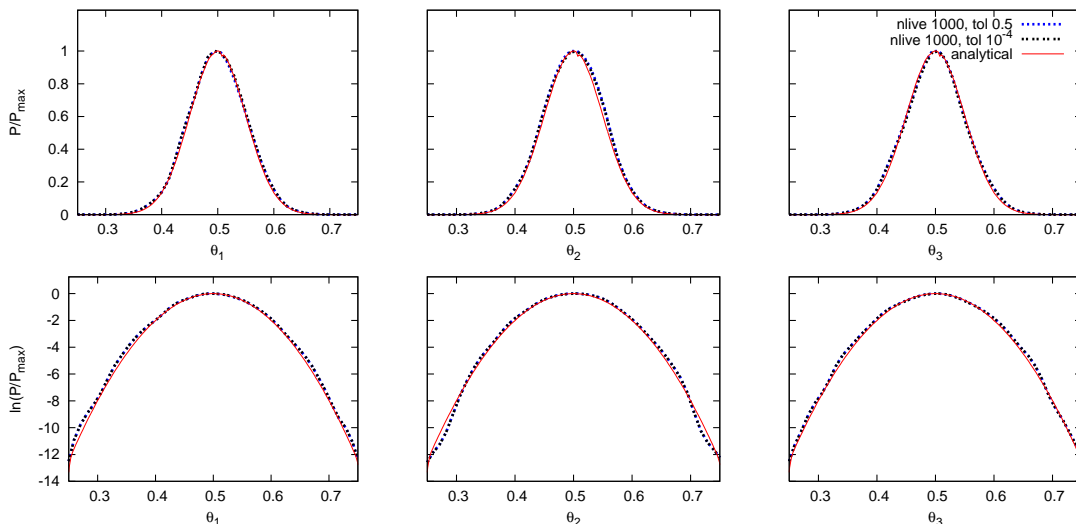
Figure 2: 1-D posterior distributions for the first 3 parameters of the uni-modal 8-D Gaussian problem. The red curve shows the analytical posterior distribution while the blue dashed and grey dotted curves show the posterior distributions recovered by MULTINEST with tol set to 0.5 and 1×10^{-4} respectively. 1000 live points were used in both cases. The upper and lower panels show the posterior and ln of posterior distributions respectively.

with the following values for the parameters defining the likelihood function:

$$\mathcal{M}_1 : w_1 = 0.98, \sigma_{1i} = 0.05 \text{ and } \mu_{1i} = 0.5\,(1 \leq i \leq 8); \tag{4.3}$$

$$\mathcal{M}_2 : w_2 = 0.01, \sigma_{2i} = 0.02\,(1 \leq i \leq 8), \mu_{21} = 0.55, \mu_{2i} = 0.5\,(2 \leq i \leq 8); \tag{4.4}$$

$$\mathcal{M}_3 : w_3 = 0.01, \sigma_{31} = 0.004, \sigma_{32} = 0.1, \sigma_{3i} = 0.02\,(3 \leq i \leq 8),$$
$$\mu_{31} = 0.3, \mu_{3i} = 0.5\,(2 \leq i \leq 8). \tag{4.5}$$

We assume uniform priors $\mathcal{U}(0,1)$ for all 8 parameters. This describes a multi-modal distribution with the broad Gaussian of the previous example centered in the middle of the parameter space. A narrow isotropic Gaussian lies slightly to the right of the broad Gaussian in θ_1 while another narrow Gaussian lies to the left. We denote these three modes by $\mathcal{M}_1$, $\mathcal{M}_2$ and $\mathcal{M}_3$ respectively. $\mathcal{M}_3$ has its variances in the first two dimensions chosen in such a way to mimic the funnel region at low $m_{1/2}$ values in the CMSSM model. The weights of the mixture model are set such that the $\mathcal{M}_1$ occupies 98% of the posterior probability mass while $\mathcal{M}_2$ and $\mathcal{M}_3$ collectively occupy only 2% of the probability mass and therefore we would expect the posterior probability distribution to be dominated by $\mathcal{M}_1$. The analytical 1-D and 2-D posterior distributions and profile likelihood functions are shown in Fig. 4 (only distributions of the first 3 parameters are shown, as parameters θ_4 onwards behave exactly as θ_3). As expected, the posterior distributions are dominated almost entirely by $\mathcal{M}_1$, the mode with the largest posterior probability mass. $\mathcal{M}_3$ barely registers in the posterior distribution as the secondary peak in the θ_1 direction, while $\mathcal{M}_2$ is completely masked by $\mathcal{M}_1$.

The profile likelihood, on the other hand, is dominated by the narrow, highly-peaked modes $\mathcal{M}_2$ and $\mathcal{M}_3$, whose peak likelihood values are 9.25 and 9.61 times higher, respectively, than the
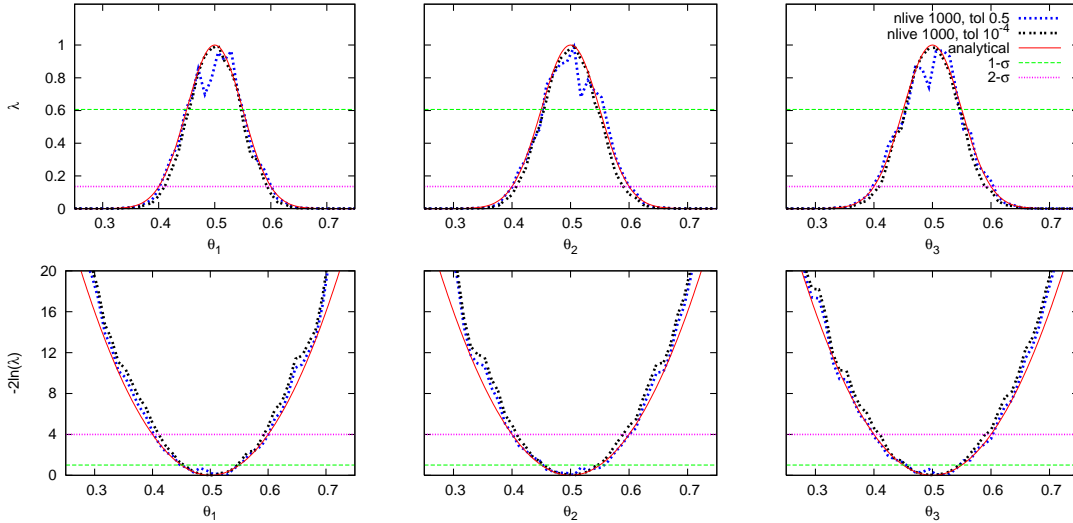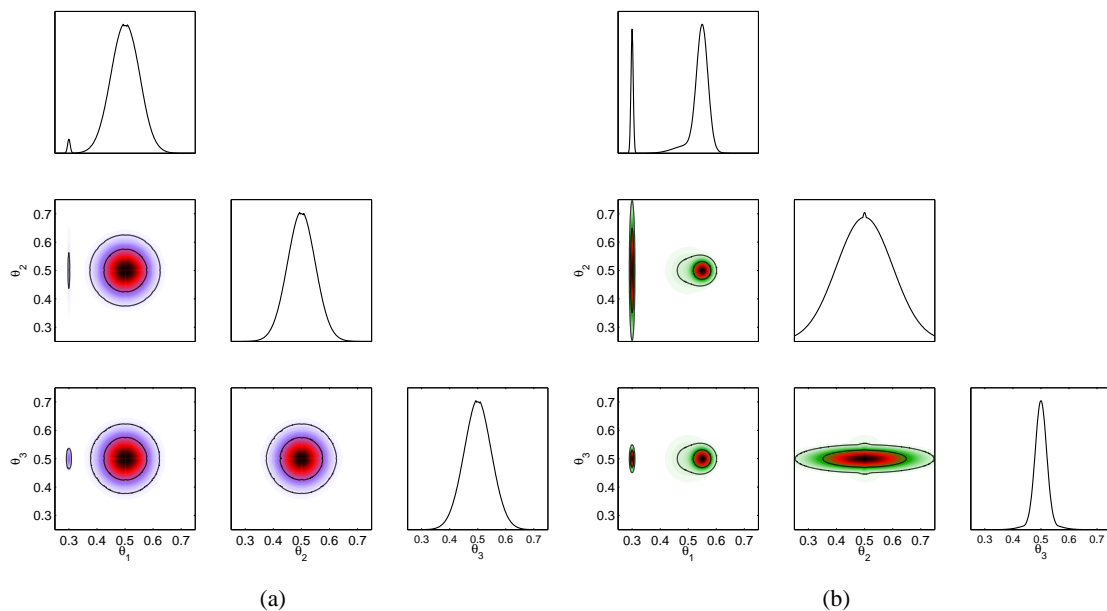
Figure 3: 1-D profile likelihood functions for the first 3 parameters of the uni-modal 8-D Gaussian problem. The red curve shows the analytical profile likelihood functions while the blue dashed and grey dotted curves show the profile likelihoods recovered by MULTINEST with tol set to 0.5 and 1×10^{-4} respectively. 1000 live points were used in both cases. The upper and lower panels show the profile likelihood and $-2\ln$ of profile likelihood functions respectively. Green and magenta horizontal lines represent the 1σ and 2σ profile likelihood intervals respectively.

likelihood value at the peak of $\mathcal{M}_1$. This is indeed reflected in the right panels of Fig. 4. The 1-D profile likelihood for θ_1 is bi-modal with $\mathcal{M}_2$ and $\mathcal{M}_3$ constituting the two modes. $\mathcal{M}_3$ only produces a heavier, non-Gaussian tail extending downwards from $\theta_1 = 0.55$ (where the maximum likelihood peak of $\mathcal{M}_2$ is located). Looking instead at the profile likelihood for θ_2, the (very subdominant) contribution from $\mathcal{M}_1$ barely shows up as a small extra peak on top of the broader peak from $\mathcal{M}_3$ at $\theta_2 = 0.5$. The profile likelihood for θ_2 consists almost entirely of $\mathcal{M}_3$ as its standard deviation in θ_2 is 5 times larger than the one of $\mathcal{M}_2$ while the likelihood value at the peak of $\mathcal{M}_3$ is only slightly smaller than the value for $\mathcal{M}_2$. We notice that confidence regions for the profile likelihood plotted in the 2-D panels are only approximately correct, as we have assumed that the likelihood distribution follows a multivariate Gaussian distribution in order to derive the confidence limits as discussed in Sec. 2.2, which is obviously not correct for this particular toy model.

These dramatic differences in the posterior and profile likelihood functions make this problem an extremely challenging one for any numerical inference algorithm to tackle. Even just finding $\mathcal{M}_2$ and $\mathcal{M}_3$ is an extremely difficult task as they are essentially spikes in the probability distribution. We show the recovered 1-D posterior distributions and profile likelihood with MULTINEST in Figs. 6 and 5, respectively. With $n_{\mathrm{live}} = 4,000$ and $\mathrm{tol} = 0.5$ (resulting in about 725,000 likelihood evaluations), MULTINEST finds all three Gaussians and the recovered posterior distributions are almost identical to the analytical posterior distributions. However, the profile likelihood functions are highly inaccurate and very noisy. This is because the high likelihood regions have not been explored in great detail as the sampling proceeds according to the posterior mass, which is

Figure 4: 1-D and 2-D (a) marginalized posterior distributions and (b) profile likelihood functions for the 8-D Gaussian mixture problem. The contours represent the 68% and 95% Bayesian credible regions in (a) and profile likelihood confidence intervals in (b) respectively. The plots were generated after dividing the analytical posterior values in 200 bins so that a fair comparison can be made with the distributions obtained using MULTINEST.

heavily concentrated in $\mathcal{M}_1$. Therefore, the modes $\mathcal{M}_2$ and $\mathcal{M}_3$ are explored with a proportionally lower number of live points, as their posterior mass is small.

This problem can be solved by running MULTINEST with $n_{\text{live}} = 20,000$ and decreasing tol to 1×10^{-4} which results in around 11 million likelihood evaluations. Increasing the number of live points ensures that $\mathcal{M}_3$ encloses more live points and is explored at higher resolution. As can be seen from Fig. 6, the recovered profile likelihood functions are much more accurate and less noisy with this setup. The 1-D profile likelihoods for θ_1 and θ_3 are almost identical to the analytical ones, but we can see some under-sampling of the profile likelihood for θ_2. Despite this, just finding $\mathcal{M}_3$ is already an extremely challenging task for any numerical technique, including GA, and although the recovered profile likelihood functions by MULTINEST are not perfect, they are reasonably accurate, certainly within the 2σ confidence region.

5. Application to the CMSSM

The Minimal Supersymmetric Standard Model (MSSM) [32, 33] with R-parity can solve the hierarchy problem and provide a candidate dark matter (DM) particle. The MSSM with one particular choice of universal boundary conditions at the grand unification scale, is called the Constrained Minimal Supersymmetric Standard Model (CMSSM) [34]. In the CMSSM, the scalar mass m_0, gaugino mass $m_{1/2}$ and tri–linear coupling A_0 are assumed to be universal at a gauge unification scale $M_{\text{GUT}} \sim 2 \times 10^{16}$ GeV. In addition, at the electroweak scale one selects $\tan\beta$, the ratio of Higgs vacuum expectation values and $\text{sign}(\mu)$, where μ is the Higgs/higgsino mass parameter
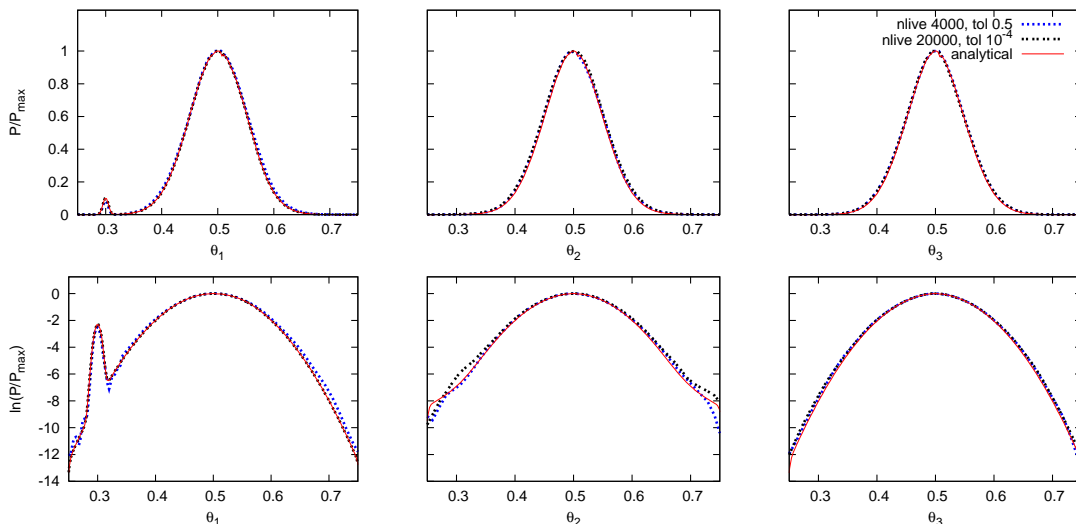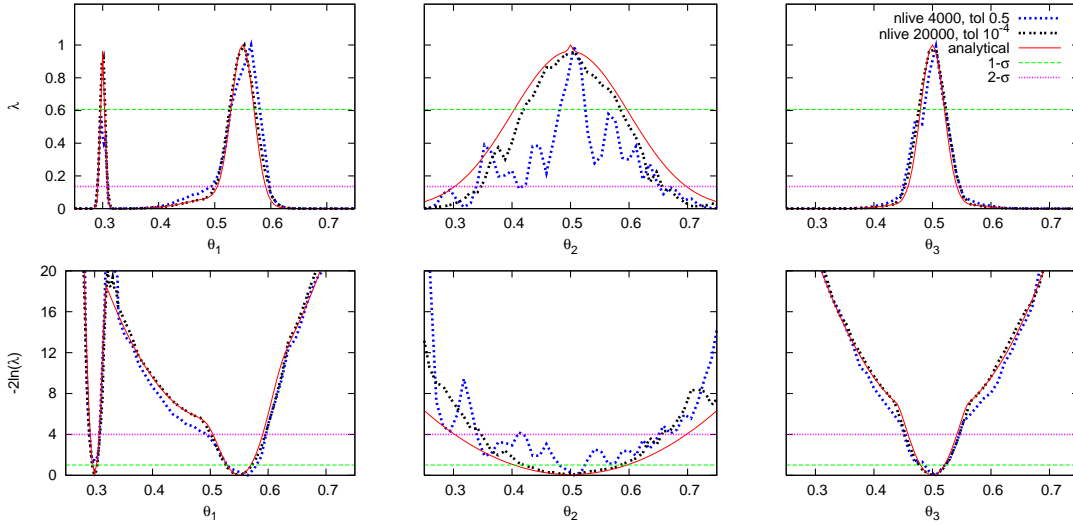
Figure 5: 1-D posterior distributions for the first 3 parameters of the 8-D Gaussian mixture problem. The red curve shows the analytical posterior distribution while the blue dashed and grey dotted curves show the posterior distributions recovered by MULTINEST with $n_{\text{live}} = 4,000$, tol $= 0.5$ and $n_{\text{live}} = 20,000$, tol $= 1 \times 10^{-4}$ respectively. The upper and lower panels show the posterior and ln of posterior distributions respectively.

whose square is computed from the potential minimisation conditions of electroweak symmetry breaking (EWSB) and the empirical value of the mass of the Z^0 boson, M_Z. The family universality assumption is well motivated since flavour changing neutral currents are observed to be rare. Indeed several string models (see, for example Ref. [35, 36]) predict approximate MSSM universality in the soft terms.

The CMSSM has proved to be a popular choice for SUSY phenomenology because of the small number of free parameters and it has recently been studied quite extensively in multi–parameter scans, both from a frequentist [4, 5, 37] and a Bayesian perspective [6–8, 14, 15, 22, 26, 38]. Bayesian and frequentist analyses are expected to give consistent inferences for the CMSSM parameters if the data are sufficiently constraining. However, several groups have concluded that we do not yet have enough constraining power in the available data to overridethe influence of priors in the Bayesian framework. Furthermore, the structure of the parameter space is such that the likelihood presents "spikes" that contain little posterior mass (for reasonable choice of priors). As a consequence, regions of high likelihood (favoured in a frequentist analysis) do not necessarily match regions of large posterior mass (favoured under a Bayesian analysis), see e.g. [14]. Therefore, Bayesian and frequentist inferences should not be expected to yield quantitatively consistent results. Both hurdles are expected to be overcome once ATLAS results become available, see [39], although the actual performance of different statistical approaches, as measured by their coverage properties, has only just begun to be scrutinized [25, 40].

Here we focus on the algorithmic problem of obtaining reliable posterior probability and profile likelihood maps with MULTINEST. The authors of Ref. [26] claimed that the GA method produces more accurate profile likelihood functions than MULTINEST. They also raised questions
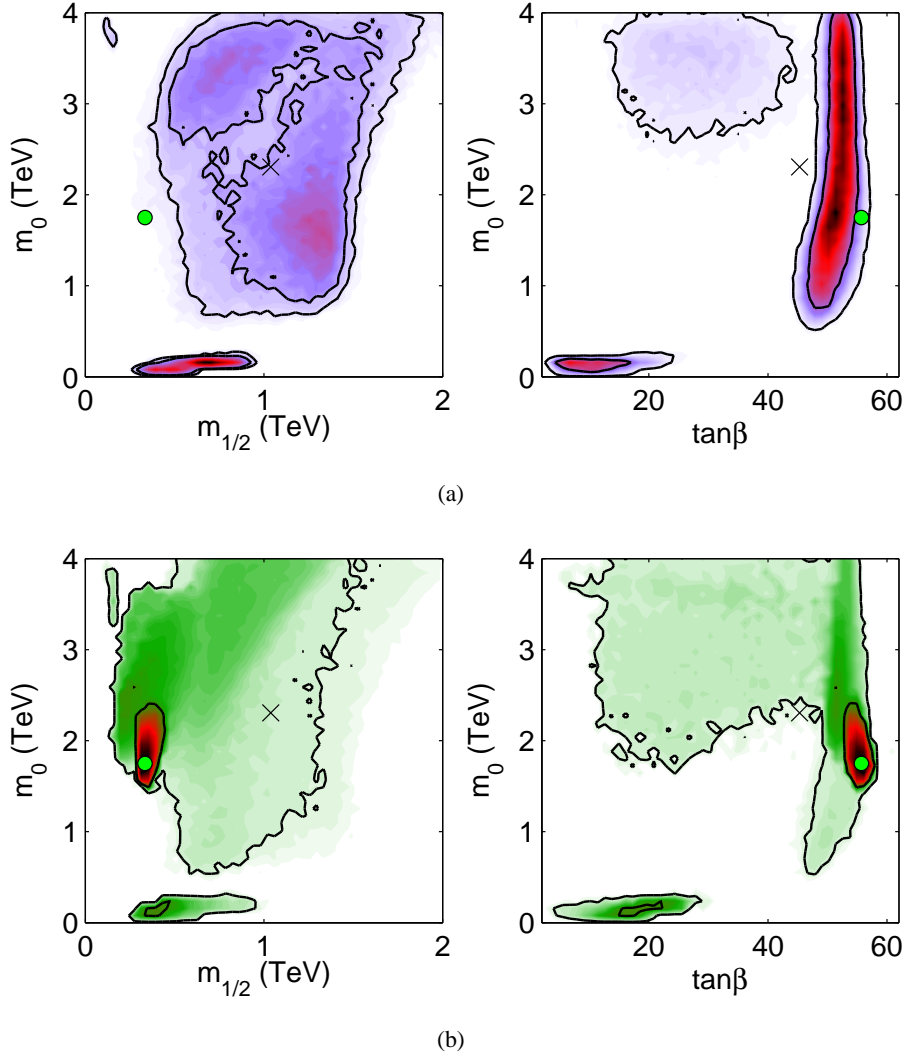
Figure 6: 1-D profile likelihood functions for the first 3 parameters of the 8-D Gaussian mixture problem. The red curve shows the analytical profile likelihood function while the blue dashed and grey dotted curves show the profile likelihoods recovered by MULTINEST with $n_{\text{live}} = 4,000$, tol $= 0.5$ and $n_{\text{live}} = 20,000$, tol $= 1 \times 10^{-4}$ respectively. The upper and lower panels show the profile likelihood and $-2 \ln$ of profile likelihood functions respectively. Green and magenta horizontal lines represent the 1σ and 2σ profile likelihood intervals respectively.

about the accuracy of the posterior distributions obtained with MULTINEST. The previous section demonstrates that the posterior distributions from MULTINEST are highly reliable, as they match very well the analytical solution for both toy problems out to more than 3σ in the tails. We stress that this is the case even for the less computationally intensive MULTINEST configuration, with $n_{\text{live}} = 1000$ and tol $= 0.5$.

We now reconstruct both the posterior distribution and the profile likelihood for the CMSSM parameters using MULTINEST, as implemented in the `SuperBayeS v1.5` code in almost exactly the same setup employed in Ref. [26]. The set of CMSSM and SM parameters together constitute an 8-dimensional parameter space with $\Theta = (m_0, m_{1/2}, A_0, \tan\beta, m_t, m_b(m_b)^{\overline{MS}}, \alpha^{\overline{MS}}, \alpha_s^{\overline{MS}}(M_Z))$ to be scanned and constrained with the presently available experimental data. The ranges over which we explore these parameters are $m_0, m_{1/2} \in (0.05, 4)$ TeV, $A_0 \in (-4, 4)$ TeV, $\tan\beta \in (2, 65)$, $m_t \in (167.0, 178.2)$ GeV, $m_b(m_b)^{\overline{MS}} \in (3.92, 4.48)$ GeV, $1/\alpha^{\overline{MS}}(M_Z) \in (127.835, 128.075)$ and $\alpha_s^{\overline{MS}}(M_Z) \in (0.1096, 0.1256)$. We use a slightly wider range for $\tan\beta$ and a smaller range for A_0 which was allowed to vary between -7 TeV and 7 TeV in [26] as there is very little probability mass as well as very few points with high likelihood with $|A_0| > 4$ TeV. The observables used in our analysis are exactly the same as given in Table 1 of Ref. [26], in order to allow a comparison.

As discussed in Secs. 3 and 4, a larger n_{live} and smaller tol should be used for profile likelihood analyses with MULTINEST. We therefore used $n_{\text{live}} = 20,000$ and tol $= 1 \times 10^{-4}$. Flat priors were imposed on all 8 parameters with parameter ranges described above. This resulted in about 5.5 million likelihood evaluations (compared to 3 million likelihood evaluations for the GA method)

Figure 7: 2-D (a) marginalized posterior distributions and (b) profile likelihoods for the CMSSM parameters using all presently available constraints. The contours represent the 68% and 95% Bayesian credible regions in (a) and profile likelihood confidence intervals in (b). Flat priors with 20,000 live points were used with tol set to 1×10^{-4} (no smoothing applied to the plots). The mean and best-fit parameters values are shown by a black cross and a green circle, respectively.

taking 6 days on 10 2.4GHz CPUs and returned an evidence value of $\log(\mathcal{Z}) = -18.62 \pm 0.08$. The resultant 2-D marginalized posterior distributions and profile likelihoods in the $m_{1/2} - m_0$ and $\tan \beta - m_0$ planes are shown in Fig. 7. By comparing the posterior distribution in the upper left panel of Fig. 7 with the upper left panel of Fig. 13 in [14], obtained with the same setup but with $n_{\text{live}} = 4,000$ and tol = 0.5, it is apparent that the posterior is stable with respect to increasing the number of samples (Fig. 7 uses ~ 20 times more samples than the corresponding figure in Ref. [14]). Indeed, the evidence value obtained with this setup is $\log(\mathcal{Z}) = -18.28 \pm 0.14$. The difference in log evidence between the two setups is thus $\Delta \log(\mathcal{Z}) = -0.34 \pm 0.16$. For the log prior scans, the difference in log evidence between the two configurations is even lower,
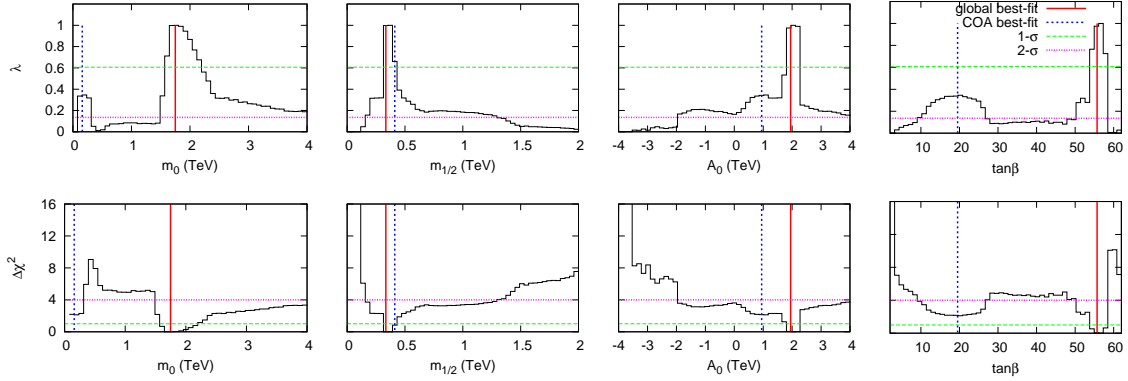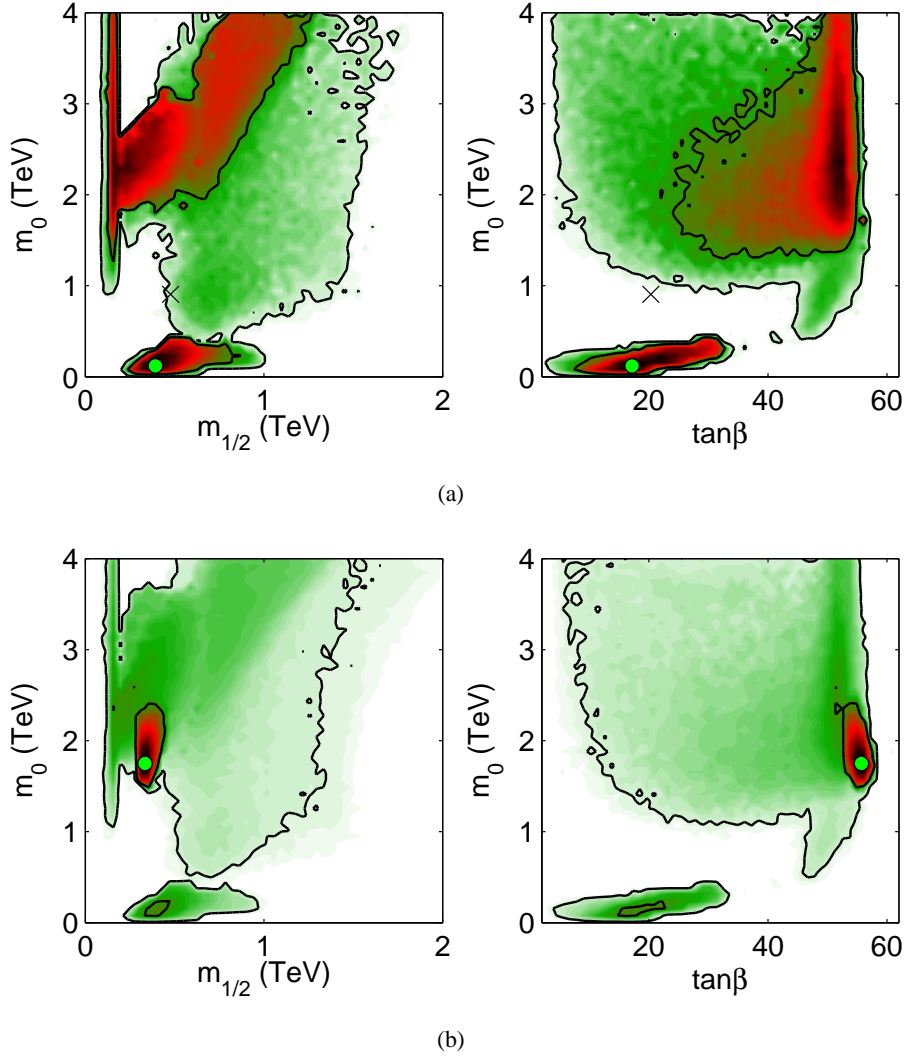
Figure 8: 1-D profile likelihoods for the CMSSM parameters normalized to the global best-fit point. The red solid and blue dotted vertical lines represent the global best-fit point ($\chi^2 = 9.26$, located in the focus point region) and the best-fit point found in the stau co-annihilation region ($\chi^2 = 11.38$) respectively. The upper and lower panel show the profile likelihood and $\Delta\chi^2$ values, respectively. Green (magenta) horizontal lines represent the 1σ (2σ) approximate confidence intervals. MULTINEST was run with flat priors, 20,000 live points and $\text{tol} = 1 \times 10^{-4}$.

$\Delta \log(\mathcal{Z}) = -0.08 \pm 0.14$. As this systematic difference is comparable with the statistical uncertainty, we can conclude that the posterior mass which is being missed by the standard, Bayesian setup is negligible (especially so given that the systematic difference is much less than the typical difference in thresholds on the Jeffreys' scale for the strength of evidence, which are at $\Delta \log(\mathcal{Z}) = 1.0, 2.5, 5.0$, see e.g. [27]). This leads us to reject the speculation in Ref. [26] regarding the validity of posterior probability distributions obtained with the standard configuration of MULTINEST.

The 2-D profile likelihood plots (bottom panels in Fig. 7) should be compared with Fig. 1(a) in Ref. [26], which were obtained with the GA[3]. Our best-fit point (i.e, the unconditional MLE) has $\chi^2 = 9.26$ and lies in the focus point region. Therefore, 1σ and 2σ regions are approximately delimited by $\chi^2 < 11.56$ and $\chi^2 < 15.43$, respectively. The MULTINEST best-fit point has a slightly better χ^2 value than the GA best-fit point ($\chi^2 = 9.35$). However, MULTINEST 1σ and 2σ regions are much larger than the corresponding GA regions. This is because MULTINEST has been able to map out more accurately the regions surrounding the best-fit than GA. By comparing 2-D posterior and profile likelihood plots in Fig. 7, we notice that most of the 1σ profile likelihood interval in the focus point region of the $m_{1/2} - m_0$ lies outside of the corresponding 2σ Bayesian credible interval. This is a clear sign that the high likelihood region in the focus point region is spike-like. These regions are not well explored by MULTINEST with the standard configuration ($\text{tol} = 0.5$). If one is interested in mapping out the posterior, however, this is fully justified, as these regions contribute almost nothing to the Bayesian evidence or the posterior probability. If however one wants to use MULTINEST for profile likelihood intervals, then it is imperative to map

---

[3]We have verified that reducing $n_{\text{live}} = 4,000$, while keeping $\text{tol} = 1 \times 10^{-4}$, leads to profile likelihoods very similar to the ones shown in Fig. 7, albeit slightly more noisy. We thus do not show those results, but we remark that this configuration reduces the computational effort by a factor of ~ 3 compared with the case where $n_{\text{live}} = 20,000$, without degrading the profile likelihood results appreciably.

Figure 9: 2-D profile likelihood functions for (a) log priors and (b) flat+log merged chains for CMSSM parameters. The contours represent the 68% and 95% profile likelihood confidence intervals. 20,000 live points were used with tol set to 1×10^{-4}. The mean and best-fit parameters values are shown by black cross and green circle respectively.

out those regions in much greater detail.

The values of global best-fit point and the best-fit point in stau co-annihilation region (COA) found by MULTINEST are listed in Tab. 1 and have similar parameter values to ones found by GA. Fig. 8, shows the 1-D profile likelihood. This figure should be compared with the corresponding 1-D profile likelihoods obtained with the GA method in Fig. 3 of Ref. [26].

In principle, the profile likelihood does not depend on the choice of priors. However, in order to explore the parameter space using any Monte Carlo technique, a set of priors needs to be defined. Any numerical sampling technique effectively defines a metric on the parameter space of interest, and different choices of this metric will generally lead to different regions of the parameter space
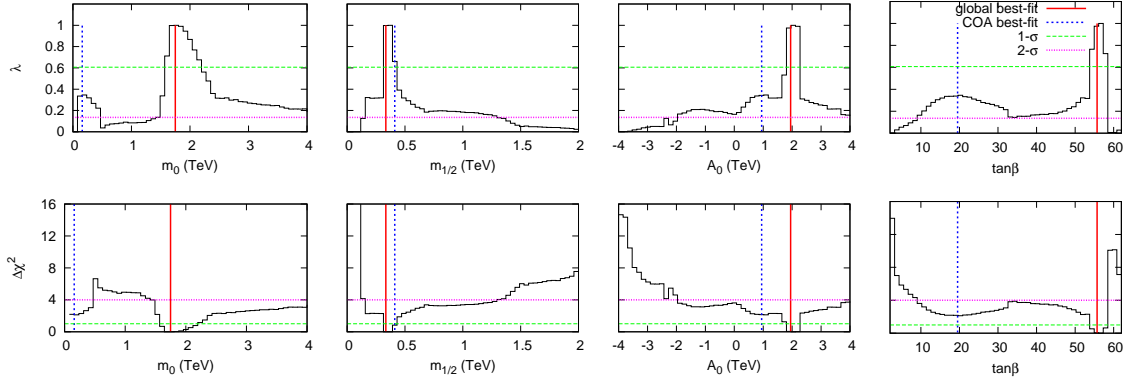
Figure 10: 1-D profile likelihoods for the CMSSM parameters for flat+log merged chains, normalized to the global best-fit point. The red solid and blue dotted vertical lines represent the global best-fit point ($\chi^2 = 9.26$, located in the focus point region) and the best-fit point found in the stau co-annihilation region ($\chi^2 = 11.38$) respectively. The upper and lower panel show the profile likelihood and $\Delta\chi^2$ values, respectively. Green (magenta) horizontal lines represent the 1σ (2σ) approximate confidence intervals. MULTINEST was run with flat priors, 20,000 live points and $\mathrm{tol} = 1 \times 10^{-4}$.

	Global BFP (located in FP region)	COA BFP
Model and Nuisance Parameters		
m_0	1748.7 GeV	159.0 GeV
$m_{1/2}$	334.1 GeV	411.4 GeV
A_0	1949.5 GeV	946.2 GeV
$\tan\beta$	55.7	19.53
m_t	173.0 GeV	173.3 GeV
$m_b(m_b)^{\overline{MS}}$	4.19 GeV	4.20 GeV
$\alpha_s^{\overline{MS}}(M_Z)$	0.1178	0.1182
$1/\alpha^{\overline{MS}}(M_Z)$	127.952	127.959
Observables		
m_W	80.367 GeV	80.371 GeV
$\sin^2\theta_{\mathrm{eff}}$	0.23156	0.23153
$\delta a_\mu^{\mathrm{SUSY}} \times 10^{10}$	6.8	13.7
$BR(\overline{B} \to X_s\gamma) \times 10^4$	3.45	2.95
ΔM_{B_s}	17.31 ps^{-1}	19.0 ps^{-1}
$BR(\overline{B}_u \to \tau\nu) \times 10^4$	1.38	1.46
$\Omega_\chi h^2$	0.11408	0.11038
$BR(\overline{B}_s \to \mu^+\mu^-)$	4.08×10^{-8}	3.89×10^{-8}

Table 1: Best-fit parameter and observable values found from flat+log merged chains. These quantities are shown for both the global best-fit point, located in the focus point (FP) region, as well as the best-fit point in the stau co-annihilation (COA) region.

to be explored in greater or lesser detail. Therefore, different choice of priors might lead to slightly different profile likelihoods. Uniform priors in $\log(m_0)$ and $\log(m_{1/2})$ (called log priors) are often employed in Bayesian analyses of CMSSM as the masses are scale parameters. In order to get a feeling for the robustness of our profile likelihoods, we show the 2-D profile likelihoods from a

scan using $\log$ priors in Fig. 9(a). As expected, the parameter space with lower m_0 and $m_{1/2}$ values has been explored more thoroughly. The best-fit point found has $\chi^2 = 11.44$ and lies in the stau co-annihilation region (COA), thus this scan with $\log$ priors has missed a a higher likelihood point in the focus point region. Conversely, the scan with flat priors has explored the COA region in similar detail. Thus, while these tunings of MULTINEST are clearly superior for profile likelihood analysis than the default settings for Bayesian analysis, there remains sensitivity to the choice of priors used in the scans as previously seen in the literature.

We can obtain more robust profile likelihoods by simply merging the flat and $\log$ prior scans, the performing the profiling over the joint set of samples (encompassing about 10.85 million points). This does not come at a greater computational cost given that a responsible Bayesian analyses would estimate sensitivity to the choice of prior as well. The resulting 2-D and 1-D profile likelihoods are shown in Figs. 9(b) and 10 respectively. The most obvious difference between the merged and flat profile likelihood distributions is the appearance of the funnel region at low $m_{1/2}$ values which was not explored thoroughly with flat priors because of the its smaller probability mass. Apart from this, it is clear that the merged profile likelihood distributions are quite similar to the distributions obtained with flat prior (see Fig. 7(b)).

In the left panel of Fig. 11, we show the 2-D profile likelihood functions spin-independent scattering cross-section of the neutralino and a proton $\sigma_{\rm p}^{\rm SI}$ versus the neutralino mass m_χ along with the latest experimental exclusion limits at the 90% confidence level from CDMS-II [41] and XENON100 [42], derived under standard halo assumptions (and in particular, assuming a local DM density $\rho_\chi = 0.3$ GeV/cm^3). Our global best-fit point has quite a large cross-section (1.85×10^{-7} pb) and almost the entire 1σ region in the FP region would already be excluded by current experimental limits if one takes them at face value. However, there are several uncertainties that need to be accounted for in order to achieve a robust exclusion: astrophysical uncertainties in the local DM density and velocity distribution [43–45], systematic uncertainties related to our position in the Milky Way halo [46], clumpiness of the DM distribution and impact of baryonic physics [47], hadronic matrix elements uncertainties [48]. Taken together, all those uncertainties can easily exceed an order of magnitude. For this reason, we have chosen not to include the current direct detection limits in the likelihood function, although we notice qualitatively that the FP best-fit point is coming under pressure from such limits. On the other hand, the best-fit $\sigma_{\rm p}^{\rm SI}$ in the COA region is relatively small (2.02×10^{-9} pb) and is well below the current experimental limits.

Regarding indirect detection prospects, we plot the 2-D profile likelihood functions for the velocity-averaged neutralino self-annihilation cross-section $\langle \sigma v \rangle$ against the neutralino mass m_χ in the right panel of Fig. 11. Although there is a small region of the parameter space around 300 GeV $\leq m_\chi \leq$ 500 GeV and $-27.5 < \log_{10} \langle \sigma v \rangle < -26.5$ with high likelihood points found by GA in [26] and missed by MULTINEST, in general we have mapped the profile likelihood more thoroughly as can be seen by comparing the right panel of Fig. 11 with Fig. 6(a) in [26]. Our global best-fit point has $\langle \sigma v \rangle = 2.29 \times 10^{-26}$cm^3 s^{-1} which is very similar to the global best-fit point found in [26] and therefore we reach the same conclusion, namely that it might be possible to cover part of this high-likelihood FP region in the future with the Large Area Telescope (LAT)[49] aboard the Fermi gamma-ray space telescope.
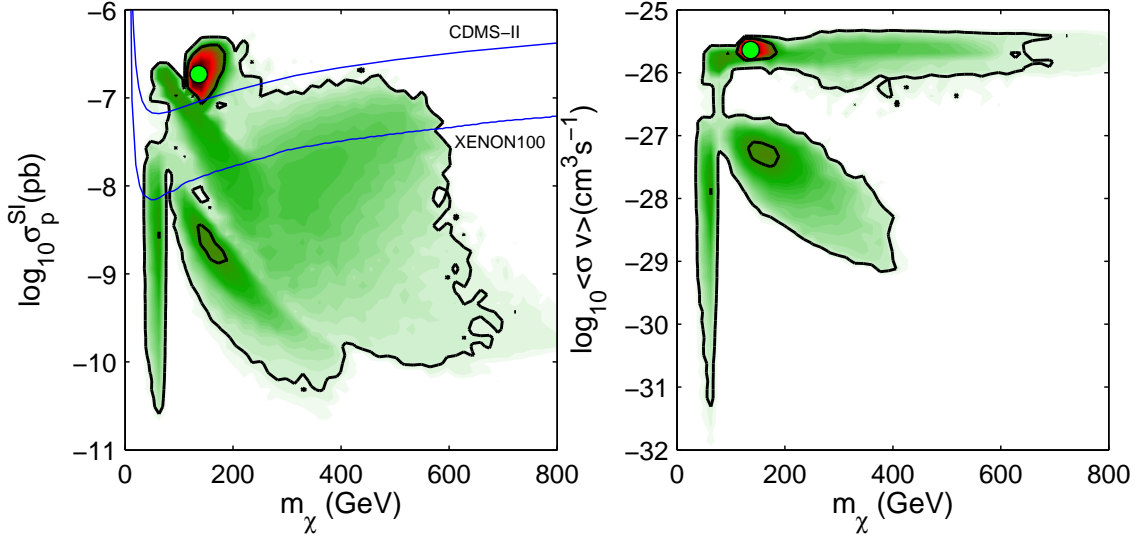
Figure 11: 2-D profile likelihood functions for flat+log merged chains for spin-independent scattering cross-section of the neutralino and a proton $\sigma_{\mathrm{p}}^{\mathrm{SI}}$ versus the neutralino mass m_χ (left panel) and velocity-averaged neutralino self-annihilation cross-section $\langle \sigma v \rangle$ versus the neutralino mass m_χ (right panel). The contours represent the 68% and 95% profile likelihood confidence intervals. 20,000 live points were used with tol set to 1×10^{-4}. The best-fit parameters values are shown by green circle. The latest experimental exclusion limits at the 90% confidence level from CDMS-II [41] and XENON100 [42] are also shown in the left panel.

	Global BFP (located in FP region)	COA BFP
m_χ	136.34 GeV	156.46 GeV
Direct Detection		
σ_{p}^{SI}	1.85×10^{-7} pb	2.02×10^{-9} pb
σ_{p}^{SD}	1.96×10^{-6} pb	4.05×10^{-6} pb
σ_{n}^{SD}	1.34×10^{-6} pb	3.00×10^{-6} pb
Indirect Detection		
$\langle \sigma v \rangle$	2.29×10^{-26} cm^3 s^{-1}	4.89×10^{-28} cm^3 s^{-1}

Table 2: Best-fit neutralino mass and dark matter direct and indirect detection observables. These quantities are shown for both the global best-fit point, located in the focus point (FP) region, as well as the best-fit point in the stau co-annihilation (COA) region.

6. Conclusions

As the LHC impinges on the most anticipated regions of SUSY parameter space, the need for statistical techniques that will be able to cope with the complexity of SUSY phenomenology is greater than ever. Early on, the data will not be strong enough to dominate the impact of priors on Bayesian inference; thus, the field is preparing to present the complementary information provided by Bayesian techniques and the traditional results from the profile likelihood.

In this paper we have shown that, when configured appropriately, MULTINEST can be succesfully employed for approximating the profile likelihood functions, even though it was primarily designed for Bayesian analyses. In particular, we have demonstrated that it is important to use a termination criterion that allows MULTINEST to explore high-likelihood regions. We have also shown that the likelihood spikes which play a key role in the profile likelihood and are missed by MULTINEST when it is run in Bayesian analysis mode, have no bearing on the posterior distribution. Despite these significant improvements, we do find that the prior used in the MULTINEST scan can influence the profile likelihood ratio, particularly if the global maximum of the likeli-

hood is spiky and in a region with extremely low prior probability. This can be partially abated by merging samples from scans with different priors, which are often available from studies of prior dependence in the corresponding Bayesian results.

We compared our results and conclusions to those reported in Ref. [26], which used genetic algorithms for mapping the profile likelihood functions. The authors of that study concluded that their sampling algorithm resulted in better approximations to the profile likelihood than those obtained with samples from Bayesian techniques like MCMC and MULTINEST. They also speculated that MULTINEST might not be sufficiently accurate even for Bayesian analyses. Our results indicate that when run appropriately, MULTINEST provides a significantly better approximation to the profile likelihood than the genetic algorithm method presented in Ref. [26], particularly in the exploration of the parameter space near the boundaries of the interval, albeit at a slightly increased computational cost. There are also indications that MULTINEST might be more reliable in identifying the overall best-fit point than the genetic algorithm, as the latter only found the overall best-fit in 1 out of 10 runs (see Fig. 8 in Ref. [26]).

We end by speculating that the ideal algorithm for profile likelihood based analyses in complex, multi-modal problems such as supersymmetric parameter spaces may require a hybrid of global scans like MCMC and MULTINEST together with dedicated maximization packages like MINUIT [50]. The challenge for most maximization packages – whether they are based on gradient descent, conjugate gradient descent, or the EM-algorithm – is that they can get stuck in local maxima. Hence, it seems clear that a more global scan will be required. However, once the scanning algorithm has located a basin of attraction the maximization packages are likely to be the best tool for refining the local maxima. As we discussed here, MULTINEST with a low tolerance acts as a maximizer, and can be subject to similar problems with local maxima. The nested structure provided by the nested sampling algorithm may provide an efficient and practical way for defining those basins of attraction, reducing the number of points that will initiate a dedicated maximization. In future work we wish to investigate these hybrid approaches and compare MINUIT and MULTINEST for this dedicated maximization step.

Acknowledgements

References

[1] The CMS Collaboration, *Search for Supersymmetry in pp Collisions at 7 TeV in Events with*

Jets and Missing Transverse Energy, `arXiv:1101.1628`.

[2] R. Roberts and L. Roszkowski, *Implications for minimal supersymmetry from grand unification and the neutrino relic abundance*, *Phys.Lett.* **B309** (1993) 329–336, [`hep-ph/9301267`].

[3] G. L. Kane, C. F. Kolda, L. Roszkowski, and J. D. Wells, *Study of constrained minimal supersymmetry*, *Phys.Rev.* **D49** (1994) 6173–6210, [`hep-ph/9312272`].

[4] J. R. Ellis, K. A. Olive, Y. Santoso, and V. C. Spanos, *Likelihood analysis of the CMSSM parameter space*, *Phys.Rev.* **D69** (2004) 095004, [`hep-ph/0310356`].

[5] S. Profumo and C. E. Yaguna, *A Statistical analysis of supersymmetric dark matter in the MSSM after WMAP*, *Phys.Rev.* **D70** (2004) 095004, [`hep-ph/0407036`].

[6] E. A. Baltz and P. Gondolo, *Markov Chain Monte Carlo Exploration of Minimal Supergravity with Implications for Dark Matter*, *Journal of High Energy Physics* **10** (Oct., 2004) 52–+, [`hep-ph/04`].

[7] R. Ruiz de Austri, R. Trotta, and L. Roszkowski, *A Markov chain Monte Carlo analysis of the CMSSM*, *Journal of High Energy Physics* **5** (May, 2006) 2–+, [`hep-ph/06`].

[8] B. Allanach and C. Lester, *Multi-dimensional mSUGRA likelihood maps*, *Phys.Rev.* **D73** (2006) 015013, [`hep-ph/0507283`].

[9] J. Skilling, *Nested Sampling*, in *American Institute of Physics Conference Series* (R. Fischer, R. Preuss, and U. V. Toussaint, eds.), pp. 395–405, Nov., 2004.

[10] D. Sivia and J. Skilling, *Data Analysis A Bayesian Tutorial*. Oxford University Press, 2006.

[11] S. S. AbdusSalam, B. C. Allanach, M. J. Dolan, F. Feroz, and M. P. Hobson, *Selecting a model of supersymmetry breaking mediation*, Phys.Rev.D **80** (Aug., 2009) 035017–+, [`arXiv:0906.0957`].

[12] S. S. Abdussalam, B. C. Allanach, F. Quevedo, F. Feroz, and M. Hobson, *Fitting the phenomenological MSSM*, Phys.Rev.D **81** (May, 2010) 095012–+, [`arXiv:0904.2548`].

[13] F. Feroz, M. P. Hobson, L. Roszkowski, R. Ruiz de Austri, and R. Trotta, *Are BR($B \to s\gamma$) and $(g-2)_\mu$ consistent within the Constrained MSSM?*, *ArXiv e-prints* (Mar., 2009) [`arXiv:0903.2487`].

[14] R. Trotta, F. Feroz, M. Hobson, L. Roszkowski, and R. Ruiz de Austri, *The impact of priors and observables on parameter inferences in the constrained MSSM*, *Journal of High Energy Physics* **12** (Dec., 2008) 24–+, [`arXiv:0809.3792`].

[15] F. Feroz, B. C. Allanach, M. Hobson, S. S. Abdus Salam, R. Trotta, and A. M. Weber, *Bayesian selection of sign μ within mSUGRA in global fits including WMAP5 results*, *Journal of High Energy Physics* **10** (Oct., 2008) 64–+, [`arXiv:0807.4512`].

[16] M. J. White and F. Feroz, *MSSM dark matter measurements at the LHC without squarks and sleptons*, *Journal of High Energy Physics* **7** (July, 2010) 64–+, [arXiv:1002.1922].

[17] G. Bertone, K. Kong, R. Ruiz de Austri, and R. Trotta, *Global fits of the Minimal Universal Extra Dimensions scenario*, *ArXiv e-prints* (Oct., 2010) [arXiv:1010.2023].

[18] L. Roszkowski, R. Ruiz de Austri, and R. Trotta, *Efficient reconstruction of constrained MSSM parameters from LHC data: A case study*, Phys.Rev.D **82** (Sept., 2010) 055003–+, [arXiv:0907.0594].

[19] P. Scott, J. Conrad, J. Edsjö, L. Bergström, C. Farnier, and Y. Akrami, *Direct constraints on minimal supersymmetry from Fermi-LAT observations of the dwarf galaxy Segue 1*, *Journal of Cosmology and Astroparticle Physics* **1** (Jan., 2010) 31–+, [arXiv:0909.3300].

[20] J. Ripken, J. Conrad, and P. Scott, *Direct constraints on the CMSSM using H.E.S.S. observations of the Galactic Centre and the Sagittarius dwarf galaxy*, *ArXiv e-prints* (Dec., 2010) [arXiv:1012.3939].

[21] L. Roszkowski, R. Ruiz de Austri, R. Trotta, Y.-L. S. Tsai, and T. A. Varley, *Global fits of the Non-Universal Higgs Model*, arXiv:0903.1279.

[22] B. C. Allanach, K. Cranmer, C. G. Lester, and A. M. Weber, *Natural priors, CMSSM fits and LHC weather forecasts*, *Journal of High Energy Physics* **8** (Aug., 2007) 23–+, [arXiv:0705.0487].

[23] F. Feroz and M. P. Hobson, *Multimodal nested sampling: an efficient and robust alternative to Markov Chain Monte Carlo methods for astronomical data analyses*, MNRAS **384** (Feb., 2008) 449–463, [0704.3704].

[24] F. Feroz, M. P. Hobson, and M. Bridges, *MULTINEST: an efficient and robust Bayesian inference tool for cosmology and particle physics*, MNRAS **398** (Oct., 2009) 1601–1614, [arXiv:0809.3437].

[25] M. Bridges, K. Cranmer, F. Feroz, M. Hobson, R. Ruiz de Austri, and R. Trotta, *A coverage study of the CMSSM based on ATLAS sensitivity using fast neural networks techniques*, *Journal of High Energy Physics* **3** (Mar., 2011) 12–+, [arXiv:1011.4306].

[26] Y. Akrami, P. Scott, J. Edsjö, J. Conrad, and L. Bergström, *A profile likelihood analysis of the constrained MSSM with genetic algorithms*, *Journal of High Energy Physics* **4** (Apr., 2010) 57–+, [arXiv:0910.3950].

[27] R. Trotta, *Bayes in the sky: Bayesian inference and model selection in cosmology*, *Contemp. Phys.* **49** (2008) 71–104, [arXiv:0803.4089].

[28] S. Wilks, *The large-sample distribution of the likelihood ratio for testing composite hypotheses*, *Ann. Math. Statist.* **9** (1938) 60–2.

[29] M. Bartlett, *Properties of sufficiency and statistical tests*, *Journal of the Royal Statistical Society. A* **160** (1937).

[30] D. R. Cox and N. Reid, *Parameter orthogonality and approximate conditional inference*, *Journal of the Royal Statistical Society. Series B (Methodological)* **49** (1987), no. 1 1–39.

[31] O. Barndorff-Nielsen and D. Cox, *Bartlett adjustments to the likelihood ratio statistic to the distribution of the maximum likelihood estimator*, *Journal of the Royal Statistical Society. Series B* **46** (1984).

[32] G. R. Farrar and P. Fayet, *Phenomenology of the Production, Decay, and Detection of New Hadronic States Associated with Supersymmetry*, *Phys. Lett.* **B76** (1978) 575–579.

[33] S. Dimopoulos and H. Georgi, *Softly Broken Supersymmetry and SU(5)*, *Nucl. Phys.* **B193** (1981) 150.

[34] L. Alvarez-Gaume, J. Polchinski, and M. B. Wise, *Minimal Low-Energy Supergravity*, *Nucl. Phys.* **B221** (1983) 495.

[35] J. P. Conlon and F. Quevedo, *Gaugino and scalar masses in the landscape*, *Journal of High Energy Physics* **6** (June, 2006) 29–+, [`hep-th/06`].

[36] A. Brignole, L. E. Ibáñez, and C. Muñoz, *Towards a theory of soft terms for the supersymmetric standard model*, *Nuclear Physics B* **422** (July, 1994) 125–171, [`hep-ph/93`].

[37] O. Buchmueller, R. Cavanaugh, A. de Roeck, J. R. Ellis, H. Flaecher, S. Heinemeyer, G. Isidori, K. A. Olive, F. J. Ronga, and G. Weiglein, *Likelihood functions for supersymmetric observables in frequentist analyses of the CMSSM and NUHM1*, *European Physical Journal C* **64** (Dec., 2009) 391–415, [`arXiv:0907.5568`].

[38] L. Roszkowski, R. Ruiz de Austri, and R. Trotta, *Implications for the Constrained MSSM from a new prediction for $b \to s\gamma$*, *Journal of High Energy Physics* **7** (July, 2007) 75–+, [`arXiv:0705.2012`].

[39] L. Roszkowski, R. Ruiz de Austri, and R. Trotta, *Efficient reconstruction of CMSSM parameters from LHC data - A case study*, *Phys. Rev.* **D82** (2010) 055003, [`arXiv:0907.0594`].

[40] Y. Akrami, C. Savage, P. Scott, J. Conrad, and J. Edsjö, *Statistical coverage for supersymmetric parameter estimation: a case study with direct detection of dark matter*, *ArXiv e-prints* (Nov., 2010) [`arXiv:1011.4297`].

[41] Z. Ahmed and et al, *Search for Weakly Interacting Massive Particles with the First Five-Tower Data from the Cryogenic Dark Matter Search at the Soudan Underground Laboratory*, *Physical Review Letters* **102** (Jan., 2009) 011301–+, [`arXiv:0802.3530`].

[42] XENON100 Collaboration, E. Aprile, K. Arisaka, F. Arneodo, A. Askin, L. Baudis, A. Behrens, K. Bokeloh, E. Brown, T. Bruch, G. Bruno, J. M. R. Cardoso, W. . Chen, B. Choi, D. Cline, E. Duchovni, S. Fattori, A. D. Ferella, F. Gao, K. . Giboni, E. Gross, A. Kish, C. W. Lam, J. Lamblin, R. F. Lang, C. Levy, K. E. Lim, Q. Lin, S. Lindemann,

M. Lindner, J. A. M. Lopes, K. Lung, T. Marrodan Undagoitia, Y. Mei, A. J. Melgarejo Fernandez, K. Ni, U. Oberlack, S. E. A. Orrigo, E. Pantic, R. Persiani, G. Plante, A. C. C. Ribeiro, R. Santorelli, J. M. F. dos Santos, G. Sartorelli, M. Schumann, M. Selvi, P. Shagin, H. Simgen, A. Teymourian, D. Thers, O. Vitells, H. Wang, M. Weber, and C. Weinheimer, *Dark Matter Results from 100 Live Days of XENON100 Data*, *ArXiv e-prints* (Apr., 2011) [arXiv:1104.2549].

[43] L. E. Strigari and R. Trotta, *Reconstructing WIMP Properties in Direct Detection Experiments Including Galactic Dark Matter Distribution Uncertainties*, *JCAP* **0911** (2009) 019, [arXiv:0906.5361]. * Brief entry *.

[44] C. McCabe, *The Astrophysical Uncertainties Of Dark Matter Direct Detection Experiments*, *Phys.Rev.* **D82** (2010) 023530, [arXiv:1005.0579].

[45] A. M. Green, *Dependence of direct detection signals on the WIMP velocity distribution*, *JCAP* **1010** (2010) 034, [arXiv:1009.0916].

[46] M. Pato, O. Agertz, G. Bertone, B. Moore, and R. Teyssier, *Systematic uncertainties in the determination of the local dark matter density*, *Phys.Rev.* **D82** (2010) 023531, [arXiv:1006.1322].

[47] F.-S. Ling, E. Nezri, E. Athanassoula, and R. Teyssier, *Dark Matter Direct Detection Signals inferred from a Cosmological N-body Simulation with Baryons*, *JCAP* **1002** (2010) 012, [arXiv:0909.2028]. * Brief entry *.

[48] J. R. Ellis, K. A. Olive, and C. Savage, *Hadronic Uncertainties in the Elastic Scattering of Supersymmetric Dark Matter*, *Phys.Rev.* **D77** (2008) 065026, [arXiv:0801.3656].

[49] W. B. Atwood, A. A. Abdo, M. Ackermann, W. Althouse, B. Anderson, M. Axelsson, L. Baldini, J. Ballet, D. L. Band, G. Barbiellini, and et al., *The Large Area Telescope on the Fermi Gamma-Ray Space Telescope Mission*, ApJ **697** (June, 2009) 1071–1102, [arXiv:0902.1089].

[50] F. James and M. Roos, *Minuit: A System for Function Minimization and Analysis of the Parameter Errors and Correlations*, *Comput. Phys. Commun.* **10** (1975) 343–367.