

# Should we still believe in constrained supersymmetry? <sup>†</sup>

---

**Csaba Balázs**

*School of Physics, Monash University, Melbourne, Victoria 3800 Australia*

*E-mail: [csaba.balazs@monash.edu](mailto:csaba.balazs@monash.edu)*

**Andy Buckley**

*School of Physics and Astronomy, University of Edinburgh, Edinburgh, EH9 3JZ UK*

*E-mail: [andy.buckley@ed.ac.uk](mailto:andy.buckley@ed.ac.uk)*

**Daniel Carter**

*School of Physics, Monash University, Melbourne, Victoria 3800 Australia*

*E-mail: [daniel.carter@monash.edu](mailto:daniel.carter@monash.edu)*

**Benjamin Farmer<sup>‡</sup>**

*School of Physics, Monash University, Melbourne, Victoria 3800 Australia*

*E-mail: [benjamin.farmer@monash.edu](mailto:benjamin.farmer@monash.edu)*

**Martin White**

*School of Physics, The University of Melbourne, Melbourne, Victoria 3010 Australia*

*E-mail: [mwhi@unimelb.edu.au](mailto:mwhi@unimelb.edu.au)*

We calculate partial Bayes factors to quantify how the feasibility of the constrained minimal supersymmetric standard model (CMSSM) has changed in the light of a series of observations. We take as “training” data the approximate knowledge that was available before LEP, and take our comparison model to be the Standard Model with a simple dark matter candidate. Partial Bayes factors are then computed, using as “inference” data the LEP2 Higgs constraints, 2011 XENON100 dark matter constraints, 2011 LHC supersymmetry search results, and the early 2012 LHC Higgs search results. We find that LEP and the LHC strongly shatter our trust in the CMSSM, reducing its posterior odds by a factor of approximately two orders of magnitude. This conclusion is robust under variation of priors, but may be avoided if the CMSSM is not required to explain the  $(g - 2)_\mu$  anomaly.

*36th International Conference on High Energy Physics*

*4-11 July 2012*

*Melbourne, Australia*

---

<sup>†</sup>This research was funded in part by the ARC Centre of Excellence for Particle Physics at the Tera-scale, and in part by the Project of Knowledge Innovation Program (PKIP) of Chinese Academy of Sciences Grant No. KJCX2.YW.W10. AB acknowledges the support of the Scottish Universities Physics Alliance. The use of Monash Sun Grid (MSG) and Edinburgh ECDF high-performance computing facilities is also gratefully acknowledged. Most numerical calculations were performed on the Australian National Computing Infrastructure (NCI) National Facility SGI XE cluster and Multi-modal Australian ScienceS Imaging and Visualisation Environment (MASSIVE) cluster.

<sup>‡</sup>Speaker.

## 1. Foreword

This proceedings paper is a summary of the results obtained in ref. [1]. To achieve the necessary brevity we will refer the reader to the full paper for many technical details. We also note that the statistical terminology used presently has been significantly altered from the original work, with the intention of improving its alignment with the relevant statistics literature. We hope these changes render the present description of our statistical methods superior to the original in clarity.

An extended version of this summary will appear in the proceedings of conference C12-06-10.

## 2. Introduction

Supersymmetry is an attractive and robust extension of the Standard Model (SM) of particle physics, the most well-studied version of which is the constrained minimal supersymmetric standard model (CMSSM). Based on experimental data, an extensive literature delineates the regions of the CMSSM where its parameters can most probably fall, however the more fundamental question is that of model selection; how likely is it that the CMSSM is an accurate description of weak-scale physics? To address this question one has several options. The most common frequentist measure is the p-value, which is the probability that more extreme<sup>1</sup> data than that observed would occur assuming the hypothesis in question to be true [2]. In the Bayesian approach model selection is based on the Bayes factor, and requires comparison to alternative hypotheses. [3].

To compute Bayes factors we need to first compute the marginalised likelihood  $P(\text{data}|H_i)$ , also called the “evidence”, for each model hypothesis  $H_i$ , e.g.

$$P(\text{data}|H_i) = \int d\theta P(\text{data}|\theta, H_i)P(\theta|H_i). \quad (2.1)$$

This requires the specification of a prior probability density  $P(\theta|H_i)$  over the parameters  $\theta$  of each model, which must reflect our knowledge (or lack thereof) of the parameters before knowing *data*. In the case where our prior knowledge is weak it is generally very difficult to specify a prior which both accurately expresses this knowledge and is “proper”, in the sense that its integral can be normalised to 1 (indeed the first criterion alone is difficult to achieve). Common choices of simple prior, such as uniform or logarithmically flat distributions, as well as most formal minimally-informative priors (such as maximum entropy [4] or “reference” [5] priors), are improper.

For inference of the model parameters themselves the use of such improper priors is generally unproblematic, as they may still result in proper posterior distributions once combined with sufficiently powerful data, however they cause major problems for model comparisons because they cannot be used to compute marginalised likelihoods. A naïve fix may be to specify cutoffs to render the original priors proper, however unless the cutoff approximates some actual prior knowledge it merely introduces an arbitrary constant into the marginalised likelihood, which thus remains useless for model comparison.

This problem is well known and a number of solutions have been proposed [6], however they generally depart from pure Bayesian methods. In this work we adopt the simplest of these solutions, which is to use so-called “partial” Bayes factors, which, although more limited in the inferences

<sup>1</sup>Here ‘more extreme’ can be defined in numerous ways.

that can be derived from them compared to more advanced methods, retain a pure Bayesian interpretation.

To compute partial Bayes factors, one takes note of the aforementioned fact that it is generally possible to obtain a proper posterior from an improper prior by incorporating sufficiently strong data via a Bayesian “update” (i.e. an iteration through Bayes’ theorem). The idea is then to use some portion of the available data to update the improper priors in this fashion, and then use the resulting posterior together with the remaining data to compute a Bayes factor as normal. The resultant Bayes factor is only “part” of the full Bayes factor that would have resulted from using all the data (if it could have been computed), and so it is termed “partial”.

To illustrate the procedure explicitly, consider the division of the available data into two sets; a “training” set  $d_1$ , and an “inference” set  $d_2$ . In principle many such divisions are possible, and each will result in a different partial Bayes factor (a “flaw” which alternate methods attempt to remedy, generally by combining the various possible partial Bayes factors in some way, in conjunction with specifying rules for choosing the divisions to use), however in our situation a roughly chronological separation is quite natural and has a useful interpretation. We describe our chosen separation in section 3. Next consider the ordinary Bayes factor, for a test of some model  $H$  against an alternate  $H_{\text{alt}}$ , for such a set of data:

$$B(d_2, d_1) = \frac{P(d_2, d_1 | H)}{P(d_2, d_1 | H_{\text{alt}})} = \frac{P(d_2 | d_1, H) P(d_1 | H)}{P(d_2 | d_1, H_{\text{alt}}) P(d_1 | H_{\text{alt}})} = B(d_2 | d_1) B(d_1) \quad (2.2)$$

Here  $B(d_2 | d_1)$  is the partial Bayes factor obtained by “training” the model priors with  $d_1$  and then performing the comparison using  $d_2$ , while  $B(d_1)$  is uncomputable or unreliable since to compute it we need to integrate over an improper prior. The product  $B(d_2 | d_1) B(d_1)$  is the standard (uncomputable) Bayes factor  $B(d_2, d_1)$ , but by discarding the uncomputable piece  $B(d_1)$  we are left with at least some inferential power, and as a bonus our sensitivity to the original improper prior is reduced (in proportion to the informativeness of  $d_1$ ).

Since we have stuck to the Bayesian rules there exists a Bayesian interpretation of  $B(d_2 | d_1)$ . Consider its place in computing the posterior odds for  $H$  vs  $H_{\text{alt}}$ :

$$\text{Odds}(H : H_{\text{alt}} | d_2, d_1) = B(d_2 | d_1) \text{Odds}(H : H_{\text{alt}} | d_1) = B(d_2 | d_1) B(d_1) \text{Odds}(H : H_{\text{alt}}) \quad (2.3)$$

The prior odds,  $\text{Odds}(H : H_{\text{alt}})$ , cannot be computed by any standard Bayesian means and must be supplied based on prior knowledge. The combination  $\text{Odds}(H : H_{\text{alt}} | d_1) = B(d_1) \text{Odds}(H : H_{\text{alt}})$  is no more computable for its extra dependence on the uncomputable  $B(d_1)$ , and so, instead of considering their personal  $\text{Odds}(SM : CMSSM)$ , we invite the reader to instead directly consider their personal  $\text{Odds}(H : H_{\text{alt}} | d_1)$ . The partial Bayes factor  $B(d_2 | d_1)$  can then be interpreted as the factor required to correctly update these personal odds to take into account the newly learned data  $d_2$  (assuming of course that the reader roughly accepts our adopted model priors and our assumptions regarding the nature of  $d_1$  and  $d_2$ ).

### 3. Training and inference data

We describe in this section the “training” data used to convert our initially improper parameter space priors into informative proper priors via a Bayesian update, and the “inference” data which is used in conjunction with the trained priors to construct partial Bayes factors.

First, our “training” data includes the WMAP measurement of the WIMP relic density  $\Omega h_\chi^2$  (used as a central constraint on the neutralino relic density), electroweak precision measurements, limits on rare B and D decays, the LEP2 lower bounds on sparticle masses, and the muon  $g - 2$  anomaly.

For inference data we use in turn: the LEP2 Higgs search limits; the 2011 XENON100 limits on the WIMP-nucleon scattering rate [7] together with the 2011 LHC  $1 \text{ fb}^{-1}$  zero-lepton sparticle search limits [8]; and the February 2012 LHC Higgs search results [9]. Partial Bayes factors are computed for the addition of each of these pieces of data in turn, giving us three such factors plus a ‘total’ partial Bayes factor which is the cumulative effect of the total inference data set. For further details of the training and inference data we refer the reader to the full description of the analysis given in ref. [1].

We also compute PBFs with  $(g - 2)_\mu$  removed from the training set, to investigate its influence and to consider the consequences of it being explained within the SM. This action destroys any hope of achieving even a weakly proper trained prior for the CMSSM, so this set of PBFs can only be interpreted as describing the damage to the  $M_0, M_{1/2} < 2 \text{ TeV}$  region of the CMSSM, not as damage to the CMSSM as a whole.

#### 4. Results

Two previously studied priors were used to compute partial Bayes factors (to allow an investigation of prior sensitivity and to remain consistent with previous literature): the ‘log’ prior [10], which is flat in  $A_0$  and  $\tan\beta$ , and flat in the logarithm of  $M_0$  and  $M_{1/2}$ , and the ‘naturalness’ prior [11] (specifically the ‘CCR’ version of this prior [12]), which assigns low prior weight to fine-tuned regions of CMSSM parameter space. The  $\mu < 0$  branch is strongly disfavoured [13] so we scan only the  $\mu > 0$  branch to reduce computational demand. We scan  $M_0$  and  $M_{1/2}$  below 2 TeV since  $(g - 2)_\mu$  sufficiently excludes model points outside this range, with  $-3 \text{ TeV} < A_0 < 4 \text{ TeV}$  and  $0 < \tan\beta < 62$  ( $\tan\beta = 0$  is of course unphysical so receives zero weight after training). The top quark mass is also scanned using a Gaussian prior with mean 172.9 GeV and standard deviation 1.1 GeV.

Scans were performed using `MultiNest v2.12`, with the CMSSM spectrum generated by `ISAJET v7.81` and further training observables computed by `micrOmegas v2.4.Q` and `SuperISO v3.1`. The LEP Higgs search likelihood is implemented with a simple error function approximation, while the LHC Higgs search likelihood is reconstructed from ATLAS results [9] using asymptotic approximations and utilising Higgs branching ratios computed by `HDECAY v4.43`. The LHC sparticle search likelihood is implemented using a Bayesian neural network trained using 50,000 model points sampled from the full 4D parameter space, using `Herwig++ 2.5.2` to generate 15,000 Monte Carlo events per model point, with `Delphes 1.9` providing a fast simulation of the ATLAS detector, and with the total SUSY production cross section computed at next-to-leading order by `PROSPINO 2.1`, in a simulation chain tuned to match the ATLAS  $1 \text{ fb}^{-1}$  zero-lepton jets+MET search described in ref. [8]. For further details and references related to these codes we refer readers to the full study [1].

The marginal likelihood values for the SM-like comparison model are computed assuming all parameters except the physical Higgs mass to be fixed, excluding parameters involved in the

Knowledge change	log prior		natural prior		Weight of evidence (against CMSSM)
	PBF	Discr. inf. (bits)	PBF	Discr. Inf. (bits)	
<b>All training data</b>					
Training → LEP+XENON100	14.7(4)	3.88(4)	18.6(6)	4.22(4)	Strong
" " → ATLAS-sparticle	2.04(5)	1.03(4)	1.97(6)	0.98(5)	Barely worth mentioning
" " → ATLAS-Higgs	6.1(2)	2.61(4)	5.4(2)	2.43(5)	Substantial
Training → All	185(5)	7.53(4)	197(6)	7.62(5)	Decisive
<b><math>(g-2)_\mu</math> excluded from training data</b> (applicable only for $M_0, M_{1/2} < 2$ TeV)					
Training → LEP+XENON100	2.72(6)	1.45(3)	2.15(6)	1.11(4)	Barely worth mentioning
" " → ATLAS-sparticle	0.72(2)*	-0.48(4)*	1.81(6)	0.86(5)	Barely worth mentioning
" " → ATLAS-Higgs	4.2(2)	2.09(4)	6.7(2)	2.74(5)	Barely worth mentioning
Training → All	8.3(1)	3.05(4)	26.1(8)	4.71(5)	Substantial - Strong <sup>†</sup>

\* This apparent slight preference back towards the CMSSM is an artefact of the reweighting process used to obtain these results from the primary scans. See ref. [1] for details.

† Robustness to change in prior is compromised by the removal of  $(g-2)_\mu$  from the training set; the results we obtain span the two listed categories of the Jeffreys scale.

**Table 1:** Summary and interpretation of our results. Column 1 indicates the “training” and “inference” data used to compute the partial Bayes factors (PBFs) in the adjacent columns (where a  $PBF > 1$  indicates that the inference data provides evidence in favour of the SM-like hypothesis); ‘Training’ indicates that the priors were trained using only the “baseline” training data described in section 3, while “ ” indicates that training was performed using all the data from the row above. ‘LEP+XENON100’, ‘ATLAS-sparticle’ and ‘ATLAS-Higgs’ indicate that the update data was the LEP2 Higgs and 2011 XENON100 dark matter search data [7], the ATLAS  $1 \text{ fb}^{-1}$  SUSY search data [8], and ATLAS  $1 \text{ fb}^{-1}$  Higgs search data [9] respectively. The ‘Discr. inf.’ columns contain the discrimination information provided by the update data (simply the base 2 logarithm of the PBF) in favour of the SM-like hypothesis (the KL divergence  $KL(P(d_2|d_1, SM) || P(d_2|d_1, CMSSM))$  being the expected value of this quantity under  $P(d_2|d_1, SM)$ ). The two pairs of PBF and Discr. inf. columns indicate the results obtained using ‘log’ and ‘natural’ priors. The final column gives an interpretation of the strength of the evidence provided by the inference data, according to the Jeffreys scale.

dark sector, which are assumed to be unaffected by any data in our inference set. The relevant 1D parameter space  $m_h$  is given an initially log prior, which becomes roughly Gaussian (peaked near 90 GeV) after training with electroweak precision data [14]. SM marginal likelihoods are computed by directly applying the inference data likelihoods to this function (using standard numerical integration tools), and are then combined with the CMSSM marginal likelihoods to obtain partial Bayes factors, which are presented in table 1.

## 5. Discussion and conclusions

Our results provide a full probabilistic justification for the current intuition in the community that if the CMSSM is a good approximation to TeV scale physics, then it is extremely surprising that no direct evidence for it has yet been observed. In addition, our computed partial Bayes factors demonstrate that the parameter space priors that enable the above conclusion have the further, and unavoidable, implication that it is now much less probable that the CMSSM will be discovered to well approximate Nature than it was before the LEP2 Higgs search results were obtained – in the

sense that the odds of this occurring vs the SM remaining valid (with dark matter and  $(g-2)_\mu$  unexplained) are a factor of approximately 200 less with our ‘inference’ data considered than when only the ‘training’ data is considered. This conclusion cannot be avoided simply by altering the priors used because strong tensions exist even at the likelihood level, particularly between  $(g-2)_\mu$  and the ATLAS Higgs search likelihood.

The only escape available is to abandon  $(g-2)_\mu$  as a constraint; even dropping the assumption that neutralinos fully account for the observed dark matter relic density does not sufficiently open up the parameter space to avoid strong conflict between  $(g-2)_\mu$  and  $m_h$ . However, our results show that the  $M_0, M_{1/2} < 2$  TeV region of the CMSSM is still disfavoured even if  $(g-2)_\mu$  is abandoned.

Finally, we also highlight that our study was performed using February 2012 data for the Higgs mass constraints and that since this time these constraints have become much stronger, such that updating them would significantly strengthen our results.

## References

- [1] C. Balazs, A. Buckley, D. Carter, B. Farmer, and M. White. [arXiv:1205.1568v3 \[hep-ph\]](#).
- [2] O. Buchmueller, R. Cavanaugh, A. De Roeck, M. Dolan, J. Ellis, *et al.* [arXiv:1110.3568 \[hep-ph\]](#).  
O. Buchmueller, R. Cavanaugh, A. De Roeck, M. Dolan, J. Ellis, *et al.* [arXiv:1112.3564 \[hep-ph\]](#).
- [3] G. D. Starkman, R. Trotta, and P. M. Vaudrevange. [arXiv:0811.2415 \[physics.data-an\]](#). M. E. Cabrera, J. Casas, V. A. Mitsou, R. Ruiz de Austri, and J. Terron. [arXiv:1109.3759 \[hep-ph\]](#).  
S. AbdusSalam, B. Allanach, H. Dreiner, J. Ellis, U. Ellwanger, *et al.* *Eur.Phys.J.* **C71** (2011) 1835,  
[arXiv:1109.3859 \[hep-ph\]](#). S. Sekmen, S. Kraml, J. Lykken, F. Moortgat, S. Padhi, *et al.*  
[arXiv:1109.5119 \[hep-ph\]](#). C. Stenge, G. Bertone, D. Cerdeno, M. Fornasa, R. de Austri, *et al.*  
[arXiv:1112.4192 \[hep-ph\]](#). L. Roszkowski, E. M. Sessolo, and Y.-L. S. Tsai. [arXiv:1202.1503 \[hep-ph\]](#).
- [4] E. Jaynes. *Systems Science and Cybernetics*, IEEE Transactions on **4** no. 3, (1968) 227–241.
- [5] J. Berger, J. Bernardo, and D. Sun. *The Annals of Statistics* **37** no. 2, (2009) 905–938.
- [6] A. O’Hagan. *J. Roy. Statist. Soc. Ser. B* **57** no. 1, (1995) pp. 99–138. J. Berger and L. Pericchi. *JASA* **91** no. 433, (1996) 109–122. J. Berger and J. Mortera. *JASA* **94** no. 446, (1999) 542–554.
- [7] E. Aprile *et al.* *Phys.Rev.Lett.* **107** (2011) 131302, [arXiv:1104.2549 \[astro-ph.CO\]](#).
- [8] ATLAS Collaboration. *Phys.Lett.* **B710** (2012) 67–85, [arXiv:1109.6572 \[hep-ex\]](#).
- [9] ATLAS Collaboration. *Phys.Lett.* **B710** (2012) 49–66, [arXiv:1202.1408 \[hep-ex\]](#). ATLAS Collaboration. *Phys.Rev.Lett.* **108** (2012) 111803, [arXiv:1202.1414 \[hep-ex\]](#). ATLAS Collaboration. *Phys.Rev.Lett.* **108** (2012) 111802, [arXiv:1112.2577 \[hep-ex\]](#). ATLAS Collaboration. *Phys.Lett.* **B710** (2012) 383–402, [arXiv:1202.1415 \[hep-ex\]](#).
- [10] S. S. AbdusSalam, B. C. Allanach, M. J. Dolan, F. Feroz, and M. P. Hobson. *Phys. Rev.* **D80** (2009) 035017, [arXiv:0906.0957 \[hep-ph\]](#). G. Bertone, D. G. Cerdeno, M. Fornasa, R. Ruiz de Austri, C. Stenge, *et al.* *JCAP* **1201** (2012) 015, [arXiv:1107.1715 \[hep-ph\]](#). F. Feroz, K. Cranmer, M. Hobson, R. Ruiz de Austri, and R. Trotta. *JHEP* **06** (2011) 042, [arXiv:1101.3296 \[hep-ph\]](#). B. Allanach. *Phys.Rev.* **D83** (2011) 095019, [arXiv:1102.3149 \[hep-ph\]](#).
- [11] B. C. Allanach, K. Cranmer, C. G. Lester, and A. M. Weber. *JHEP* **0708** (2007) 023, [arXiv:0705.0487 \[hep-ph\]](#). M. E. Cabrera, J. A. Casas, and R. Ruiz de Austri. *JHEP* **03** (2009) 075, [arXiv:0812.0536 \[hep-ph\]](#).
- [12] L. Roszkowski, R. Ruiz de Austri, and R. Trotta. *Phys. Rev.* **D82** (2010) 055003, [arXiv:0907.0594 \[hep-ph\]](#).
- [13] F. Feroz, B. C. Allanach, M. Hobson, S. S. AbdusSalam, R. Trotta, *et al.* *JHEP* **0810** (2008) 064, [arXiv:0807.4512 \[hep-ph\]](#).
- [14] ALEPH, CDF, D0, DELPHI, L3, OPAL and SLD Collaborations, and the LEP E.W., Tevatron E.W. and SLD E.W. and H.F. groups. [arXiv:1012.2367 \[hep-ex\]](#).