Quantified naturalness from Bayesian statistics

Sylvain Fichet

International Institute of Physics, Natal

20/09/12

Presentation based on arXiv:1204.4940, to appear in PRD

Overview

1. Introduction
2. Bayesian model comparison
3. Naturalness from Bayesian statistics
4. Various implications
5. Combined dark matter and electroweak fine-tuning
6. Summary

Introduction

What are naturalness and fine-tuning ? (loosely speaking)

- Notions caracterizing a model in terms of its propensity to reproduce experimental observations.
- When employed, modify our degree of belief about model(s) under consideration.

What are naturalness and fine-tuning ? (loosely speaking)

- Notions caracterizing a model in terms of its propensity to reproduce experimental observations.
- When employed, modify our degree of belief about model(s) under consideration.
- Typical example: some parameters of a model need to be tuned very precisely to satisfy an experimental constraint. The model is fine-tuned, i.e. not natural. This typically decreases the degree of belief in the model.
- (typical examples: gauge-hierarchy problem, cosmological constant problem)

But the previous considerations are very rough…

- They are not quantified.
- They are totally subjective !

But the previous considerations are very rough…

- They are not quantified.
- They are totally subjective !

So to extract some objective and useful information one would need

- A consistent measure
- A rule telling how to interpret it (i.e. relating our degree of belief to this measure)

Some propositions exist …

- First measure of fine-tuning in the gauge-hierarchy problem (SUSY context) : sensitivity of the observable with respect to parameters

$$c = \max_i \left| \frac{\partial \log m_Z^2}{\partial \log \theta_i} \right|_{m_Z = m_{Z\ ex}}$$

Ellis Enqvist Nanopoulos Zwirner ’86, Barbieri Giudice ‘88

- For variations and alternative approaches, see

  Anderson Castano ’94, Ciafaloni Strumia ’97, Chan Chattopadhyay Nath ’98, Barbieri Strumia ’98, Giusti Romanino Strumia ’98, Athron Miller ’07

Some propositions exist …

- First measure of fine-tuning in the gauge-hierarchy problem (SUSY context) : sensitivity of the observable with respect to parameters

$$c = \max_i \left| \frac{\partial \log m_Z^2}{\partial \log \theta_i} \right|_{m_Z = m_{Z\ ex}}$$

Ellis Enqvist Nanopoulos Zwirner ’86, Barbieri Giudice ‘88

- For variations and alternative approaches, see
Anderson Castano ’94, Ciafaloni Strumia ’97, Chan Chattopadhyay Nath ’98, Barbieri Strumia ’98, Giusti Romanino Strumia ’98, Athron Miller ’07

But have conceptual flaws

- Arbitrary functional form giving inequivalent results, ill-defined concepts, ...

- The worst being the interpretation: link between degree of belief and the c measure.

Bayesian model comparison

- Definition of probability: degree of belief about a proposition.

- Definition of probability: degree of belief about a proposition.

- Apply Bayes' law to hypothesis H and data d :

$$p(H|d,I) = p(H|I)\frac{p(d|H,I)}{p(d,I)}$$

- Definition of probability: degree of belief about a proposition.

- Apply Bayes' law to hypothesis H and data d :

$$\underbrace{p(H|d,I)}_{\text{posterior probability}} = \underbrace{p(H|I)}_{\text{prior probability}} \underbrace{\frac{p(d|H,I)}{p(d,I)}}_{\text{likelihood}}$$

- Definition of probability: degree of belief about a proposition.

- Apply Bayes' law to hypothesis H and data d :

$$\underbrace{p(H|d,I)}_{\text{posterior probability}} = \underbrace{p(H|I)}_{\text{prior probability}} \frac{\overbrace{p(d|H,I)}^{\text{likelihood}}}{p(d,I)}$$

- Model comparison :

$$\frac{p(\mathcal{M}_0|d)}{p(\mathcal{M}_1|d)} = \frac{p(\mathcal{M}_0)}{p(\mathcal{M}_1)} \frac{p(d|\mathcal{M}_0)}{p(d|\mathcal{M}_1)}$$

Bayes factor B_{01}

with $p(d|\mathcal{M}) = \int_{\mathcal{D}} d\theta p(d|\theta,\mathcal{M})p(\theta|\mathcal{M})$

likelihood

internal prior

- Definition of probability: degree of belief about a proposition.

- Apply Bayes' law to hypothesis H and data d :

posterior probability →
$$p(H|d,I) = p(H|I)\frac{p(d|H,I)}{p(d,I)}$$
← likelihood

↑ prior probability

- Model comparison :
$$\frac{p(\mathcal{M}_0|d)}{p(\mathcal{M}_1|d)} = \frac{p(\mathcal{M}_0)}{p(\mathcal{M}_1)}\frac{p(d|\mathcal{M}_0)}{p(d|\mathcal{M}_1)}$$
← Bayes factor B_{01}

with
$$p(d|\mathcal{M}) = \int_{\mathcal{D}} d\theta p(d|\theta,\mathcal{M})p(\theta|\mathcal{M})$$
likelihood, internal prior

- Jeffreys' scale

$\lvert\log B_{01}\rvert$	Odds	Strength of evidence
< 1.0	$\lesssim 3:1$	Inconclusive
1.0	$\sim 3:1$	Weak evidence
2.5	$\sim 12:1$	Moderate evidence
5.0	$\sim 150:1$	Strong evidence

Naturalness from Bayesian Statistics

- Consider a model $\mathcal{M}$ with n parameters $\theta = (\theta_1, \ldots, \theta_n)$ spanning the parameter space $\mathcal{D}$, and m ``observables'' $\mathcal{O}(\theta) = (\mathcal{O}_1(\theta), \ldots, \mathcal{O}_m(\theta))$ among the avalaible data.

- Consider a measurement $\mathcal{O}_{ex}$ with uncertainty Σ, such that $\mathcal{O} = \mathcal{O}_{ex}$ is satisfied over the subspace $\mathcal{D}_{ex}$ of dimension $n - m$. Other data are called d .

- Usual definition of naturalness:
  « Sensitivity of $\mathcal{O}$ around a point θ_{ex} belonging to $\mathcal{D}_{ex}$ »

⇨ $c = \max \left| \frac{\partial \log \mathcal{O}}{\partial \log \theta_i} \right|_{\theta=\theta_{ex}}$ or $c = \sqrt{\left(\frac{\partial \log \mathcal{O}}{\partial \log \theta_i}\right)^2}\Bigg|_{\theta=\theta_{ex}}$ or…

- Another definition for naturalness:
  « Probability of having $\mathcal{O} = \mathcal{O}_{ex}$ in the model $\mathcal{M}$ »

$$\Longrightarrow \quad p(\mathcal{O} = \mathcal{O}_{ex} | \mathcal{M}, d)$$

- Not normalized as a function of the hypothesis, but one can build a well-defined Bayes factor.

$$\Longrightarrow \quad \boxed{B = \frac{p(\mathcal{O} = \mathcal{O}_{ex} | \mathcal{M}, d)}{p(\mathcal{O} = \mathcal{O}_{ex} | \mathcal{M}', d')}}$$

- This is a general measure of naturalness. Many possibilities, depending on what is chosen for $\mathcal{M}, \mathcal{M}', d, d'$.

- OK for relative naturalness. For absolute naturalness, one needs to define $\mathcal{M}'$ as an absolute reference.

- Assume that the uncertainty is sufficiently small, such that one can take Laplace's approximation (i.e approximate likelihood to a multivariate normal distribution).

$$\mathcal{L}(\theta) = \mathcal{L}_{max} \exp\left(\frac{1}{2} \frac{\partial^2 \mathcal{L}}{\partial \log \mathcal{O}_i \partial \log \mathcal{O}_j} J_{\mathcal{O}\,ik} J_{\mathcal{O}\,jl} (\theta - \theta_{ex})_k (\theta - \theta_{ex})_l \right)$$

Covariance matrix (relative uncertainty) $\equiv \Sigma_{ij}^{-1}$

Jacobians

- The probability $p(\mathcal{O} = \mathcal{O}_{ex} | \mathcal{M}, d)$ takes then the form

uncertainty

Jacobian factor

$$p(\mathcal{O} = \mathcal{O}_{ex} | \mathcal{M}, d) = \mathcal{L}_{max} \frac{|\Sigma|^{1/2}}{|V|^{1/2}} \int_{\mathcal{D}_{ex}} \frac{1}{C} d\sigma(\theta)$$

prior volume

integration measure on $\mathcal{D}_{ex}$

- $$p(\mathcal{O}=\mathcal{O}_{ex}|\mathcal{M},d)=\mathcal{L}_{max}\frac{|\Sigma|^{1/2}}{|V|^{1/2}}\int_{\mathcal{D}_{ex}}\frac{1}{C}d\sigma(\theta)$$

  with $C=\left|\det\left(J_{\log\mathcal{O}}J^{t}_{\log\mathcal{O}}\right)\right|^{1/2}$

  ⇒ Generalized form of the sensitivity measure

- Tells how much information is contained in $\mathcal{O}=\mathcal{O}_{ex}$ regardless of the uncertainty.

- $$p(\mathcal{O}=\mathcal{O}_{ex}|\mathcal{M},d)=\mathcal{L}_{max}\frac{|\Sigma|^{1/2}}{|V|^{1/2}}\int_{\mathcal{D}_{ex}}\frac{1}{C}d\sigma(\theta)$$

  with $C=\left|\det\left(J_{\log\mathcal{O}}J^t_{\log\mathcal{O}}\right)\right|^{1/2}$

  ⇒ Generalized form of the sensitivity measure

- Tells how much information is contained in $\mathcal{O}=\mathcal{O}_{ex}$ regardless of the uncertainty.

- C^{-1} reduces to the intuitive sensitivity c^{-1} for
  - A single source of fine-tuning
  - Logarithmic priors
  - And punctual priors i.e selecting a point θ_{ex} of $\mathcal{D}_{ex}$,

such that $\int_{\mathcal{D}_{ex}}C^{-1}d\sigma(\theta)\rightarrow C^{-1}\big|_{\theta=\theta_{ex}}$

- $$p(\mathcal{O}=\mathcal{O}_{ex}|\mathcal{M},d)=\mathcal{L}_{max}\frac{|\Sigma|^{1/2}}{|V|^{1/2}}\int_{\mathcal{D}_{ex}}\frac{1}{C}d\sigma(\theta)$$

  with $C=\left|\det\left(J_{\log\mathcal{O}}J^t_{\log\mathcal{O}}\right)\right|^{1/2}$

  ⇒ Generalized form of the sensitivity measure

- Tells how much information is contained in $\mathcal{O}=\mathcal{O}_{ex}$ regardless of the uncertainty.

- C^{-1} reduces to the intuitive sensitivity c^{-1} for
  - A single source of fine-tuning
  - Logarithmic priors
  - And punctual priors i.e selecting a point θ_{ex} of $\mathcal{D}_{ex}$,

such that $\int_{\mathcal{D}_{ex}} C^{-1}d\sigma(\theta)\rightarrow C^{-1}\big|_{\theta=\theta_{ex}}$

- ! Still impossible to interpret if not embedded in the probability framework

Back to the naturalness measure...

- Consider a measure of relative naturalness, $B_{01} = \frac{p(\mathcal{O} = \mathcal{O}_{ex}|\mathcal{M}_0, d_0)}{p(\mathcal{O} = \mathcal{O}_{ex}|\mathcal{M}_1, d_1)}$
  With what we learnt before, it becomes

$$B_{01} = \underbrace{\frac{|V_1|^{1/2}}{|V_0|^{1/2}}}_{\text{ratio of prior volumes}} \underbrace{\int_{\mathcal{D}_{ex}\,0} C_0^{-1} d\sigma(\theta) \left(\int_{\mathcal{D}_{ex}\,1} C_1^{-1} d\sigma(\theta) \right)^{-1}}_{\text{ratio of sensitivities}}$$

Back to the naturalness measure...

- Consider a measure of relative naturalness, $B_{01} = \frac{p(\mathcal{O} = \mathcal{O}_{ex}|\mathcal{M}_0, d_0)}{p(\mathcal{O} = \mathcal{O}_{ex}|\mathcal{M}_1, d_1)}$
  With what we learnt before, it becomes

$$B_{01} = \underbrace{\frac{|V_1|^{1/2}}{|V_0|^{1/2}}}_{\text{ratio of prior volumes}} \underbrace{\int_{\mathcal{D}_{ex}\,0} C_0^{-1} d\sigma(\theta) \left(\int_{\mathcal{D}_{ex}\,1} C_1^{-1} d\sigma(\theta) \right)^{-1}}_{\text{ratio of sensitivities}}$$

- Comparing two points θ_0, θ_1 of the same model, the measure reduces to $B_{01} = \frac{C_1}{C_0}$

⇨ Interpretation of sensitivity is calibrated by Jeffreys' scale

- $B_{01} \approx 3, 12, 150$ corresponds to thresholds of weak, moderate, strong fine-tuning of θ_1 with respect to θ_0 .

How to define $\mathcal{M}'$ as a reference in $B = \dfrac{p(\mathcal{O} = \mathcal{O}_{ex}|\mathcal{M})}{p(\mathcal{O} = \mathcal{O}_{ex}|\mathcal{M}')}$?

Fully natural — Infinitely fine-tuned

$C = 0$ — $C \to \infty$

$\mathcal{X}$

How to define $\mathcal{M}'$ as a reference in $B = \dfrac{p(\mathcal{O} = \mathcal{O}_{ex}|\mathcal{M})}{p(\mathcal{O} = \mathcal{O}_{ex}|\mathcal{M}')}$?

Fully natural — "puzzle" — Infinitely fine-tuned

$C = 0$ — $C = 1$ — $C \to \infty$

$\mathcal{X}$ — $\mathcal{P}$

How to define $\mathcal{M}'$ as a reference in $B = \dfrac{p(\mathcal{O} = \mathcal{O}_{ex}|\mathcal{M})}{p(\mathcal{O} = \mathcal{O}_{ex}|\mathcal{M}')}$?

Fully natural	"puzzle"	Infinitely fine-tuned
$C = 0$	$C = 1$	$C \to \infty$
$\mathcal{X}$	$\mathcal{P}$	

- Comparison to the ideal model $\mathcal{X}$ such that $p(\mathcal{O} = \mathcal{O}_{ex}|\mathcal{X}) = \mathcal{L}_{max}$ is not very interesting.

How to define $\mathcal{M}'$ as a reference in $B = \dfrac{p(\mathcal{O} = \mathcal{O}_{ex}|\mathcal{M})}{p(\mathcal{O} = \mathcal{O}_{ex}|\mathcal{M}')}$?

Fully natural — "puzzle" — Infinitely fine-tuned

$C = 0$ — $C = 1$ — $C \to \infty$

$\mathcal{X}$ — $\mathcal{P}$

- Comparison to the ideal model $\mathcal{X}$ such that $p(\mathcal{O} = \mathcal{O}_{ex}|\mathcal{X}) = \mathcal{L}_{max}$ is not very interesting.

- Comparison to the model $\mathcal{P}$ in which $\mathcal{O}$ is an input parameter, such that $C_{\mathcal{P}} = 1$, is more interesting.

prior volume of the observable itself

$$B_{\mathcal{MP}} = \frac{p(\mathcal{O} = \mathcal{O}_{ex}|\mathcal{M})}{p(\mathcal{O} = \mathcal{O}_{ex}|\mathcal{P})} = \frac{|V_{\mathcal{O}}|^{1/2}}{|V|^{1/2}} \int_{\mathcal{D}_{ex}} \frac{1}{C} d\sigma(\theta)$$

What does specify $|V_{\mathcal{O}}|$?

Two principles leads to two different conditions leading to the same result.
(Here shown for dimensionful observables)

What does specify $|V_{\mathcal{O}}|$?

Two principles leads to two different conditions leading to the same result.
(Here shown for dimensionful observables)

- Indifference principle (setting the most objective prior).
  Invariance of $p(\mathcal{O}|\mathcal{P})$ under a change of unit scale, i.e $\mathcal{O} \to \mathcal{O} \times b$, implies $p(\mathcal{O}|\mathcal{P}) \propto \mathcal{O}^{-1}$.

- Consistency
  Asking that the measure $B_{\mathcal{MP}}$ is the same whatever the dimension of $\mathcal{O}$, i.e invariant under $\mathcal{O} \to \mathcal{O}^a$, implies $p(\mathcal{O}|\mathcal{P}) \propto \mathcal{O}^{-1}$ again.

What does specify $|V_{\mathcal{O}}|$?

Two principles leads to two different conditions leading to the same result.
(Here shown for dimensionful observables)

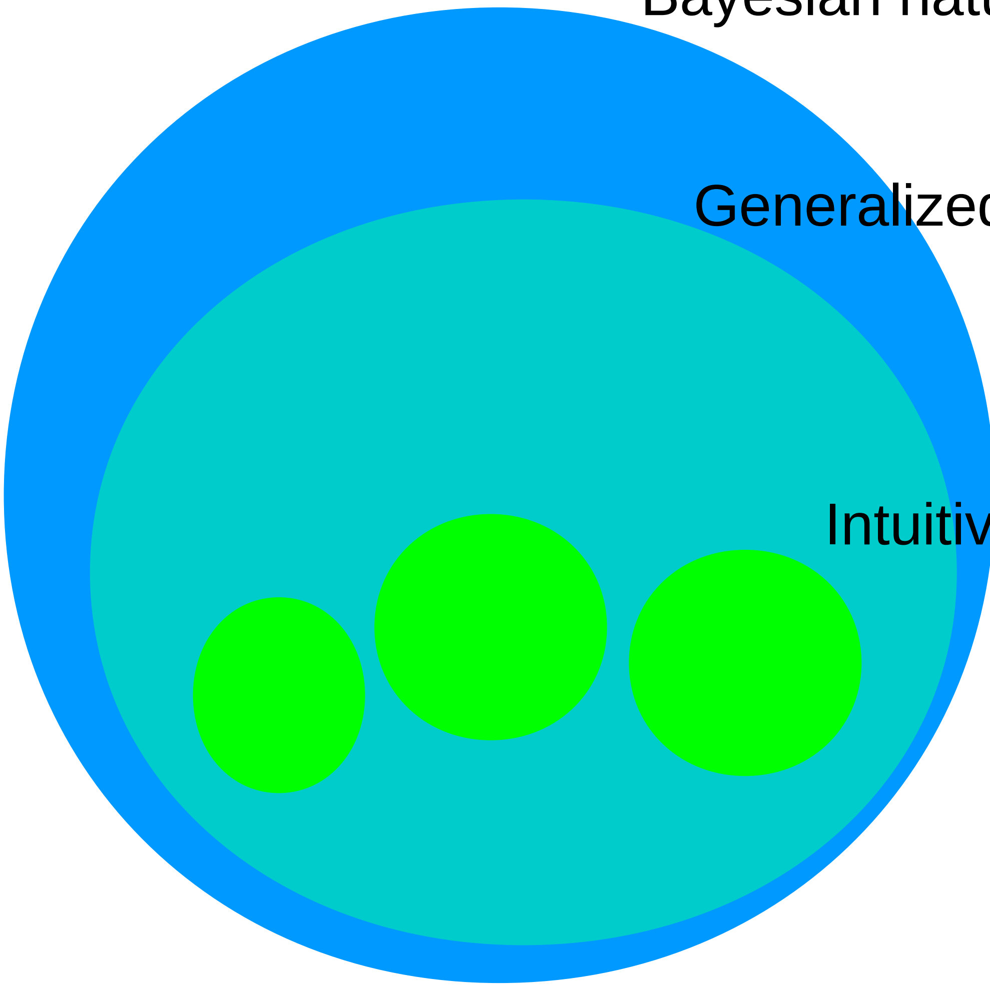
- Indifference principle (setting the most objective prior).
  Invariance of $p(\mathcal{O}|\mathcal{P})$ under a change of unit scale, i.e $\mathcal{O} \to \mathcal{O} \times b$, implies $p(\mathcal{O}|\mathcal{P}) \propto \mathcal{O}^{-1}$.

- Consistency
  Asking that the measure $B_{\mathcal{MP}}$ is the same whatever the dimension of $\mathcal{O}$, i.e invariant under $\mathcal{O} \to \mathcal{O}^a$, implies $p(\mathcal{O}|\mathcal{P}) \propto \mathcal{O}^{-1}$ again.

- Consequences :
  - $|V_{\mathcal{O}}| = \int d\log\mathcal{O}$
  - Measure invariant under $\mathcal{O} \to b \times \mathcal{O}^a$
  - $C = \partial \log \mathcal{O} / \partial \ldots$
  - The C measure is completely fixed with this approach

⇨ Conceptual problems are solved

In short...



$$\int_{\mathcal{D}_{ex}} C^{-1} d\sigma(\theta)$$

$$C = \left|\det\left(J_{\log\mathcal{O}} J^t_{\log\mathcal{O}}\right)\right|^{1/2}$$

Various implications

What happens for two sources of fine-tuning ?

- $$C = \left|\det(J_{\mathcal{O}} J^t_{\mathcal{O}})\right|^{1/2} = \|\nabla \log \mathcal{O}_1 \wedge \nabla \log \mathcal{O}_2\|$$

or $$C = \left(\|\nabla \log \mathcal{O}_1\|^2 \, \|\nabla \log \mathcal{O}_2\|^2 - (\nabla \log \mathcal{O}_1 . \nabla \log \mathcal{O}_2)^2\right)^{1/2}$$

or $C = C_1 C_2 \sqrt{1-\rho^2}$ or $$\rho = \frac{|\nabla \log \mathcal{O}_1 . \nabla \log \mathcal{O}_2|}{\|\nabla \log \mathcal{O}_1\| \, \|\nabla \log \mathcal{O}_2\|}$$

- C is maximal when $\rho = 0$ i.e. when the two observables are independently predicted. It decreases when observables are correlated in the model.

What happens for two sources of fine-tuning ?

- $$C = \left|\det(J_{\mathcal{O}} J_{\mathcal{O}}^t)\right|^{1/2} = \|\nabla \log \mathcal{O}_1 \wedge \nabla \log \mathcal{O}_2\|$$

  or $$C = \left(\|\nabla \log \mathcal{O}_1\|^2 \, \|\nabla \log \mathcal{O}_2\|^2 - (\nabla \log \mathcal{O}_1 . \nabla \log \mathcal{O}_2)^2\right)^{1/2}$$

  or $$C = C_1 C_2 \sqrt{1-\rho^2}$$ or $$\rho = \frac{|\nabla \log \mathcal{O}_1 . \nabla \log \mathcal{O}_2|}{\|\nabla \log \mathcal{O}_1\| \, \|\nabla \log \mathcal{O}_2\|}$$

- C is maximal when $\rho = 0$ i.e. when the two observables are independently predicted. It decreases when observables are correlated in the model.

- Formula becomes invalid (i.e observables not separately informative) when ρ is no longer small with respect to the experimental correlation coefficient. C is then reduced to a one-dimensional measure associated to $\mathcal{O} \sim\propto \mathcal{O}_{1,2}$.

Shall the top Yukawa appear in the EW fine-tuning measure ?

i.e does $y_t \in \{p_i\}$ in $C_{m_Z} = \left(\sum_i \left(\frac{\partial \log m_Z}{\partial \log p_i} \right)^2 \right)^{1/2}$?

- Argument 1 : y_{top} is a parameter of the model. Answer is yes.

  Argument 2 : It is fixed by $m_{top} = m_{top\,ex}$, so it is like a constant, not a free parameter. Answer is no ...

Shall the top Yukawa appear in the EW fine-tuning measure ?

i.e does $y_t \in \{p_i\}$ in $C_{m_Z} = \left(\sum_i \left(\frac{\partial \log m_Z}{\partial \log p_i} \right)^2 \right)^{1/2}$?

- Argument 1 : y_{top} is a parameter of the model. Answer is yes.
  Argument 2 : It is fixed by $m_{top} = m_{top\,ex}$, so it is like a constant, not a free parameter. Answer is no ...

- Solution : the problem is in argument 1. Letting y_{top} be free, the constraint $m_{top} = m_{top\,ex}$ has to be added to the set of experimental constraints. As it is a priori not independent from the $m_Z = m_{Z\,ex}$ prediction, one should consider the combined measure $C_{m_Z, m_{top}}$ to measure naturalness.

- The computation shows that $C_{m_Z, m_{top}}$ equals C_{m_Z} without the y_{top} term.
  $\Rightarrow$ Answer is actually no for argument 1.

Various implications

- Measure associated to two sources of fine-tuning ? $C = \|\nabla \log \mathcal{O}_1 \wedge \nabla \log \mathcal{O}_2\|$
- Shall the Yukawa couplings be taken into account in the measure of EW fine-tuning ? No.
- When including a "naturalness prior" in studies of Bayesian inference, prior of parameters should be consistent with the C measure used.
- If one needs to fine-tune parameters to select a region of parameter space with low fine-tuning, a "second-order fine-tuning" appears.
- ...

Dark matter and electroweak fine-tuning

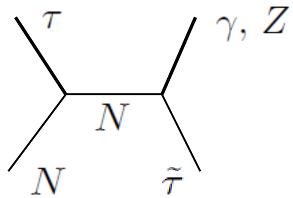
- Supersymmetry (SUSY) broken at low energy improves the gauge-hierarchy problem and can provide stable neutral particle explaining dark matter.

- cMSSM: a classic supersymmetric model, well studied (though not well motivated). Contains the neutralino, a good dark matter candidate.

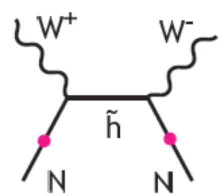
$$p_i = (m_{1/2}, m_0, a_0, \mu, B_\mu)$$

gaugino masses — scalar parameters — Higgs parameters

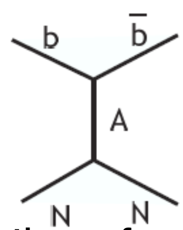
- Typically too much dark matter compared to $\Omega h^2_{ex} = 0.1126 \pm 0.0036$ (WMAP7)
  Requires enhanced annihilation mechanisms $\Rightarrow$ fine-tuning



cohannilihation region



``focus point'' region

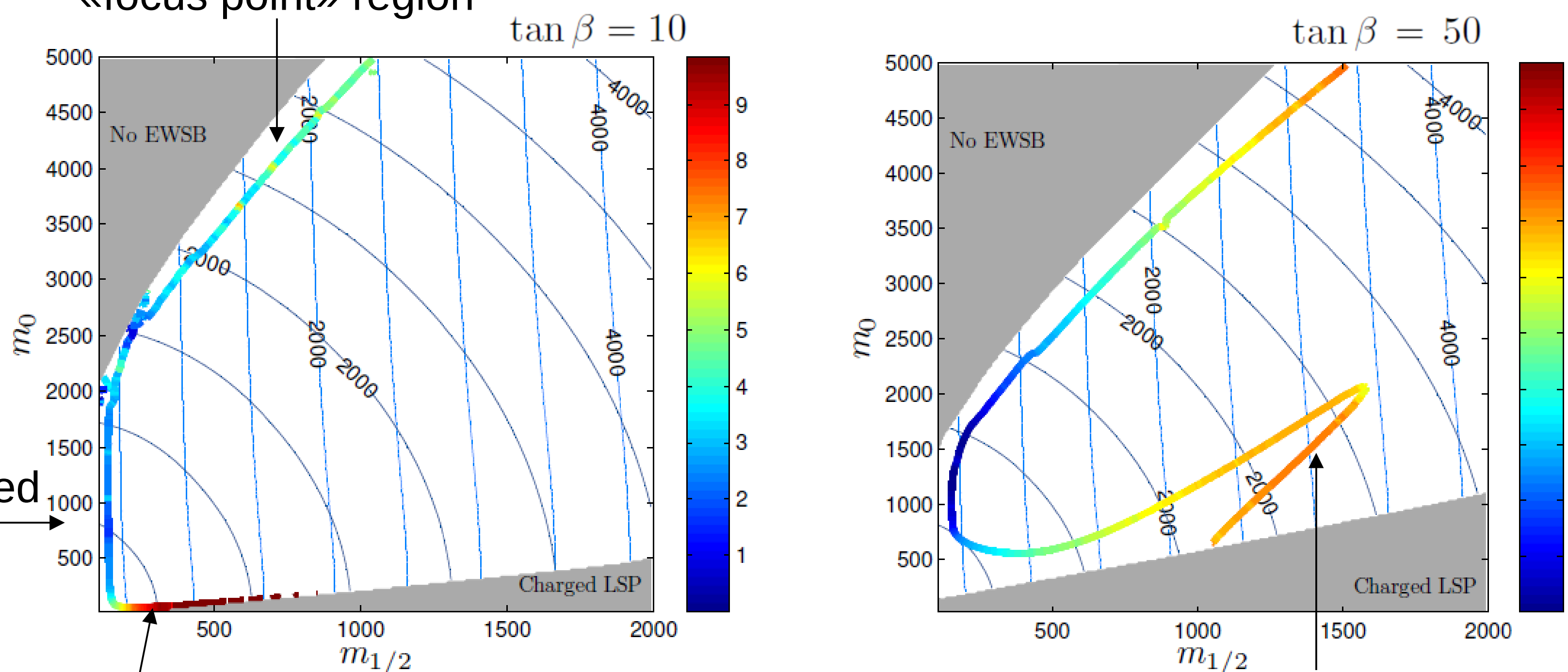


Higgs funnel

- **Two** sources of fine-tuning $\begin{cases} \Omega = \Omega_{ex} \\ m_Z = m_{Z\,ex} \end{cases} \Rightarrow$

$$\boxed{C_{m_Z,\Omega} = \left\| \frac{\partial \log m_Z}{\partial \log p_i} \wedge \frac{\partial \log \Omega h^2}{\partial \log p_i} \right\|}$$

«focus point» region



LHC excluded

cohannilihation region

Higgs funnel

- $\Delta \log C = 1,\ 2.5,\ 5$ corresponds to weak, moderate, strong (relative) fine-tuning
- Focus point region strongly favored with respect to coannihilation region, Higgs funnel in between.

Summary

Summary

- Usual measures of naturalness have conceptual problems. We propose a consistent approach, based on Bayesian statistics.
- Our approach contains the usual sensitivity C in a generalized form: several observables, arbitrary priors, non-local measure...
- Interpretation of the C measure is under control, given by Jeffreys' scale.
- The top yukawa does not enter in the C_{m_Z} measure.
- We studied combined DM and EW fine-tuning in the cMSSM. Solid conclusions are made concerning relative naturalness of the different regions.

Thank you very much !!!!!!