SFitter: From LHC Data back to the MSSM Lagrangian

Michael Rauch

University of Edinburgh

Scottish Universities Physics Alliance

Outline

- Determining SUSY parameters
- Parameter Scans
- Weighted Markov Chains
- mSUGRA as a toy model

Determining SUSY parameters

nowadays:

Parameters in the Lagrangian

$M_{\{1,2,3\}}, \mu, \tan(\beta), m_0, \ldots$

Feynman diagrams, RG evolution, . . .

Observables:

- Masses
- Cross sections
- Branching ratios
- . . .

after SUSY discovery:

Observables

$m_{h^0}, \Delta m_{\tilde{g}\chi_1^0}$, BR, . . .

?

Lagrangian parameters

On loop-level observables depend on every parameter

Simple inversion of the relations not possible

$\Rightarrow$ Parameter scans

Error estimates on parameters in the minimum

Parameter Scans

MSSM parameter space is high-dimensional:

- SM: 3+ parameters (m_t, α_s, α, $\ldots$)
- mSUGRA: 5 parameters (m_0,$m_{\frac{1}{2}}, A_0, \tan(\beta), \text{sgn}(\mu)$)
- General MSSM: 105 parameters

Parameter Scans

MSSM parameter space is high-dimensional:

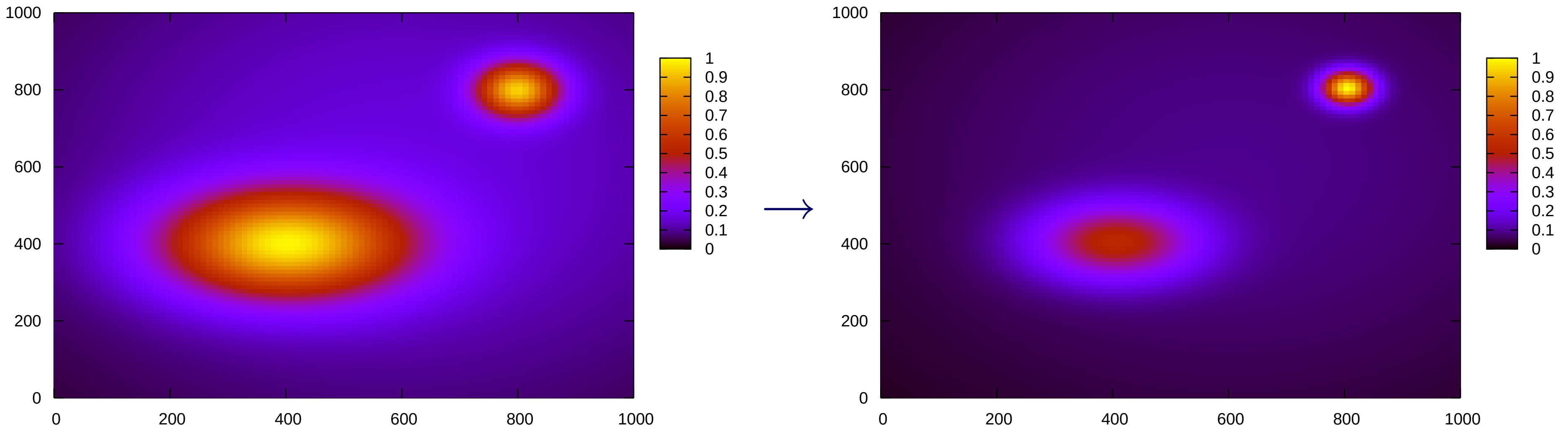
- SM: 3+ parameters (m_t, α_s, α, $\ldots$)
- mSUGRA: 5 parameters (m_0,$m_{\frac{1}{2}}, A_0, \tan(\beta), \mathrm{sgn}(\mu)$)
- General MSSM: 105 parameters

How to scan best?

Parameter Scans

How to scan best?

- Gradient search (Minuit)
  - Reasonably fast
  - Limited convergence radius
  - Can only find best fit



Parameter Scans

How to scan best?

- Gradient search (Minuit)
- Grid scan
  - Can scan complete parameter space
  - For high dimensions many points needed $\left(\text{(points per dimension)}^{\text{dimension}}\right)$

Parameter Scans

How to scan best?

- Gradient search (Minuit)
- Grid scan
- Simulated Annealing ($\longrightarrow$ Fittino)
- Markov Chains

Markov Chains

Markov Chain (MC):

- Sequence of points, chosen by an algorithm (Metropolis-Hastings), only depending on its direct predecessor
- Picks a set of "average" points according to a potential V (e.g. inverse log-likelihood, $1/\chi^2$)
- Point density resembles the value of V (i.e. more points in region with high V)
- Scans high dimensional parameter spaces efficiently
  [Baltz, Gondolo 2004]
- mSUGRA MC scans with current exp. limits
  [Allanach, Lester, Weber 2005/6; Roszkowski, Ruiz de Austra, Trotta 2006]

Improvements

Improved evaluation algorithm for binning: [Plehn, MR]

- Weight points with value of V
- Take care of
  - Overcounting because point density is already weighted ($\frac{\text{number of points}}{\sum_{\text{points}} 1/V(\text{point})}$) [based on Ferrenberg, Swendsen 1988]
  - Correct account for regions with zero probability (maintain additional chain which stores points rejected because $V(\text{point}) = 0$)

Maximum finder:

- Finds best match in N-D parameter space
- Refinement of MC match by additional gradient search
- Tries to find second-, third-, . . . -best distinct match

mSUGRA as a Toy Model

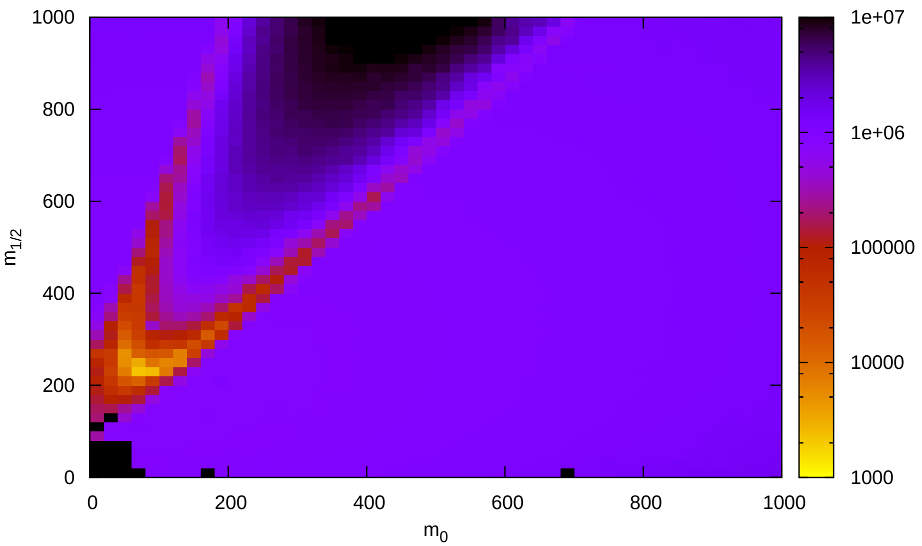
mSUGRA with LHC measurements (SPS1a kinematic edges): [Lafaye, Plehn, MR, D Zerwas]

Free parameters:

m_0, $m_{1/2}$, $\tan(\beta)$, A_0, $\mathrm{sgn}(\mu)$, m_t

SFitter output 1:

Probability Map $\longrightarrow$



- $P(m|d)$ (read: probability of the model m given the data d)
- Bayes' theorem: $P(m|d) = P(d|m)P(m)/P(d)$
- log-likelihood: $P(d|m) \propto |\mathcal{M}|^2$
- Theorists' prejudice: $P(m)$ ($P(d)$ removed via normalization)

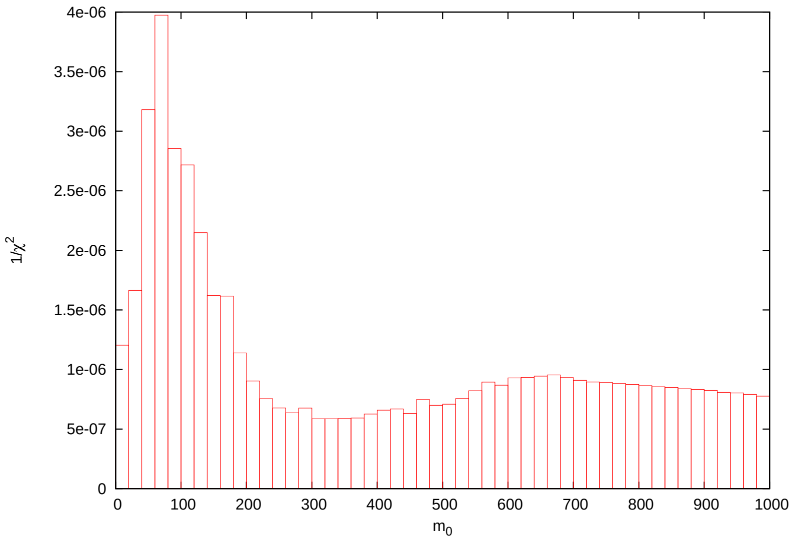
mSUGRA as a Toy Model

mSUGRA with LHC measurements (SPS1a kinematic edges): [Lafaye, Plehn, MR, D Zerwas]

Free parameters:

m_0, $m_{1/2}$, $\tan(\beta)$, A_0, $\mathrm{sgn}(\mu)$, m_t

SFitter output 1:
Probability Map $\longrightarrow$



- $P(m|d)$ (read: probability of the model m given the data d)
- Bayes' theorem: $P(m|d) = P(d|m)P(m)/P(d)$
- log-likelihood: $P(d|m) \propto |\mathcal{M}|^2$
- Theorists' prejudice: $P(m)$ ($P(d)$ removed via normalization)

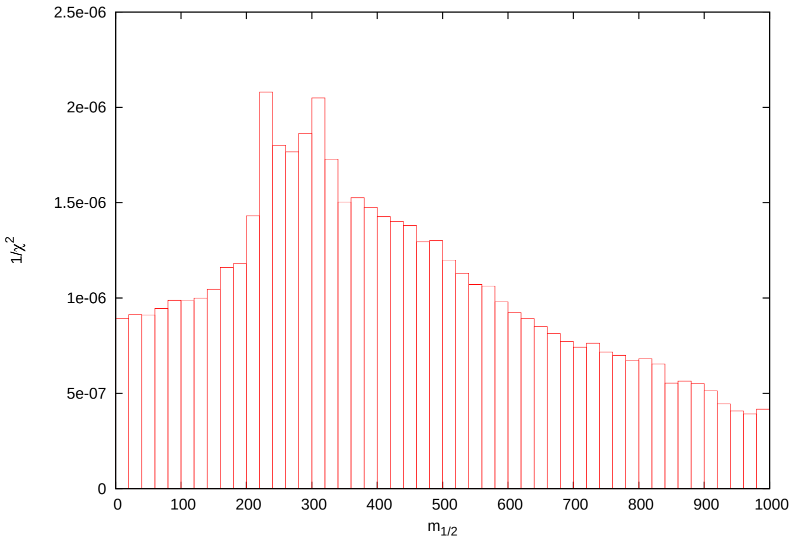
mSUGRA as a Toy Model

mSUGRA with LHC measurements (SPS1a kinematic edges): [Lafaye, Plehn, MR, D Zerwas]

Free parameters:

m_0, $m_{1/2}$, $\tan(\beta)$, A_0, $\text{sgn}(\mu)$, m_t

SFitter output 1:
Probability Map $\longrightarrow$



- $P(m|d)$ (read: probability of the model m given the data d)
- Bayes' theorem: $P(m|d) = P(d|m)P(m)/P(d)$
- log-likelihood: $P(d|m) \propto |\mathcal{M}|^2$
- Theorists' prejudice: $P(m)$ ($P(d)$ removed via normalization)

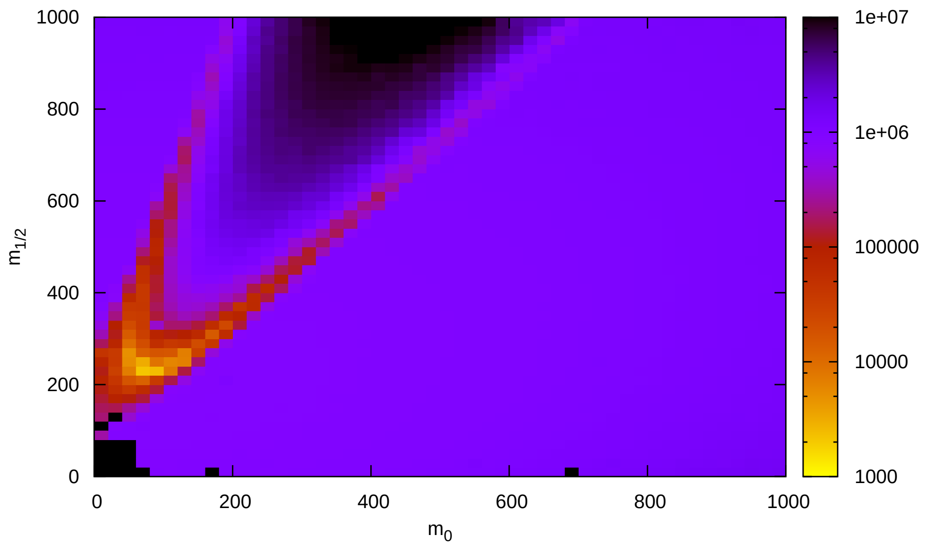
mSUGRA as a Toy Model

mSUGRA with LHC measurements (SPS1a kinematic edges): [Lafaye, Plehn, MR, D Zerwas]

Free parameters:

m_0, $m_{1/2}$, $\tan(\beta)$, A_0, $\mathrm{sgn}(\mu)$, m_t

SFitter output 1:
Probability Map $\longrightarrow$



SFitter output 2:
ranked list of minima:

	χ^2	m_0	$m_{1/2}$	$\tan(\beta)$	A_0	μ	m_t
1)	$3 \cdot 10^{-4}$	100.0	250.0	10.0	-99.9	+	171.4
2)	27.42	99.7	251.6	11.7	848.9	+	181.6
3)	54.12	107.2	243.4	13.3	-97.4	$-$	171.1
4)	70.99	108.5	246.9	13.9	26.4	$-$	173.6
5)	88.53	107.7	245.9	12.9	802.7	$-$	182.7

...

mSUGRA as a Toy Model

Ranked list of minima:

	χ^2	m_0	$m_{1/2}$	$\tan(\beta)$	A_0	μ	m_t
1)	$3 \cdot 10^{-4}$	100.0	250.0	10.0	−99.9	+	171.4
2)	27.42	99.7	251.6	11.7	848.9	+	181.6
3)	54.12	107.2	243.4	13.3	−97.4	−	171.1
4)	70.99	108.5	246.9	13.9	26.4	−	173.6
5)	88.53	107.7	245.9	12.9	802.7	−	182.7

Errors on the best-fit result:

	m_0	$m_{1/2}$	$\tan(\beta)$	A_0
ΔLHC(masses)	3.9	1.7	1.1	33
ΔLHC(edges)	1.2	1.0	0.9	20
ΔSLHC(edges)	0.7	0.6	0.7	10
ΔILC	0.09	0.13	0.12	4.8

Secondary Solutions:
small χ^2, but often strange characteristics:

- atypical errors (thin bands)
- atypical correlation matrix
- only one measurement contributes significantly to χ^2

$\Rightarrow$ Can reconstruct mSUGRA points at the LHC

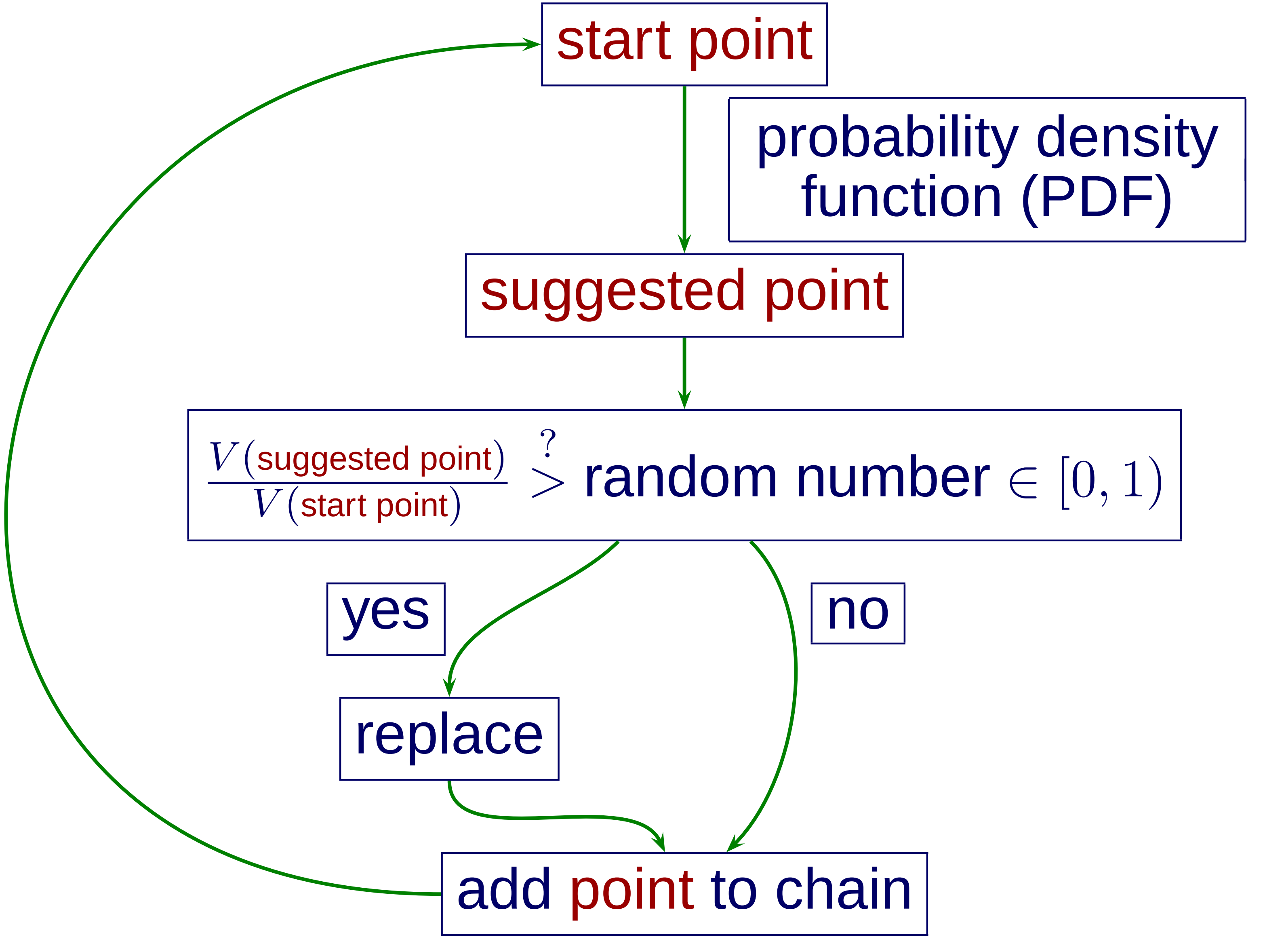
Summary & Outlook

- Parameter scans important to determine Lagrangian parameters from observables
- Problem of high-dimensional parameter spaces
- Markov Chains can do this effectively
- Improved algorithm developed
- Tested with mSUGRA SPS1a:
  can reconstruct SPS1a from (simulated) LHC data

- Repeat procedure with more general MSSM
  (20+ parameters)
- SFitter (despite its name) not tied to SUSY
  $\longrightarrow$ extend to other models/problems

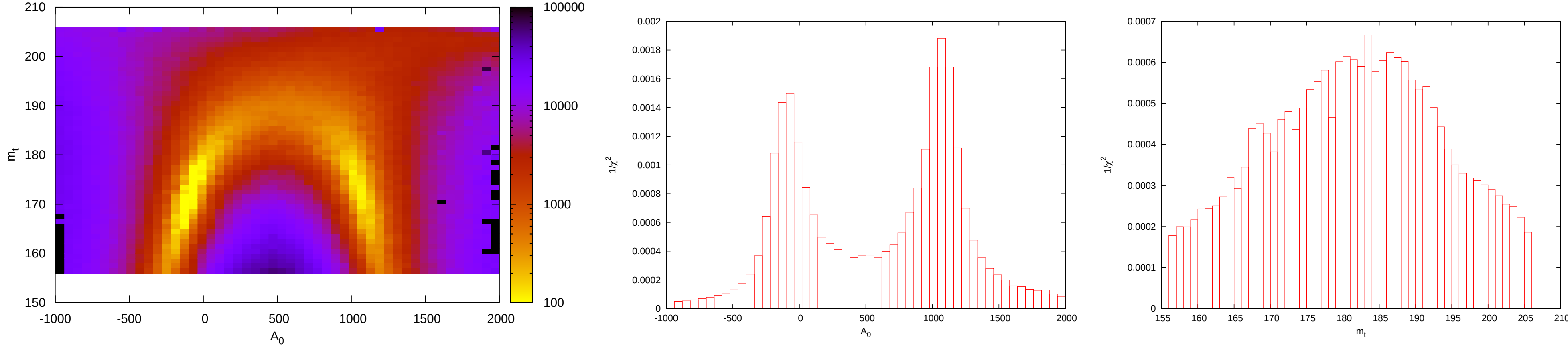
Backup Slides

Metropolis-Hastings Algorithm

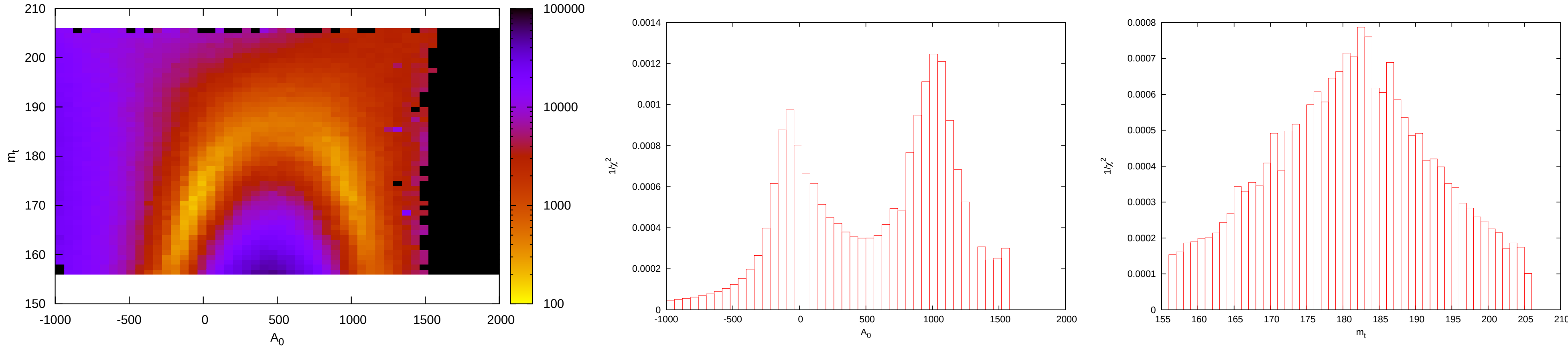


mSUGRA around Minima

$\mathrm{sgn}(\mu) = +1$ A_0 m_t



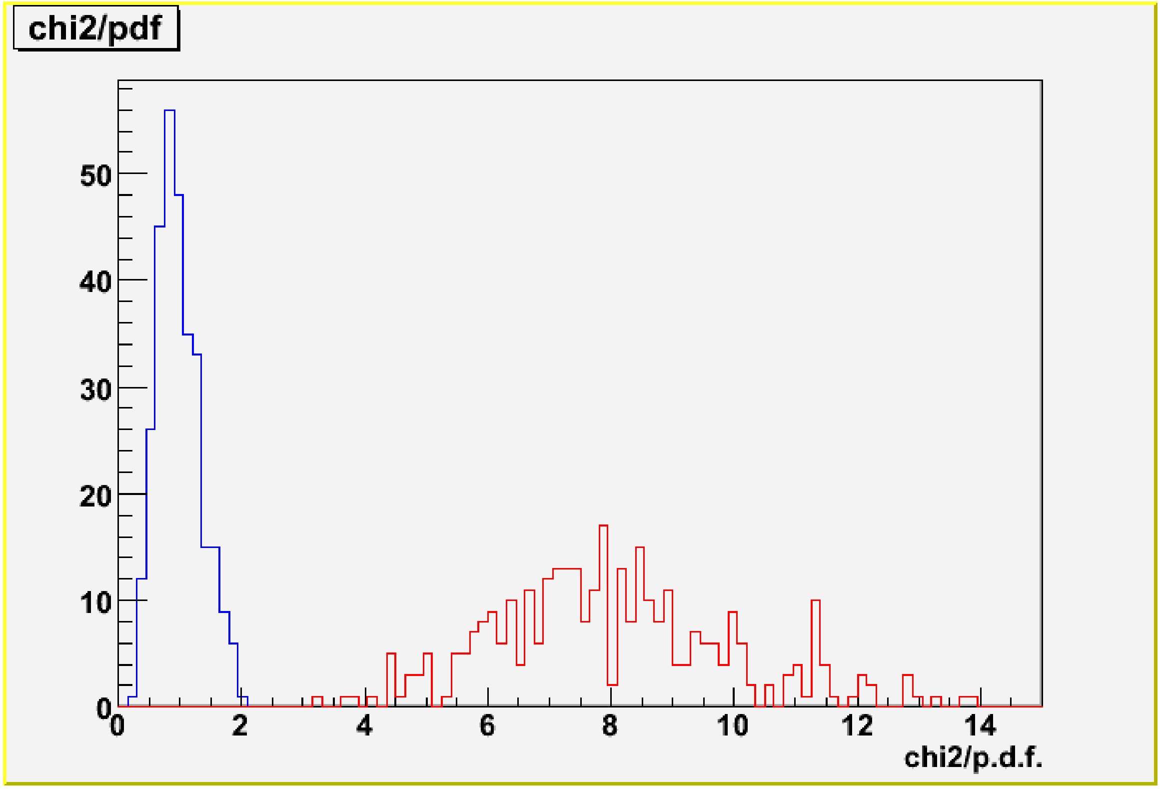
$\mathrm{sgn}(\mu) = -1$ A_0 m_t



mSUGRA Minima

mSUGRA:
Correct solution vs. negative μ solution

Experimental results smeared by random number distributed as Gaussian around central value



[plot by D Zerwas 2006]

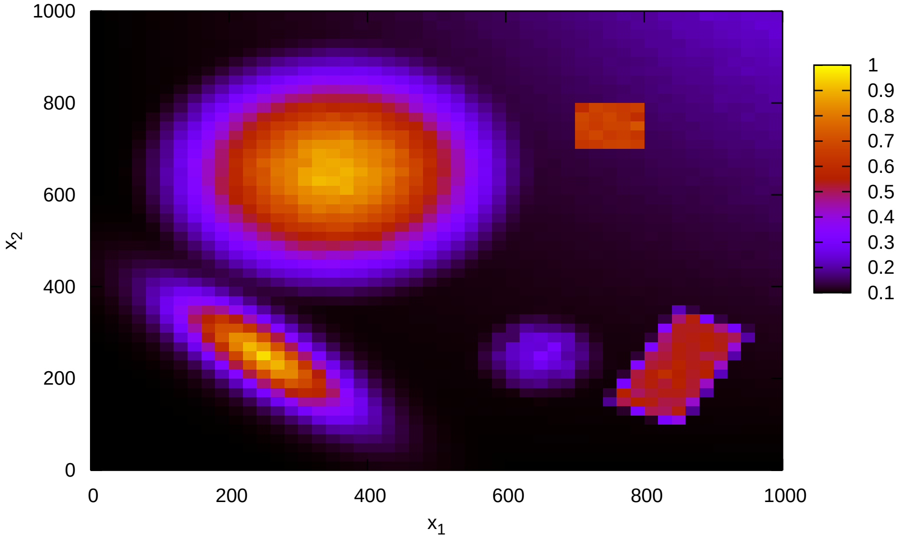
Experimental Input (edges)

(Obs)	= (meas)	± (exp)	± (theo)	
m_{h^0}	$= 109.53 \pm$	$0.25 \pm$	2.0	
m_t	$= 171.4 \pm$	$1.0 \pm$	0.0	
$\Delta m_{\tilde{\mu}_L,\chi_1^0}$	$= 106.26 \pm$	$1.6 \pm$	0.1	
$\Delta m_{\tilde{g},\chi_1^0}$	$= 509.96 \pm$	$2.3 \pm$	6.0	
$\Delta m_{\tilde{c}_R,\chi_1^0}$	$= 450.52 \pm$	$10.0 \pm$	4.2	
$\Delta m_{\tilde{g},\tilde{b}_1}$	$= 98.971 \pm$	$1.5 \pm$	1.0	
$\Delta m_{\tilde{g},\tilde{b}_2}$	$= 64.016 \pm$	$2.5 \pm$	0.7	
Edge(χ_2^0,$\tilde{\mu}_R$,χ_1^0)	$= 79.757 \pm$	$0.03 \pm$	0.08	(m_{ll}^{max})
Edge($\tilde{c}_L$,χ_2^0,χ_1^0)	$= 446.44 \pm$	$1.4 \pm$	4.3	(m_{llq}^{max})
Edge($\tilde{c}_L$,χ_2^0,$\tilde{\mu}_R$)	$= 316.51 \pm$	$0.9 \pm$	3.0	(m_{lq}^{low})
Edge($\tilde{c}_L$,χ_2^0,$\tilde{\mu}_R$,χ_1^0)	$= 392.8 \pm$	$1.0 \pm$	3.8	(m_{lq}^{high})
Edge(χ_4^0,$\tilde{\mu}_R$,χ_1^0)	$= 257.41 \pm$	$2.3 \pm$	0.3	$(m_{ll}^{\text{max}}(\chi_4^0))$
Edge(χ_4^0,$\tilde{\tau}_L$,χ_1^0)	$= 82.993 \pm$	$5.0 \pm$	0.8	$(m_{\tau\tau}^{\text{max}})$
Threshold($\tilde{c}_L$,χ_2^0,$\tilde{\mu}_R$,χ_1^0)	$= 211.95 \pm$	$1.6 \pm$	2.0	(m_{llq}^{min})
Threshold($\tilde{b}_1$,χ_2^0,$\tilde{\mu}_R$,χ_1^0)	$= 211.95 \pm$	$1.6 \pm$	2.0	(m_{llb}^{min})

Example

Test function (5-dim):

- Small Hypersphere $r = 100$, $V_{\text{max}} = 75$ @ $(650, 250, 350, 350, 350)$
- Cuboid $d = (173, 120, 200, 200, 200)$, $V_{\text{max}} = 60$ @ $(850, 225, 650, 650, 650)$
- Cube $d = (100, 100, 300, 300, 300)$, $V_{\text{max}} = 25$ @ $(750, 750, 450, 450, 450)$
- Gaussian $\sigma = (50, 150, 150, 150, 150)$, $V_{\text{max}} = 16$ @ $(250, 250, 550, 550, 550)$
- Big Hypersphere $r = 300$, $V_{\text{max}} = 12$ @ $(350, 650, 650, 650, 650)$
- Background $V = 0.1 + 4 \cdot 10^{-30} \cdot x_1^2 x_2^2 x_3^2 x_4^2 x_5^2$



1. V=74.929@(655.00,253.72,347.83,348.57,349.59)
2. V=59.972@(850.04,224.99,650.00,649.99,654.56)
3. V=58.219@(849.97,225.01,587.08,650.01,650.02)
4. V=25.110@(750.00,749.99,450.00,450.01,450.01)
5. V=16.042@(245.45,253.44,552.51,542.58,544.75)
6. V=12.116@(350.70,650.40,650.36,650.40,650.38)
7. ...

Plot Details

- Parameters: $x_1, \ldots, x_5 \in [0, 1000]$
- Bins: 50×50
- PDF: Breit-Wigner ($\frac{1}{1+\Delta x_i^2/\sigma^2}$) with $\sigma = 100$
- Number of Markov chains: 9
- Number of points per chain: 10^7
- Number of function evaluations: $33, 797, 153$
- Acceptance ratio: 0.19
- Final r (measure of convergence): 1.815
- CPU time (3 GHz): 150 min