Supersymmetric Parameter Determination at the LHC using Neural Networks

Nicki Bornhauser
In Collaboration with Manuel Drees

May 21st, 2013
Planck Conference Bonn

- The Large Hadron Collider (LHC) is working quite well. So far around $5\,\mathrm{fb}^{-1}$ of delivered data from proton–proton collisions for a center of mass energy of $7\,\mathrm{TeV}$ and $23\,\mathrm{fb}^{-1}$ for $8\,\mathrm{TeV}$
- Soon we may see signs of new physics. But even with the knowledge of the underlying theory the model parameter values would not automatically be known

→ In most new physics theories the relation mapping the measured observables onto the model parameters is unknown

- Use artificial neural networks to find this unknown relation
- In the following, as an example the constrained supersymmetric model CMSSM (mSUGRA) is examined to demonstrate the ability of neural networks for parameter determination

- Look at four different reference regions of the CMSSM at the LHC with a center of mass energy of 14 TeV
- Generally neural networks can also be used for any other model as long as the observables are chosen appropriately

- We look at 84 observables for the events after cuts
- Total number of events as well as 12 lepton classes with each 7 observables (minus one double information)

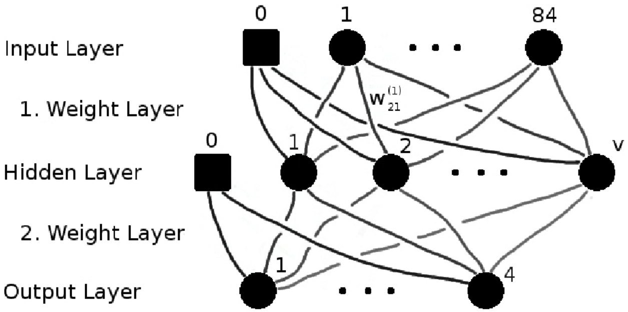
- Lepton classes: 0ℓ $1\ell^-$ $1\ell^+$ $2\ell^-$ $2\ell^+$ $\ell_i^+\ell_i^-$ $\ell_i^+\ell_{j;\ j\neq i}^-$ $\ell_i^-\ell_j^-\ell_j^+$ $\ell_i^+\ell_j^+\ell_j^-$ $\ell_i^-\ell_j^-\ell_{k;\ k\neq j,i\,\mathrm{for}\,+}^\pm$ $\ell_i^+\ell_j^+\ell_{k;\ k\neq j,i\,\mathrm{for}\,-}^\pm$ $4\ell+$
- Observables: n/N $\langle\tau^-\rangle$ $\langle\tau^+\rangle$ $\langle b\rangle$ $\langle j\rangle$ $\langle j^2\rangle$ $\langle H_T\rangle$

$n =$ number of class events $N =$ total number of events

- The usefulness of the observables for supersymmetric parameter determination was checked
- 283 parameter set pairs for a MSSM with 15 parameters, which were claimed to be indistinguishable within a LHC experiment (using 1808 mostly kinematical observables),[a] were reconsidered
- 260 out of these 283 pairs can be distinguished with a 95 % confidence level including systematic errors.[b] Without systematic errors even all pairs can be discriminated

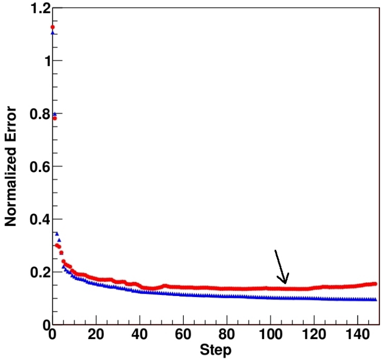
[a] N. Arkani–Hamed *et al.*, *JHEP* **0608** (2006) 070 [hep–ph/0512190]

[b] N. Bornhauser, M. Drees, *Phys. Rev.* **D86** (2012) 015025 [hep–ph/1205.6080]



- Neural network with two weight layers for 84 input values (observables) and 4 output values (CMSSM parameters)

- Neural network is trained to learn the mapping between the observables and the CMSSM parameters within a considered parameter region
- Training sets of known input and output values are used
- In each learning step the neural network weights are changed appropriately to better reproduce these training sets
- Introduce control sets to determine the "optimal" weights



- Training (blue) and control error (red) evolution for a neural network

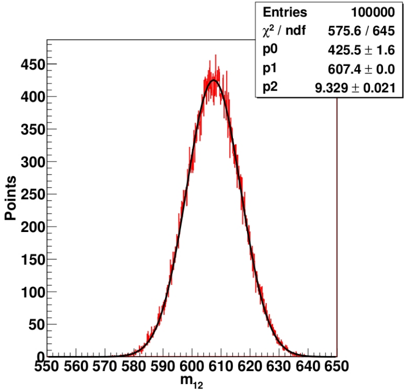
- Observables have statistical uncertainties

→ Considering the variances and covariances of the observables improves the neural network performance

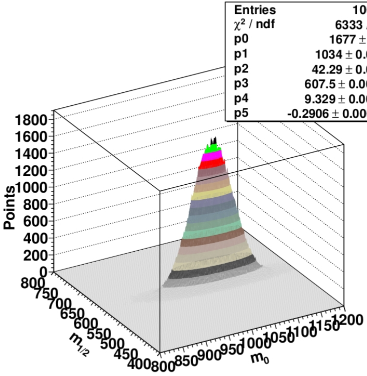
- Increase the luminosity of the training and control sets
- Require minimal event numbers for the observables
- Create (correlated) Gaussian distributed copies of each training set → network is confronted with observable uncertainties
- Furthermore, easier network specialization by creating one network for each CMSSM parameter

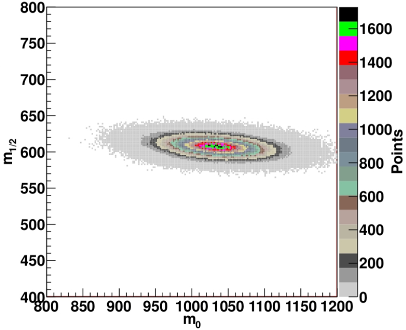
- So far, the neural networks would calculate the CMSSM parameters for a given measurement
- But the errors and correlations of the parameters are not given

- There are two different ways to calculate them using the variances and covariances of the measured observables:
  - Propagation of uncertainty
  - Feed networks with Gaussian distributed measurement copies



- One–dimensional output distribution for Gaussian distributed measurement copies





- Two–dimensional output distribution for Gaussian distributed measurement copies

- Four reference points with each around 1,000 events after cuts for an integrated luminosity of $10\,\mathrm{fb}^{-1}$
- Training and control sets are each chosen from parameter ranges around these points

- Events are generated with Herwig++[a]
- Furthermore use SOFTSUSY[b], SUSY–HIT[c], and FastJet[d]
- The events have to pass certain cuts to reduce Standard Model background

---

[a] M. Bähr *et al.*, Eur. Phys. J. C **58** (2008) 639 [arXiv:hep–ph/0803.0883]

[b] B. C. Allanach, Comput. Phys. Commun. **143** (2002) 305 [arXiv:hep–ph/0104145]

[c] A. Djouadi *et al.*, Acta Phys. Polon. B **38** (2007) 635 [arXiv:hep–ph/0609292]

[d] M. Cacciari, G. P. Salam, Phys. Lett. B **641** (2006) 57 [arXiv:hep–ph/0512210]

- Results for $10\,\mathrm{fb}^{-1}$ (top) and $500\,\mathrm{fb}^{-1}$ (bottom) with searched values in parenthesis

	Point 1	Point 2
$\overline{m}_0 \pm \sigma_{m_0}$	171.47 ± 34.53 (150)	1998.93 ± 92.20 (2000)
$\overline{m}_{1/2} \pm \sigma_{m_{1/2}}$	702.72 ± 11.07 (700)	446.55 ± 11.30 (450)
$\overline{\tan\beta} \pm \sigma_{\tan\beta}$	21.35 ± 5.96 (10)	15.91 ± 17.69 (10)
$\overline{A}_0 \pm \sigma_{A_0}$	463.43 ± 326.00 (0)	1406.37 ± 2898.67 (0)

$\overline{m}_0 \pm \sigma_{m_0}$	156.40 ± 4.88 (150)	2004.05 ± 10.61 (2000)
$\overline{m}_{1/2} \pm \sigma_{m_{1/2}}$	701.59 ± 0.89 (700)	451.15 ± 1.00 (450)
$\overline{\tan\beta} \pm \sigma_{\tan\beta}$	9.39 ± 0.70 (10)	15.12 ± 2.09 (10)
$\overline{A}_0 \pm \sigma_{A_0}$	-43.61 ± 55.37 (0)	261.47 ± 474.68 (0)

- Results for $10\,\mathrm{fb}^{-1}$ (top) and $500\,\mathrm{fb}^{-1}$ (bottom) with searched values in parenthesis

	Point 3	Point 4
$\overline{m}_0 \pm \sigma_{m_0}$	1055.97 ± 47.26 (1000)	482.94 ± 61.23 (400)
$\overline{m}_{1/2} \pm \sigma_{m_{1/2}}$	607.45 ± 11.53 (600)	695.48 ± 7.88 (700)
$\overline{\tan\beta} \pm \sigma_{\tan\beta}$	23.41 ± 37.42 (10)	26.00 ± 10.66 (30)
$\overline{A}_0 \pm \sigma_{A_0}$	1861.38 ± 1181.25 (1500)	-73.52 ± 628.16 (0)
$\overline{m}_0 \pm \sigma_{m_0}$	1015.66 ± 4.49 (1000)	391.86 ± 7.70 (400)
$\overline{m}_{1/2} \pm \sigma_{m_{1/2}}$	598.75 ± 1.21 (600)	700.71 ± 0.88 (700)
$\overline{\tan\beta} \pm \sigma_{\tan\beta}$	20.79 ± 5.20 (10)	30.54 ± 1.23 (30)
$\overline{A}_0 \pm \sigma_{A_0}$	1033.4 ± 314.44 (1500)	183.05 ± 101.48 (0)

- The knowledge of the variances and covariances of the measured observables is crucial for a successful parameter determination
- For around 1,000 events after cuts, the CMSSM parameters m_0 and $m_{1/2}$ can be determined reliably, with errors as small as 1 %
- For around 50,000 events after cuts, also the parameters $\tan\beta$ and A_0 can be determined quite accurately
- Overall, neural networks give very reliable results. With the right set of observables they can be used for an arbitrary model

Thank you for your attention!